DUET: Distinct but United Event-based Timelines

Andy Luu

Thesis submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Master's of Science
in
Computer Science and Applications

Sang Won Lee, Chair
Sol Lim
Yan Chen

December 12th, 2024
Blacksburg, Virginia

Keywords: timeline, visualization, event-based, usability, data-processing

DUET: Distinct but United Event-based Timelines

Andy Luu

(ABSTRACT)

Through innovations in sensors and storage, event-based data collection has increased significantly in recent years. Personal security cameras allow users to collect event data in their homes, Google Maps allows users to view places they have been throughout the day, and web browsers collect a history of the sites users visit. Despite the innovations in event-based data collection, the Graphical User Interfaces (GUIs) designed for these systems have not changed for years. Many applications stick to a linear timeline to display the events, leading to large downtime gaps when sparse events occur throughout the day. However, linear timelines create a natural sense of temporal perception since users can easily tell what time of day certain events occur. Our project seeks to find a balance between the natural temporal perception that linear timelines offer and improving the usability of visualizing event-based data using timelines. Through our design, users will be able to leverage the simple temporal perception that a linear timeline provides and avoid unnecessary searching to locate specific events. By seamlessly integrating a linear timeline with density visualizations and a timeline of events, we aim to allow users to search and view events easily. To evaluate our interface, we conducted two usability studies, first with specialists who view security camera information often and then with all individuals. Our findings inform the design of future event-based timeline visualizations.

DUET: Distinct but United Event-based Timelines

Andy Luu

(GENERAL AUDIENCE ABSTRACT)

Through innovations in sensors and storage, event-based data collection has increased significantly in recent years. Personal security cameras allow users to collect event data in their homes, Google Maps allows users to view places they have been throughout the day, and web browsers collect a history of the sites users visit. Despite the innovations in event-based data collection, the Graphical User Interfaces (GUIs) designed for these systems have not changed for years. Many applications stick to a linear timeline to display the events, leading to large downtime gaps when sparse events occur throughout the day. However, linear timelines create a natural sense of temporal perception since users can easily tell what time of day certain events occur. Our project seeks to find a balance between the natural temporal perception that linear timelines offer and improving the usability of visualizing event-based data using timelines. Through our design, users will be able to leverage the simple temporal perception that a linear timeline provides and avoid unnecessary searching to locate specific events. By seamlessly integrating a linear timeline with density visualizations and a timeline of events, we aim to allow users to search and view events easily. To evaluate our interface, we conducted two usability studies, first with specialists who view security camera information often and then with all individuals. Our findings inform the design of future event-based timeline visualizations.

Dedication

This is dedicated to my mother and father who raised me, fed me, helped me grow as a person, and sacrificed so much for me to be able to grow up in the United States and be the first of our family to earn a college degree. This is dedicated to my aunt, uncle, and grandmother from my father's side who helped my mother and father care for my siblings and I when my parents were working long hours at their jobs. This is dedicated to my aunts, uncles, grandmother, grandfather, and granduncle from my mother's side who always welcomed me into your homes, restaurants, and stores when I visited and fed me well. This is dedicated to my brother Denny and sister Wennie who no matter how different we all are, I know will always be there for me when I need their support. This is dedicated to my cousins Jacky, Jenny, Jammy, and Kelly who I do not see as often as I would like, but I always treasure the time I get to spend with them. This is dedicated to all of my friends who I have met during my time at Virginia Tech, thank you for making my college years some of the best years of my life.

Acknowledgments

I would like to acknowledge Dr. Sang Won Lee, who I am forever grateful for introducing me to the world of research, and who has been a mentor to me since my sophomore year at Virginia Tech. I would like to acknowledge my committee members Dr. Sol Lim and Dr. Yan Chen who without them, I would not have been able to complete this research. I would like to acknowledge former members Amber, Tausif, Donghan, Momen, Muskan, Emily, Daniel V., Daniel E., Drew, Yi, and current members Jelson, Danny, Carlos, Jaehoon, Xinyi, Adnan, and many more members of the EchoLab who without all of you, I would have never made so many lasting memories during my time working on my research.

Contents

Chapter 1

Introduction

1.1 Background

As humans continue to use technology more and more in their daily lives, collecting large amounts of event-based data has become easier than ever. Event-based data, or information related to specific events, including when, why, and how they occurred, can be collected digitally as people use their devices or through a combination of sensors and computer vision in different environments [25]. For example, as people browse the web daily, their web browsers keep track of every website that they visit with a short timestamp so that the user remembers when they last accessed that link. Google Maps recently implemented a similar feature for users that allows the application to track their location throughout the day continuously, storing the paths that users take and the places they visit [17]. These types of event-based data can easily be used to identify patterns and trends in users' lives and how they can improve certain aspects of their lives [25]. However, despite the many improvements that event-based data bring to our lives, the methods that have been used to display these events to users have stagnated in their improvements for many years. Currently, most applications display these events through a linear horizontal timeline. For example, Ring and Feit Electric security cameras use an infinitely scrollable container where each hour is contained by a certain number of pixels, and then events are placed in those containers as they occur throughout the day [12]. Although the benefits of this linear timeline approach

are clear, since it offers great temporal perception for events since users can easily tell when events occur throughout the day, the horizontal timeline visualizations are not user-friendly when there are few events. If there are only a couple of events that occur throughout the day, users may have to scroll through the entire timeline to find when those events occur. Additionally, there is also an overwhelming visualization when there are too many events. Other applications have attempted to use different methods of displaying events like browser search histories that display them in a vertical list without the gaps in between but users have limited information to temporally perceive event occurrences when presented with this format since the vertical lists do not display the duration or much visual information about the events.

1.2 Motivation

Our project seeks to find a balance between the natural temporal perception that linear timelines offer and improving the usability of visualizing event-based data using timelines. We test this balance through the DUET (Distinct but United Event-based Timelines) project, an interface that we have developed that allows users to utilize multiple different timeline visualizations to search for certain events in a video. Currently, the website supports multiple security camera videos from the publicly available VIRAT Video dataset with twelve different types of events that occur throughout the videos [21]. The website also has five different timeline visualizations for users to use to find specific events inside the security camera videos and allows users to filter by the different types of events so they can narrow their search to the types that they are looking for.

1.3 Research Questions

Through the interface that we created, we are interested in learning more about:

1. What difficulties do users face when searching for specific events in long timelines with sparse events?

2. What tools or features can be implemented into existing navigation methods to ease searching for specific events in long timelines with sparse events?

To gather preliminary thoughts about these questions and improve our interface before conducting a study with the general public, we conducted a pilot study (N=5), where we interviewed various types of people who review security camera footage. Specifically, we interviewed Police Officers, a Homeowner, an Undergraduate Researcher working on a farm, and a Government Office Security Specialist. Through the initial study, we were able to gather preliminary feedback about the initial interface that we developed. From the feedback that we gathered, we implemented improvements to the interface such as adding an adjustable playback rate for the videos and the ability to zoom into the event thumbnails. After implementing the improvements we conducted a more general user study (N=72), where we recruited individuals to test out the timeline visualizations by asking them to complete tasks using each of the timeline visualizations. During the user study, we gave participants two tasks to complete per video using a different timeline visualization for each video and then a short exit questionnaire based on their thoughts about the timeline visualization. From the user study, we found significant differences between the System Usability Scale scores and how the different timeline visualizations affected the participant's ability to find specific events when answering questions. From the results of our studies, we believe that the timeline visualizations that we created can be used to inform the design of future

event-based timeline visualizations for other domains.

Chapter 2

Review of Literature

2.1 Use Cases for Event-based Data

Many researchers in the scientific community have discovered the enormous potential of exploiting the amount of available event-based data due to the increase in the types of sensors that capture different environmental conditions. For example, in a survey of existing literature focused on the processing and analysis of event-based data, researchers found event-based data used in domains such as digital preservation, journalism, social media, and more [27].

2.1.1 Video Systems

One especially relevant domain to our study was how event-based data was used in video systems. One example of previous research proposes an ontology to represent prior knowledge related to the analysis of event-based data in videos. The proposed ontology was composed of two parts, the basic concepts and the domain-specific extensions. To demonstrate the usability and effectiveness of their proposed ontology, the researchers created a specialized version for underground video surveillance systems [24]. Another example of previous research focuses mainly on how to use attribute grammar to detect anomalies in video footage where the researchers found that in confined environments, it was possible to distinguish

most anomalies that occur. [26]. Aside from the previous examples, a different set of researchers focused mainly on proposing a procedure for constructing time-space diagrams to visualize the progression and evaluation of the coordination of traffic signals at arterials using event-based data. The findings of this study show that time-space diagrams can be constructed using automatically collected high-resolution event-based traffic data [30]. Lastly, to showcase other examples outside of video surveillance systems, one study used event-based data to propose a general framework for event detection and summary generation in broadcast sports videos. From their findings, they proposed different approaches for the detection of the plays in baseball, football, and sumo wrestling, one being rule-based inference, and another using HMM for high-level inference [15].

2.1.2 Audio Systems

Similarly to video surveillance systems, multiple studies have shown that event-based data can also be used for audio purposes by using event-based data to detect and classify different audio events. One example of previous research focused mainly on detecting audio events in noisy environments aiming to create a method to detect abnormal situations using audio clues instead of visual cues. To detect these abnormal audio cues, the researchers created a system that segmented audio segments into shot and normal classes using feature extraction with a false rejection rate of under ten percent [5]. Additionally, another study expanded its scope to audio events that occur in our daily environment, specifically in a traditional office setting by developing an audio-based surveillance system that can monitor any environment and classify the information that it collects in a meaningful fashion. The researchers concluded that an audio-based surveillance system that combines unsupervised and supervised clustering methods would be effective once relevant events were identified [8].

2.1.3 Life-logging Systems

Another especially relevant domain was the way event-based data was used to record the lives of people and animals. One example of previous research focused on fostering the concept of life-logging as a mainstream activity in everyday life. The researchers saw that with the decreasing costs of computer storage and the advancements of sensing technology, tracking life activities has become easier than ever which could help understand human performance in a variety of tasks. Thus the researchers created a comprehensive summary of life logging, including its research history, current technologies and applications, and its limitations [7]. Another set of researchers focused on analyzing life logs to support creating applications using an individual's life logs. The researchers demonstrated a method to help analyze life logs by automatically augmenting descriptions to life log events by using semantic web technologies. From their experiments, the researchers found that their method created vivid descriptions of life log events, opening greater possibilities for using life logs in the future [19]. Some researchers then drew inspiration from their peers and studied using life logging as a method to aid people with dementia in maintaining their identities. The researchers utilized SenseCam images to create meaningful discussions about the user's recent memories to construct a version of the user's identity during therapy sessions to explore the potential of life-logging technology in supporting health care professionals [22]. Lastly, aside from research focused on humans, one set of researchers focused on using life logs to detect wildlife morbidity events to reduce threats to wildlife populations. The researchers developed the Wildlife Morbidity and Mortality Event Alert System which integrates pre-diagnostic clinical data in near real-time from a network of wildlife rehabilitation organizations to detect early signs of unusual wildlife morbidity and mortality events. While demonstrating their system, the researchers were able to alert events associated with both common and emerging diseases in various wildlife populations [13]. From these existing works, it is clear that there are use

cases for processing event-based data across multiple different domains.

2.2 Augmenting Existing Navigation Interfaces

As previously mentioned, linear timelines are one of the most popular methods for navigating through events due to the natural sense of temporal perception that they offer. Many applications take advantage of this benefit whether that be in the form of a progress slider in a video or a recollection of significant events in history. However, the traditional horizontal slider is no longer effective enough for users hence why researchers in some domains have attempted to augment current interfaces with new visualizations. One example of previous research focuses mainly on improving the navigation of educational videos using data generated from user interactions. The researchers used the interaction data to implement LectureScape which included a 2D timeline with a density graph based on user interactions, a method to search for keywords with an interactive transcript, and video highlights using storyboards, a word cloud, and bookmarks. From their findings, the researchers saw that their augmentations drew attention to important and confusing points in the videos and allowed users to have better control and flexibility during navigation [14]. Another example of a 2D timeline interface can be seen with BrowseLine, which focuses on a timeline visualization for users to identify temporal patterns from their browsing history. The researchers believed that by using the other web browsing the user was doing at a certain time, a user may be able to recall when they visited the page they were looking for. To test their hypothesis, the researchers implemented a 2D timeline with rows for each hour and a horizontal list of which websites were visited in that hour grouped by domains. They also implemented domain filtering and thumbnail and metadata previews to further help users recall their past search history. The findings of this study show the temporal data from traditional browsing

histories can be leveraged through visual pattern recognition to enhance a user's ability to recall past browsing history [9]. Aside from web browsing history, some researchers have focused on users' personal history through LifeLines, a visualization software for searching through a person's biological data. LifeLines presents a youth record on a single screen, providing direct access to all details of certain information, and promoting critical alerts on the overview level. The findings of this study emphasize the importance of creating and iterating on visualizations of personal data, especially life events that happen every day [18]. From these works, it is clear that traditional navigation methods, although effective, benefit significantly in usability from new augmentations.

2.3 Personal Security Systems

As previously discussed, methods of displaying event-based data, especially traditional video surveillance systems, have not been improved for years, and neither have methods to create affordable personal security systems. One set of researchers from the NRI Institute of Information Technology aimed to study different methods to address current issues with video surveillance systems. Due to a lack of an existing standard, they decided to create a review of requirements, current approaches, and explanations of concepts to new users. The requirements were split into functional requirements and non-functional requirements. As for current approaches, the researchers saw that a computer-based approach provided high stability, great compatibility with cameras, simultaneous operation, and more. Lastly, the researchers explained concepts for new users such as how motion detection, object tracking, and human activity recognition work in these systems [11]. Other researchers also noticed the popularity of video surveillance systems but saw that manufacturers each had unique protocols so they sought to create low-cost software for integrating those systems. The

researchers combined a low-range server for video streaming, a Raspberry PI for motion notifications, and a cellphone application for remote video access and SMS notifications. The findings of their study introduce a low-cost, scalable video surveillance system using a Raspberry Pi which can be augmented with other modules to become a complete security system in the future [1]. Another example of previous research also focused on a low-cost system but sought to improve its usability with some new features instead. The researchers began with a TP-LINK model NC220 camera which had WiFi connectivity, night vision, motion and sound detection, alerts, and a mobile application for live streaming. To improve the current architecture, the researchers used Telegram to allow users to activate and deactivate alerts and notify users of motion alerts. The researchers evaluated their proposed system against the original system using the SUS score with the results heavily favoring the proposed system. The findings of this study introduce an improved low-cost video surveillance system that enhances the current system's alert and notification capabilities [2]. It is clear from these works that there is value in implementing new software and designs for current personal video surveillance systems.

Chapter 3

Implementation

After reviewing the previous literature, we were still left seeking an improved understanding of three aspects. The first aspect was what augmentations users want in timeline visualizations to aid with searching for events. The second aspect was the users' difficulties when searching for events using existing timeline visualizations. The third aspect was about accuracy since we wanted to know more about whether adding different augmentations led users to find the correct events when searching. From these gaps in knowledge, we began to start thinking about what to include when implementing the DUET website. When we initially started implementing the website, we looked into different existing systems to gather features that our timeline visualizations should have and realized that we wanted to ensure that we had a visualization that offered a great sense of temporal perception, a visualization that was space efficient when searching for events, and a visualization that provided visual information about the events.

3.1 Base Interface

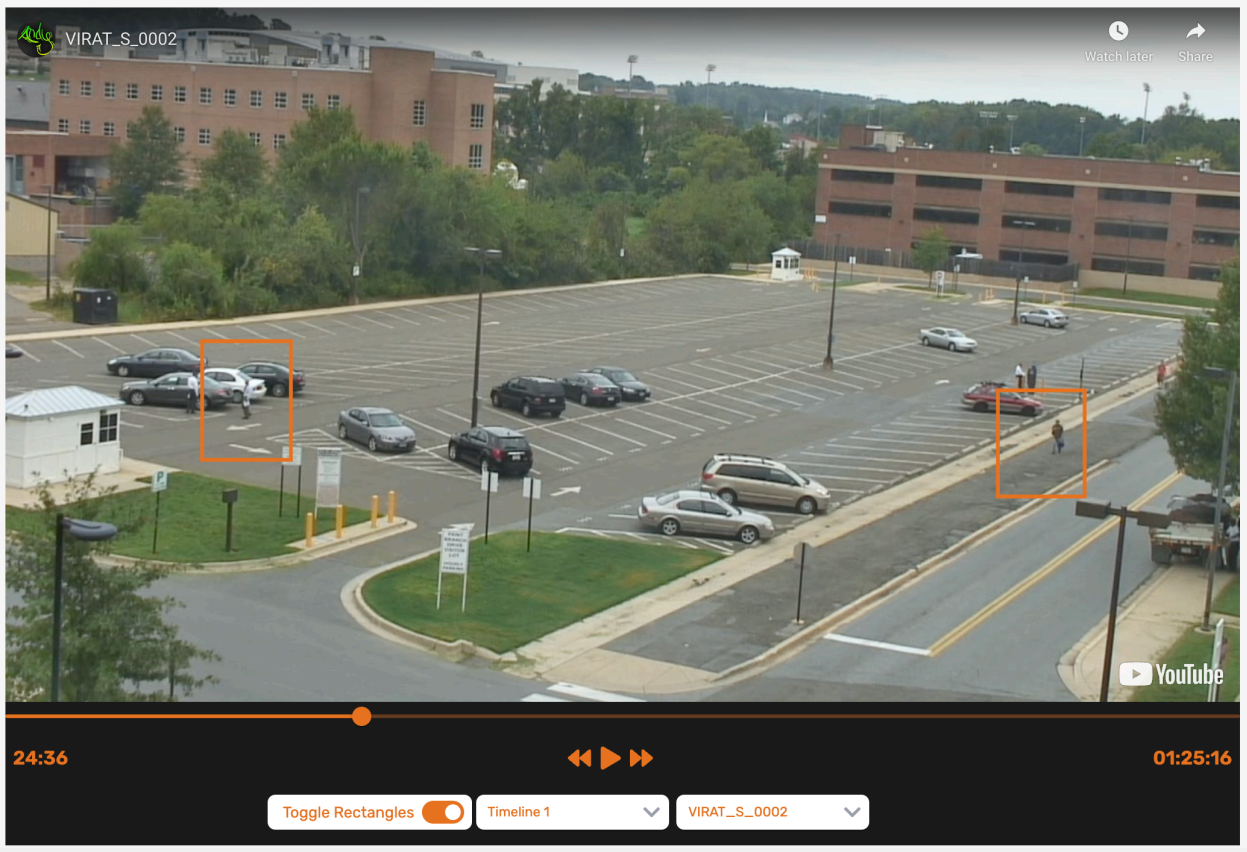


Figure 3.1: Base Interface

When you first launch the website, you are greeted with a large view of a single security camera video provided by the VIRAT Video Dataset [21]. Underneath the video player is a standard linear timeline with the elapsed time of the video, basic video controls, and the total duration of the video similar to existing video players like YouTube [29]. After the linear timeline with video controls, the next feature is a toggle for showing or hiding the bounding rectangles for events. We implemented this toggle since some events may have information that the bounding rectangles may cover but the user wants to see. The last initial features are dropdown menus for selecting a video from a list of available videos and selecting the

timeline type. The base interface that we designed for DUET is shown in Figure 3.1.

3.2 Density Graph

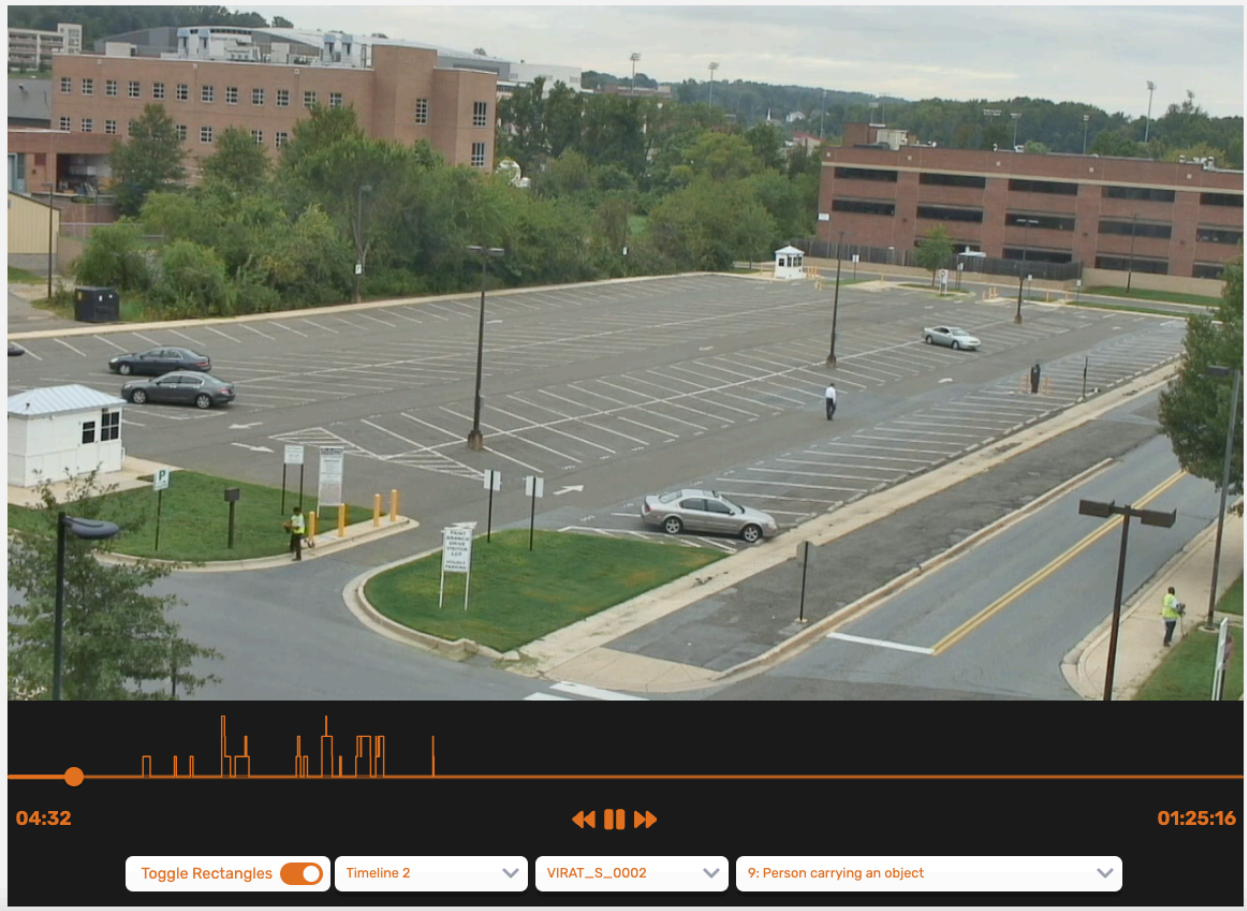


Figure 3.2: Density Graph

After implementing the initial linear timeline, we created some new features that we believed would improve how users navigate through the long videos with sparse events. The first new feature we implemented was a density graph above the existing linear timeline in the base interface, similar to other video navigation platforms like YouTube [14, 29]. To create the density graph, we parsed through all of the events in the current video, incremented the

graph at the starting point of each event, and decremented the graph after the stopping point of each of the events. The main purpose of the density graph was to create a 2D timeline to inform the user temporally when events are occurring in the video [4, 20, 28]. After we created the density graph which showed all the events in the video, we believed that a user should have the ability to filter through the different event types in the video. To implement filtering based on event types, we added a third dropdown menu with all of the different event types, allowing users to select one of the event types in the video. After the user selects one of the event types, the 2D timeline is repopulated with just the events of that event type in the video. To separate the density graph from the base interface, we currently refer to the base interface with the density graph as Timeline 2 which is shown in Figure 3.2.

3.3 Event Blocks

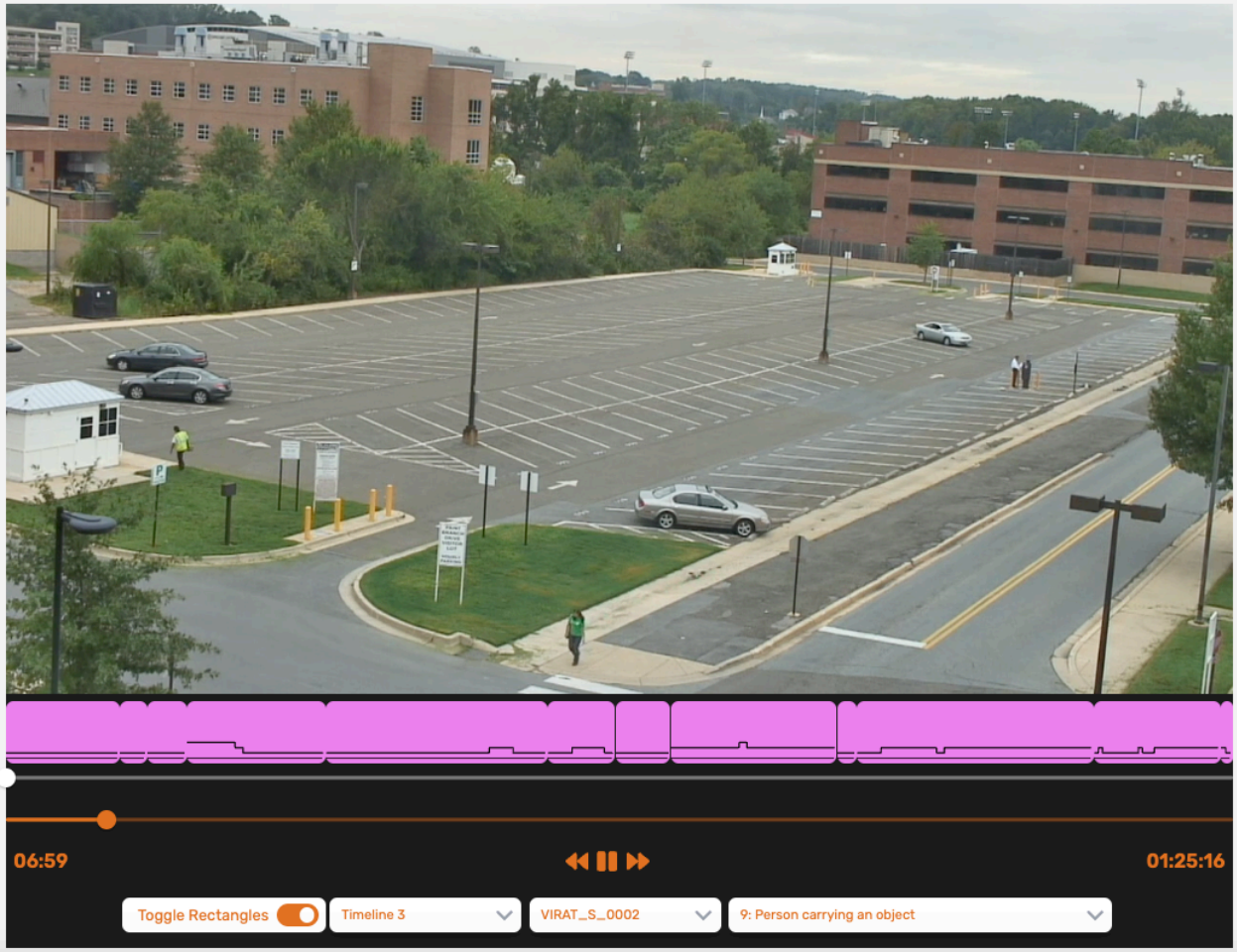


Figure 3.3: Event Blocks

After implementing the density graph and filtering by event type, we wanted to implement a timeline that was space-efficient for a user which led to creating a linear timeline with event blocks [16]. In our implementation, an event block is a block generated from the timeline based on sections of the timeline where events occur during that block of time [4, 9, 28]. To generate the event blocks, we repurposed the density graph and took the points where events started to happen (the points where the density graph started to rise) and the points where events ended (the points where the density graph started to fall), and then grouped the events

between those points into one block. We also created a separate color for each event block based on their event type and allowed users to click on an event block to fast-forward the user to the first event that occurs in the event block [6]. We then placed all the event blocks horizontally above the linear timeline in the base interface with no periods of downtime in between them. After placing the event blocks, we synced the event blocks to a new linear timeline underneath the event blocks which would follow along with when an event occurred in the original timeline and allow the users to navigate between the different event blocks. Through the event blocks and new linear timeline, we created a 2D timeline where the user can navigate through the events without the large periods of downtime that the video may have, creating a space-efficient method of event navigation. To separate the event blocks from the base interface and density graph, we currently refer to the linear timeline with the event blocks as Timeline 3 which is shown in Figure 3.3.

3.4 Density Graph + Event Blocks

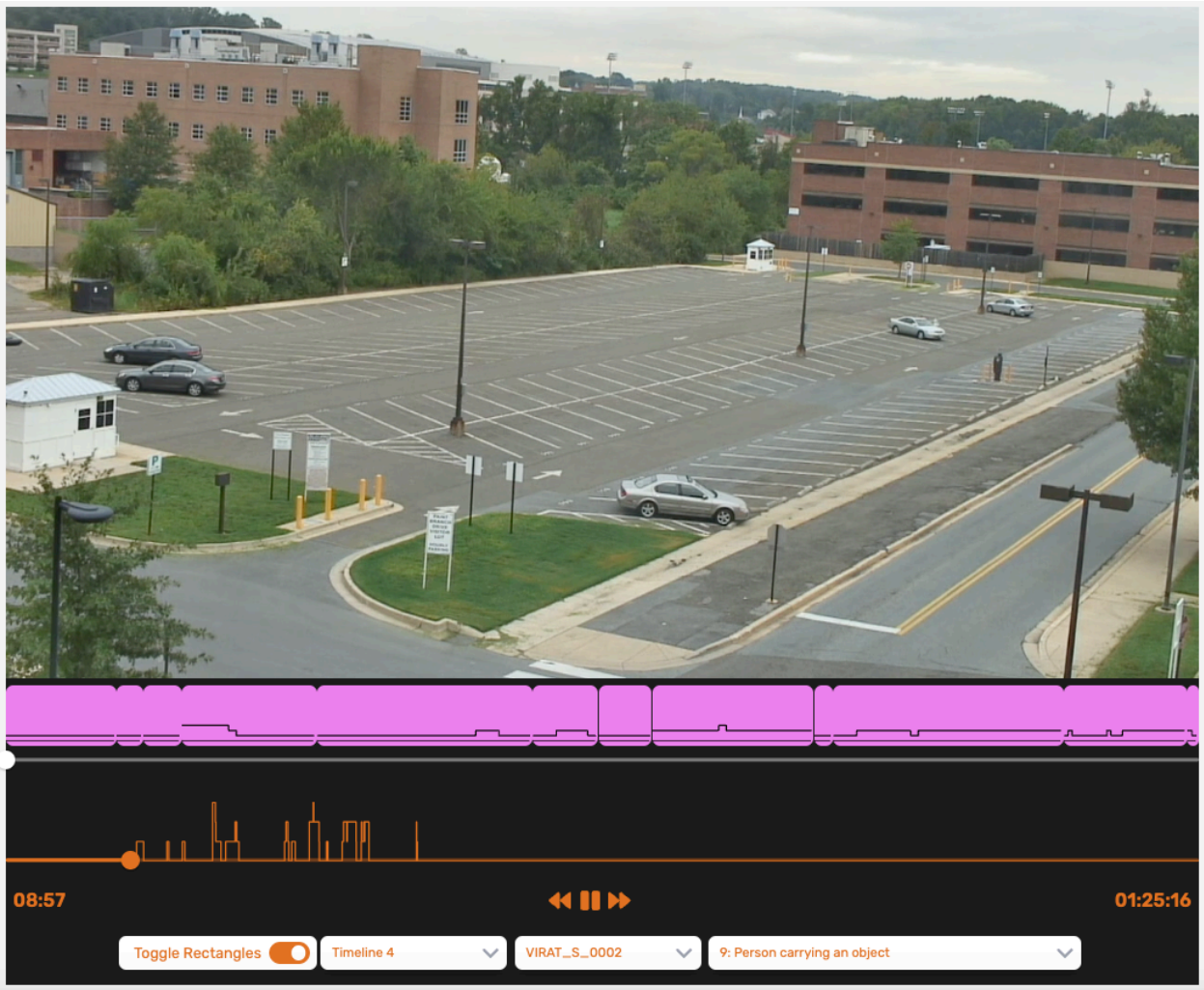


Figure 3.4: Density Graph + Event Blocks

Once we finished implementing the density graph above the linear timeline on the base interface and the new linear timeline with the event blocks, we created a combined version of the two features that placed the new linear timeline with the event blocks above the density graph. The main purpose of this combined version was for users to gain the benefits of both of the features, being able to see temporally when events occur in the timeline as well as have an improved method of navigating between those events. To separate this iteration

from the others, we currently refer to the timeline with the density graph and the event blocks as Timeline 4 which is shown in Figure 3.4.

3.5 Event Thumbnails

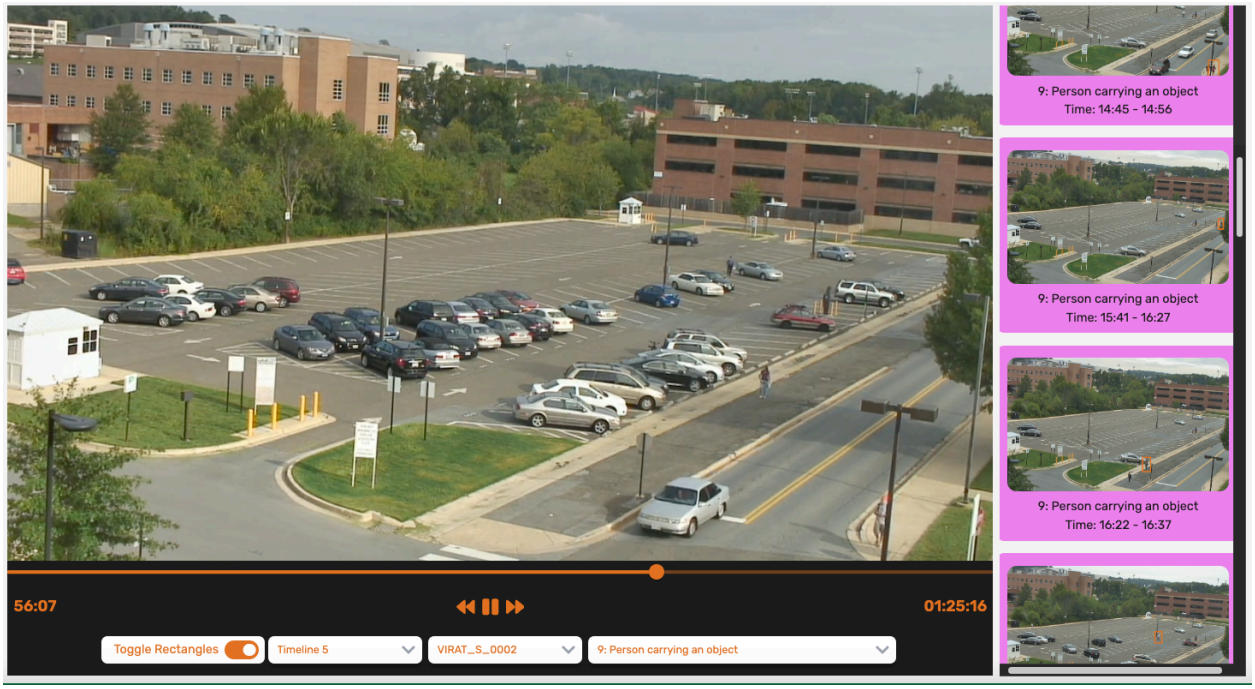


Figure 3.5: Event Thumbnails

Lastly, we decided to implement one more timeline visualization which shows a list of the events on the left side of the video player. Each of the events in the lists was displayed with a thumbnail of the first frame of the event, the event type, and the start time of that event [4, 9, 28]. We were able to generate the thumbnails using a provided MatLab script that the VIRAT ground dataset uses to draw the bounding rectangles onto each of the frames [21]. We also implemented a function where when a user clicks on one of the events, they are fast-forwarded to the start time of that event so they can watch that event happen. The main purpose of creating the event list with thumbnails was to provide users with a

visual method of navigation, similar to other video-watching platforms [29]. To separate this feature from the others, we currently refer to the linear timeline with a list of events with thumbnails as Timeline 5 which is shown in Figure 3.5.

Chapter 4

Pilot Study

4.1 Study Design

After implementing the timeline visualizations, we decided to conduct interviews with users who have occupations where they view security camera footage to discuss in depth about what difficulties they face with their current methods [23]. The interviews were conducted either in person in a conference room or remotely via Zoom according to the preference of the participant. Each interview took approximately an hour to complete and each participant was compensated with a $30 electronic gift card for their time. We recorded the entire interview through Zoom and then utilized its transcription feature to transcribe the audio from each interview. The audio for all of the interviews was stored in a password-protected, encrypted OneDrive folder managed by Virginia Tech. Our interviews were split into three sections.

4.1.1 Intake Interview

The first section mainly focused on expanding on each participant's responses to the sign-up survey. During the sign-up survey, we asked each participant this set of questions:

- What is your current role at the organization and what responsibilities do you have in this role?

- How often do you view the security camera footage as a part of your work?
  - Every Day
  - Every Week
  - Every Month
  - Every Year
  - Other
- What are the typical events or incidents are the ones that you look for when you view the footage?
- What information do you typically use and are given as clues to spot the events? (time window, profile, or nothing?)
- Could you share the most challenging and time-consuming tasks that you had to do when reviewing camera footage? How long does it usually take for you to review the camera footage to find what you are looking for?
- What software does your organization use to view the footage that you collect from the security cameras?
- If there is software you use from other companies, are there any features that they share in common?
- Does your software have any special features for identifying events (i.e., computer vision for event identification)? If so, how does that help you find more quickly? When is it not helpful?
- Are there any current issues that you have with the current software that your organization uses to search for events across long periods of footage?

Before the interview, we looked through each participant's responses to the questions asked in the sign-up survey and highlighted any interesting sections that we wanted to discuss further during the interview. After expanding on the participant's responses from the sign-up survey, we sent the participant a link to the DUET website with each of the timeline visualizations that we developed. During the tutorial, we showed them the base interface, how to switch between the different types of timelines and videos, and a brief overview of each of the timeline visualizations that we implemented.

4.1.2 Usability Study

After familiarizing the participants with the timeline visualizations we developed, we had each participant attempt to complete at least one of three types of tasks that we created to represent different real-world scenarios that security camera operators may face.

The first task involved understanding the sequence between two different events. First, we had each participant find when a white car drove into the parking lot from the back entrance of the parking lot. We also required the man exiting the white car to be wearing a white shirt to narrow down the possibilities. Next, we had the participant calculate how long until after the first event occurs, that a white car and a black car pull into the parking lot and people exit from both cars. We designed this task to mimic when a security camera operator has to find an event that happened after a previous event that occurred in the footage that a bystander remembered.

The second task involved counting the number of people that fit certain criteria. First, we had each participant switch to a video showing an intersection with a parking lot and stop sign where people waited for their friends to come and pick them up from. Then we asked them to focus on the stop sign and count how many people who were waiting at the stop

sign were carrying a suitcase. This task was designed to mimic when an operator has to review footage to find people who fit a given profile.

The third task involved a scenario where the person that the participant is trying to find does not exist in the given security camera footage they are reviewing. First, we had each participant find when a woman with a white shirt unloaded her purse from a gray car which was a real task that occurred in the video. Next, we asked each participant to focus on the white security office building next to the parking lot and find the shirt color of the man that exists for that building. However, there is not a man who ever exits from the white security office building. This task was designed to mimic when an operator can not find a matching event from a given description.

4.1.3 Exit Interview

After the three tasks, we asked each participant some further questions about our timeline visualizations to gauge their opinions on our system. Specifically, we asked:

- What was the timeline representation that you liked using the most?
  - Timeline 1 (Linear Timeline)
  - Timeline 2 (Linear Timeline + Density Graph)
  - Timeline 3 (Linear Timeline + Timeline of Event Blocks)
  - Timeline 4 (Linear Timeline + Density Graph + Timeline of Event Blocks)
  - Timeline 5 (Linear Timeline + Event Thumbnails)
- Please explain your answer to the previous question.
- What was the timeline representation that you liked using the least?

  - Timeline 1 (Linear Timeline)
  - Timeline 2 (Linear Timeline + Density Graph)
  - Timeline 3 (Linear Timeline + Timeline of Event Blocks)
  - Timeline 4 (Linear Timeline + Density Graph + Timeline of Event Blocks)
  - Timeline 5 (Linear Timeline + Event Thumbnails)

- Please explain your answer to the previous question.
- What are your overall opinions on our current system?
- Are there any other features you think would be nice to implement in our current system?
- Do you believe the tasks we had you perform were realistic to those you look for when you browse for events in your security camera footage?

After the exit questions, we answered any questions the participants asked, thanked them for their time, and let them know the timeline for compensation.

4.2 Results

During the pilot study interviews, we interviewed two Police Officers from different departments in our local area. We recruited these two police officers by calling the police stations and asking if they would participate in a short interview about how they use security camera footage in their roles and to test the initial timeline visualizations we developed. The first police officer we interviewed was a detective for their precinct. The officer mentioned that for his role, he gets assigned various cases including white-collar crimes, violent crimes,

and sexual violence crimes so he has to view security camera footage very frequently for his role while investigating these cases. The second police officer we interviewed was a crime prevention officer for our university's police department. The officer mentioned that for his role, he performs security assessments of buildings, reports crime statistics, trains recruits to the academy, and helps with patrols or investigations when needed.

After interviewing the police officers, we tried to find other types of people who used security camera footage to interview leading, us to Prolific. We used Prolific to create an advertisement for our pilot study since the website is designed to help researchers advertise their studies across the world. Since we had previous luck using Zoom for the interviews with the police officers, we decided that Zoom interviews would be sufficient for any participants recruited through Prolific. The first participant we recruited from Prolific has been a homeowner for twenty years. The homeowner mentioned having a Ring Camera and an ADT home security system for safety purposes since they live in an urban community. They mainly view the security camera footage daily to ensure that no criminal breaks into their home or cars, keep an eye out for anyone walking around their house that they do not recognize, and help look out for their neighbor's homes and cars. The second participant we recruited from Prolific was a student who worked on a farm during her undergraduate studies as an undergraduate researcher. The student mentioned that they mainly viewed the security camera footage weekly to ensure that no animals like deer, raccoons, groundhogs, and more wander onto the farm and eat the crops since those crops are what her research was based on. To collect the security camera footage, the farm used GoPros which they placed in various areas around the farm and used the software the GoPros came with to review the footage as well as a camera and software made by a company named Access which would send them motion notifications on their mobile devices. The third participant we recruited from Prolific was a Security Specialist who works in an office for the water

district of a government agency. The security specialist mentioned that they mainly viewed the security camera footage daily to ensure that they could respond quickly to any incidents that occurred with customers in the building lobby or the parking lot outside of the office. To collect the security camera footage, the office used cameras they placed in various areas around and outside the office which were wired with Ethernet for a stable internet connection and streamed the footage to the security office's computers.

After the pilot study, we gathered the transcriptions from the five interviews and created a codebook with each code consisting of a label, definition, notes about the code, and the participants that mentioned something related to that code. After collecting all of the codes, we split the code labels into two types of labels based on whether the participant was talking about our timeline visualizations or not. We then used Miro to organize our codes into different categories. For the entire codebook we created, please refer to Appendix A.

To begin, we wanted to look at some codes that did not focus on our timeline visualizations but influenced improvements to our website. The first subcategory was about Difficulties with Reviewing Camera Footage with the first notable code being C03 about how there is no standard across different businesses regarding the video footage format or software used to view the footage. There was also code C09 about how some cameras have AI features but they mainly only detect motion in view and do not give any distinguishing information about the motion. Lastly, there were codes related to the massive amount of information users may need to review on occasion with code C29 stating that when there are large timeframes, it is basically impossible to review the footage.

The second subcategory was about Useful Video Reviewing Software Features. In this subcategory, the first notable code was C18 which referred to being able to mark/annotate timestamps in the footage for later review/storage. Another notable code was C10 which was about how existing systems have motion visually annotated on the timeline/video or

alerted to the user through notifications. Lastly, there was code C42 which talked about a feature to zoom in or zoom out of the timeline bar using the scroll wheel to make the timeline event density change.

Now we will expand on the codes related to comments about our timeline visualizations and how we organized them further into different subcategories. The first subcategory of codes related to comments about our timeline visualizations was the Navigability of the Current System. In this subcategory, we had codes related to Event Filtering with code E19 describing how filtering for different events could narrow the search time and code E48 noting that our current filtering method is more informative than existing home systems. We had codes related to Reducing Search Time as well with code E20 mentioning how jumping to events by clicking on an event block can make searching faster as well as code E28 highlighting that when using the density graph, you do not need to watch the downtime periods and can just focus on the peaks. We also had a code related to Temporal Cues with code E21 stating that the Event Blocks and Density Graph give a better idea of where the event is and what is occurring during the time of the event. Lastly, there was code E36 related to how the UI is user-friendly and simplistic which contrasted with existing interfaces that make their UI very cluttered with features along with code E39 which emphasized that our system allows advanced features to be toggled on and off which is helpful when the features block the view of the video for the user.

The second subcategory of codes related to comments about our timeline visualizations was Timeline Type Preference Reasons. In this subcategory, we had a code related to Finding Known Events with code E23 stating that some participants prefer Thumbnails when looking for a person carrying an object. Similarly, there was a code about Immediate Visuals with code E56 stating that some participants prefer Thumbnails for an immediate visual of what is occurring in an event. There was also a code related to Confirmation where code E55

stated that the participant preferred the thumbnails for checking if an event was categorized incorrectly or not. Lastly, there was a code related to Control Over Intelligent Features where code E40 stated that a participant preferred the Density Graph over Event Blocks since he did not like how the downtime periods were cut out since there could be evidence in those periods of downtime.

The third subcategory of codes related to comments about our timeline visualizations was Suggested Video Reviewing Software Features. In this subcategory, we had a code related to Video Traversal with code E41 talking about a feature to grab the linear timeline slider and move it along with the video. There was a code related to an improvement for the Temporal Cues as well, code E25 pointed out a feature to change the opacity of an event when a user has viewed it so there is some visual indicator to prevent users from getting confused about which event they already viewed. Fast forwarding was mentioned as well with code E37 which was related to implementing a feature to increase or decrease the playback rate of the video. Lastly, there was code E26 where a participant mentioned that more Advanced Identification / Filtering would be helpful to implement in the future.

The last subcategory of codes related to comments about our timeline visualizations was Limitations with the Current Features. In this subcategory, we had codes related to Thumbnails Misguiding Users where code E59 stated that thumbnails may lead to participants assuming some events are not the ones they are looking for, E24 and E50 which were about some participants relying on the thumbnail and how one frame is not enough information to describe the entire event. Navigation was also mentioned with code E43 talking about how there needs to be a method to move the timeline in small chunks to notice subtle changes. There were codes about the massive amount of material to review when a video has a lot of events as well with E58 and E49 describing how the density graph may be overwhelming with lots of events and when the event blocks and density graph are combined, they give

too much information. Lastly, there was a negative code about Temporal Cues where code E57 mentioned a participant did not prefer the event blocks because it is harder to see when events start and end.

4.3 Improvements

After organizing the codes into various subcategories, we used the information collected to inform improvements to our current timeline visualizations.

4.3.1 Base Interface

Figure 4.1: Base Interface

For the improvements that we made to Timeline 1 (Base Interface) which are shown in Figure 4.1, we started by adding a method to speed up and slow down the playback rate of the video since it was a heavily requested feature from the participants of the pilot study. We used YouTube's player to play our videos so we implemented speed-up and slow-down buttons which would increase and decrease the playback rate respectively based on the preset rates that YouTube already has. We also added an indicator underneath the linear timeline to report to the user the video's current playback rate.

4.3.2 Density Graph + Event Blocks

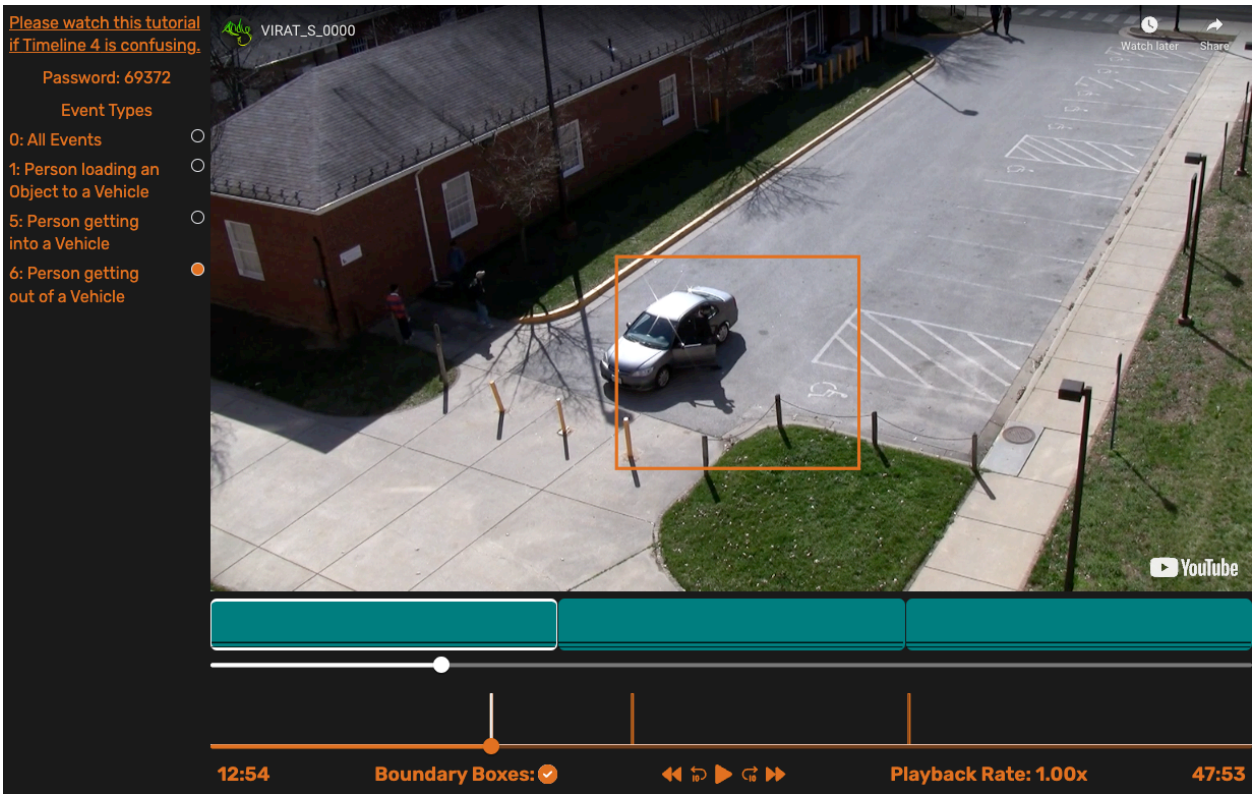


Figure 4.2: Density Graph + Event Blocks

As for the improvements made to Timelines 2-4 which are shown in Figure 4.2, we changed the dropdown menu for the event types to a menu where all the event types are shown to the user at once, and then when the user selects an event type, a radio button is checked. We decided to change the dropdown menu because we wanted users to be able to see all of the event types present in the video immediately when the interface is loaded. We added this new select menu to a permanent sidebar that is on the left side of the video player.

We also implemented a feature where when the timeline is within an event, the density graph at that event will be highlighted in white so that the user knows which event is currently happening. We implemented a similar feature for Timeline 3 (Event Blocks) where when an

event is occurring in the event block, the event block will have a white outline around it to notify the user. After implementing these two features, we also made a change to Timeline 4 (Density Graph + Event Blocks) where when a user hovers over an event block, the event block has a green outline and the density graph for that event block is highlighted in green so that the user can tell which event block corresponds to which peak on the density graph before they click that event block.

4.3.3 Event Thumbnails

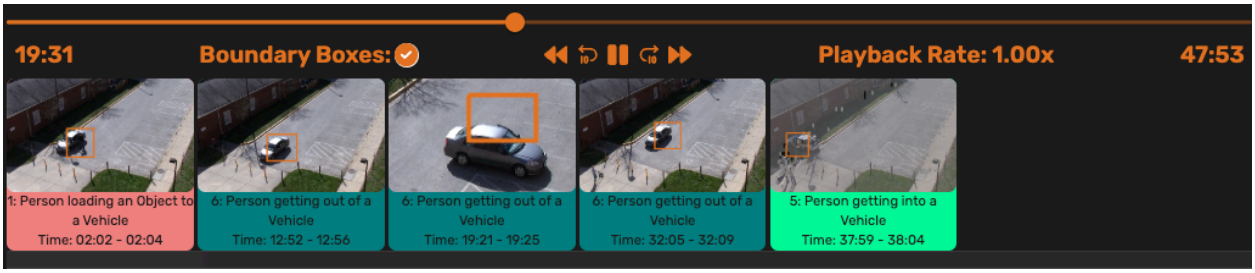


Figure 4.3: Event Thumbnails

Lastly, for the improvements made to Timeline 5 (Event Thumbnails) which are shown in Figure 4.3, since the sidebar on the left was not being used by the menu for the event types, we made the list of thumbnails a horizontal list and placed it underneath the linear timeline and video controls. We also added a feature based on feedback from the pilot study where when a user clicks on one of the thumbnails, it will zoom into the image and allow the user to get a closer look at what action is happening in the thumbnail.

Chapter 5

General User Study

5.1 Study Design

After we made the improvements based on the pilot study, we wanted to modify our study to be a general user study. We decided to use QuestionPro as the main platform for the general user study since, after looking at our pilot study, we believed we could convert the interview questions and tasks into a survey format, which would allow us to be able to gather more results since the survey could be completed asynchronously by the participants. For our survey, the first section presented to the participant was the consent form outlining all of the necessary preliminary information for this study, if a participant did not want to participate in the study after reading the consent form, they could reply with "I do NOT consent to participate in this research." which would terminate the survey. If they did choose to participate in the study, they would then be presented with the introduction section. In the introduction, we asked participants these questions:

- Please enter your email address.
- What gender do you identify as?
  - Male
  - Female

  - Non-Binary
  - Prefer not to disclose
  - Other

- How old are you?
  - Under 18
  - 18 - 24
  - 25 - 34
  - 35 - 44
  - 45 - 54
  - 55 - 64
  - Above 64

- What race or ethnicity describes you?
  - American Indian or Alaskan Native
  - Asian / Pacific Islander
  - Black or African American
  - Hispanic
  - White / Caucasian
  - Other

- Have you ever viewed surveillance video footage, if so, how often do you view surveillance video footage (Examples: footage of a site (home, street, retail stores, building, bird watchers, etc.))?

  - Daily
  - Weekly
  - Monthly
  - Yearly
  - Never

- If you have previously viewed surveillance video footage, what are the typical events or incidents are the ones that you look for when you view the footage?
- If you have previously viewed surveillance video footage, what information do you typically use and are given as clues to spot the events? (time window, profile, or nothing)?
- If you have previously viewed surveillance video footage, what software/products/systems did you use to view the surveillance video footage?

We initially asked participants basic gender, age, and ethnicity questions followed by questions based on their previous usage of security camera video footage to gauge how diverse our participant pool was. After the participant answered the questions in the introduction section, we moved on to the tutorial section, where the participant was asked to watch a brief tutorial about how to use all five of the different timeline visualizations and then confirm that they had watched the tutorial by answering five questions. Each of the five questions showed the participant a picture of a timeline visualization we created along with this question:

- Please select the timeline visualization that is shown in the image below.
  - Timeline 1 (Base Interface)

- Timeline 2 (Density Graph)
- Timeline 3 (Timeline of Event Blocks)
- Timeline 4 (Density Graph + Timeline of Event Blocks)
- Timeline 5 (Event Thumbnails)

5.1.1 Usability Study

After finishing the tutorial block, we created a Question Pro block for each of the five videos. Each block begins by asking the participant to open a link to our DUET website with the video associated with that block playing and a randomized timeline visualization. We then asked the participant to enter the password displayed on the DUET website into a question for that block, which was used to handle the timeline randomization. After entering the password, we asked the participants to use the current timeline visualization to solve two questions for that block. These two questions were our method of testing if the different timeline visualizations affected the user's correctness when finding specific events in the videos. For each question, we also added an option for a user to choose if they could not find the answer within five minutes so that a user would not get stuck on a question for too long. After attempting to solve the two questions for that specific event block using the randomized timeline visualization, we asked participants to select the timeline visualization that they used to complete the two questions and then asked them to judge the timeline visualization using the System Usability Scale or SUS scale which as the name suggests, was our main metric for measuring the user's perceived usability of each timeline visualization they used. Each block would record the time that the user spent within the block as well so that we could see if using the different timeline visualizations would affect the time it took the user to complete the tasks. Here is an example of one of the question sets for the

VIRAT-S-0002 Question Block:

- Task: The driver of this car gets out of the car and has a conversation with another person. What is the car that the other person gets into later?
  - Car 1
  - Car 2
  - Car 3
  - Car 4
  - Car 5
  - I could not find the answer within five minutes.
- Task: During which timeframe do you see this man unloading a blue shoulder bag from his car?
  - 11:00 – 14:00
  - 18:00 – 20:00
  - 20:00 – 22:00
  - 32:00 – 35:00
  - 43:00 – 45:00
  - 13:00 – 15:00
  - 16:00 – 18:00
  - 22:00 – 23:00
  - 34:00 – 35:00
  - 45:00 – 47:00

  - I could not find the answer within five minutes.

- Please select the timeline visualization that you used to complete the previous tasks.
  - Timeline 1 (Base Interface)
  - Timeline 2 (Density Graph)
  - Timeline 3 (Timeline of Event Blocks)
  - Timeline 4 (Density Graph + Timeline of Event Blocks)
  - Timeline 5 (Event Thumbnails)

- System Usability Scale (SUS)
  - I think that I would like to use this UI frequently.
    - Strongly Disagree
    - Disagree
    - Neutral
    - Agree
    - Strongly Agree
  - I found the UI unnecessarily complex.
    - Strongly Disagree
    - Disagree
    - Neutral
    - Agree
    - Strongly Agree
  - I thought the UI was easy to use.
    - Strongly Disagree

    * Disagree
    * Neutral
    * Agree
    * Strongly Agree

- I think that I would need the support of a technical person to be able to use this UI.
    * Strongly Disagree
    * Disagree
    * Neutral
    * Agree
    * Strongly Agree

- I found the various functions in this UI were well integrated.
    * Strongly Disagree
    * Disagree
    * Neutral
    * Agree
    * Strongly Agree

- I thought there was too much inconsistency in this UI.
    * Strongly Disagree
    * Disagree
    * Neutral
    * Agree
    * Strongly Agree

- I would imagine that most people would learn to use this UI very quickly.
  - Strongly Disagree
  - Disagree
  - Neutral
  - Agree
  - Strongly Agree
- I found the UI very cumbersome to use.
  - Strongly Disagree
  - Disagree
  - Neutral
  - Agree
  - Strongly Agree
- I felt very confident using the UI.
  - Strongly Disagree
  - Disagree
  - Neutral
  - Agree
  - Strongly Agree
- I needed to learn a lot of things before I could get going with this UI.
  - Strongly Disagree
  - Disagree
  - Neutral
  - Agree

    * Strongly Agree

The questions for the other question blocks can be seen in Appendix B.

5.1.2 Exit Interview

After completing tasks using the five different videos with the five different timeline visualizations, we asked each participant some exit questions before they completed the survey so that we could collect the user's thoughts about the different timeline visualizations. Specifically, we asked:

- Please rank (1-5) the following in order of which timeline visualization you liked the most:
  - Timeline 1 (Base Interface)
  - Timeline 2 (Density Graph)
  - Timeline 3 (Timeline of Event Blocks)
  - Timeline 4 (Density Graph + Timeline of Event Blocks)
  - Timeline 5 (Event Thumbnails)
- Please explain your preference for the ranked choices you made in the previous question.
- Are there any other features you think would be nice to implement in our current system?
- Do you believe the tasks we had you perform were realistic to those you look for when you browse for events in surveillance camera footage?
- Do you have any questions for us?

After the exit questions, we thanked the participants for their time, and let them know the timeline for compensation which was a 1 out of 10 chance to win a $10 Amazon Gift Card.

Chapter 6

Results

6.1 Participants

During the user study, we recruited 72 participants through our University's mailing lists, local police stations, and other forms of online advertisements. Out of the participants, 57 identified as Male, 14 identified as Female, and 1 identified as Non-Binary. As for the age ranges of the participants, 62 were between the ages of 18 - 24 years old, 8 were between the ages of 25 - 34 years old, and 2 were between the ages of 35 - 44 years old. For the ethnic diversity of the participants, 1 participant was American Indian or Alaskan Native, 31 participants were Asian or Pacific Islander, 11 participants were Black or African American, 1 participant was Hispanic, and 26 participants were White or Caucasian. Lastly, for their self-reported amount of time spent viewing security camera footage, 29 participants had Never viewed security camera footage, 13 participants viewed security camera footage on a Weekly basis, 12 participants viewed security camera footage on a Monthly basis, and 18 participants viewed security camera footage on a Yearly basis.

6.2 Data Analysis

To analyze the data collected from the survey and user study, we decided to focus on four different aspects, the number of questions the user got correct based on the timeline visu-

alization they used, the participant's SUS Questionnaire score for the timeline visualization that they used, the time the participant took to complete each question block, and the participant's personal ranking of the five timeline visualizations.

6.3 Correctness Results

For the Correctness Results, we calculated a participant's correctness score based on how many questions the participant got correct for each timeline visualization. If a participant got none of the questions correct, their score was zero, if the participant got one question correct, their score was 1 and if the participant got both questions correct, their score was 2. After we got the scores from each participant for each of the five different timeline visualizations, we calculated the mean and standard error for Correctness for each timeline visualization which is shown in Figure 6.1.

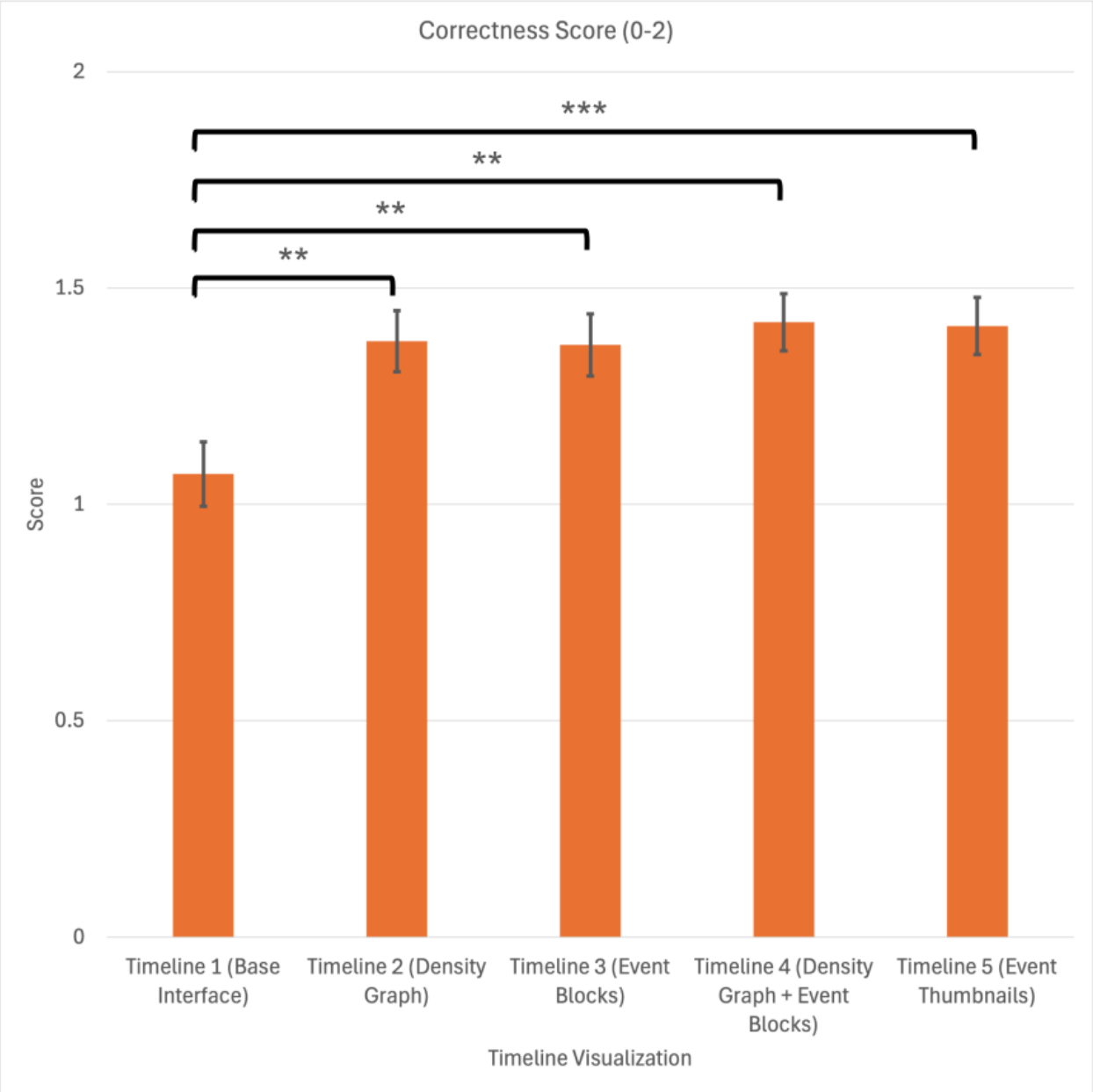


Figure 6.1: Correctness Bar Graph

For Timeline 1 (Base Interface), the mean Correctness score was 1.070, with a standard error of 0.074. For Timeline 2 (Density Graph), the mean Correctness score was 1.377, with a standard error of 0.071. For Timeline 3 (Event Blocks), the mean Correctness score was 1.368, with a standard error of 0.071. For Timeline 4 (Density Graph + Event Blocks),

the mean Correctness score was 1.421, with a standard error of 0.066. Lastly, for Timeline 5 (Event Thumbnails), the mean Correctness score was 1.412, with a standard error of 0.066. Running a Kruskal Wallis test against the Correctness scores for all of the timeline visualizations, we see our Correctness scores are statistically significant ($p = 0.002274$). Thus, we can conclude from the Correctness scores that the timeline visualization affects the participant's ability to find the correct events while searching for events in the video footage. For our post hoc analysis of the Correctness scores, we ran a Dunn's Multiple Comparison test which showed that all four other timeline visualizations led to higher correctness scores than Timeline 1 (Base Interface).

6.4 Time Results

For the Time Results, we calculated the mean and standard error for the time taken to complete a question block using each of the timeline visualizations which is shown in Figure 6.2.

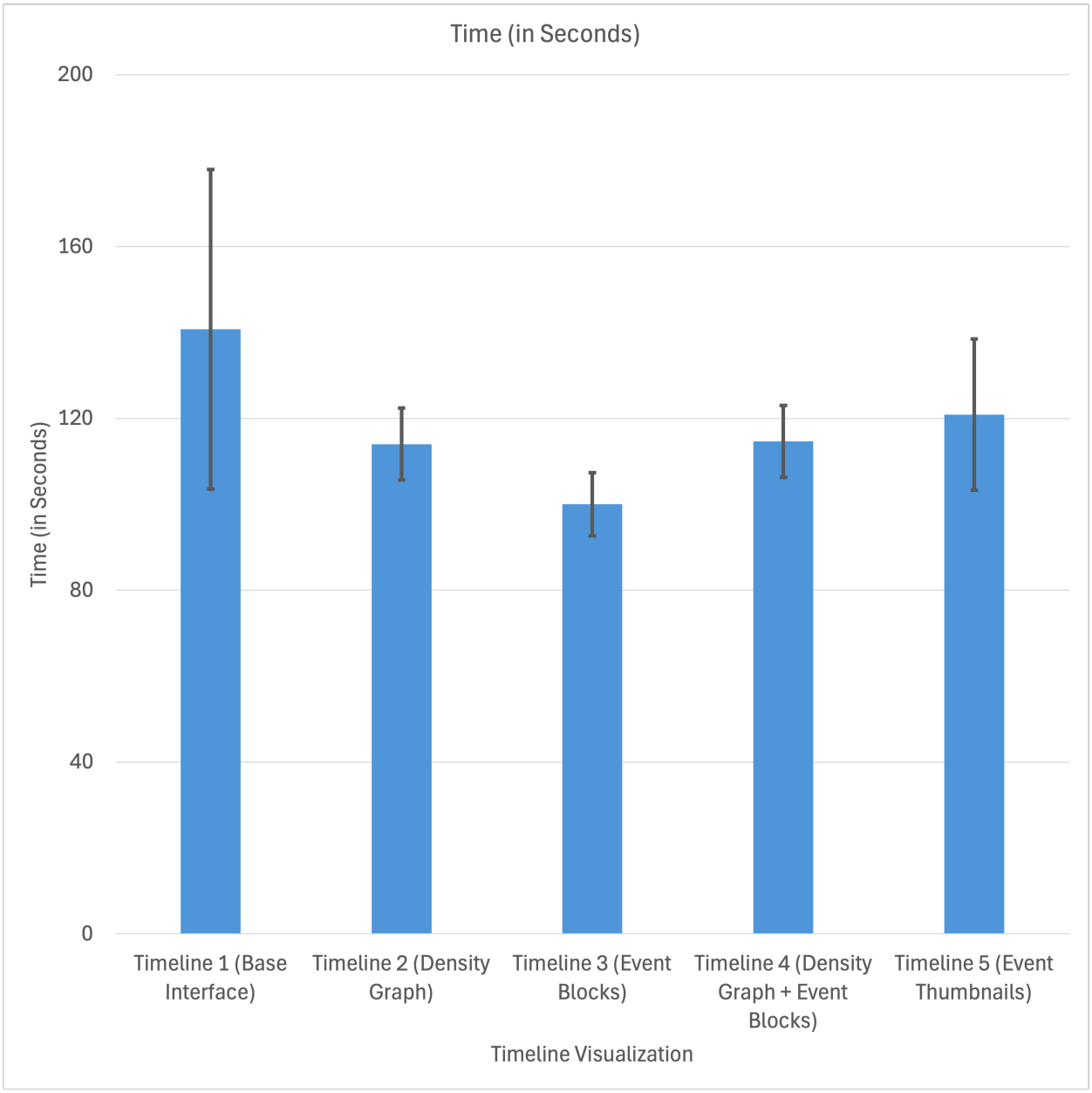


Figure 6.2: Time Bar Graph

For Timeline 1 (Base Interface), the mean time taken to complete the tasks in a question block was 140.750 seconds, with a standard error of 37.197 seconds. For Timeline 2 (Density Graph), the mean time taken to complete the tasks in a question block was 114.014 seconds, with a standard error of 8.378 seconds. For Timeline 3 (Event Blocks), the mean time taken

to complete the tasks in a question block was 100.014 seconds, with a standard error of 7.334 seconds. For Timeline 4 (Density Graph + Event Blocks), the mean time taken to complete the tasks in the question block was 114.639 seconds, with a standard error of 8.393 seconds. For Timeline 5 (Event Thumbnails), the mean time taken to complete the tasks in a question block was 120.875 seconds, with a standard error of 17.568 seconds. Running a Kruskal Wallis test against the time taken to complete a question block using each timeline visualization, we see that the data we collected about the time taken to complete the blocks is not statistically significant ($p = 0.5604$). Thus, we can not conclude that the time taken to complete a question block for each timeline visualization affects the effectiveness of each timeline visualization.

6.5 SUS Questionnaire Results

For the SUS Questionnaire Results, we calculated the mean and standard error for the SUS Questionnaire scores for each of the timeline visualizations, which is shown in Figure 6.3.

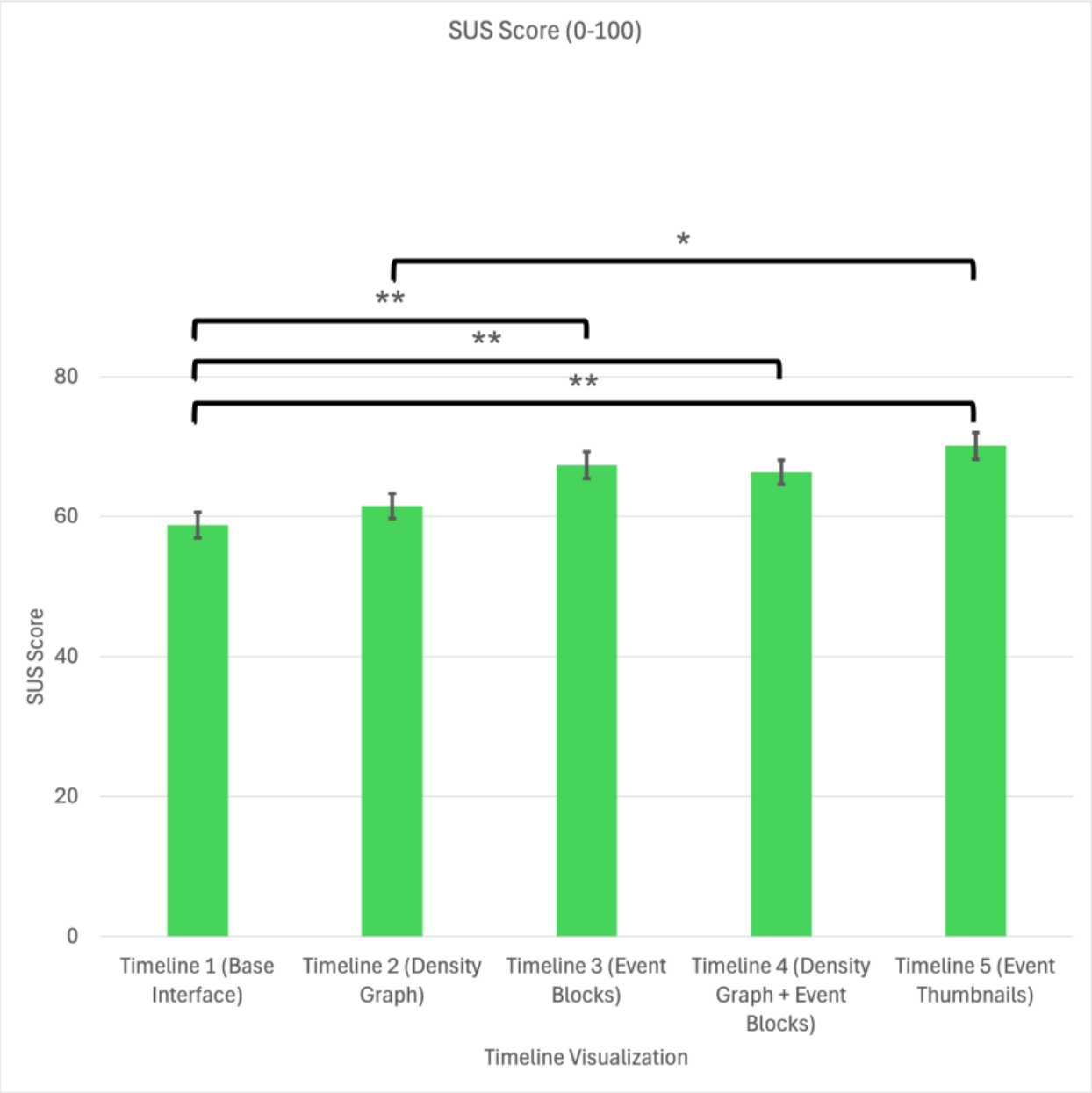


Figure 6.3: SUS Bar Graph

For Timeline 1 (Base Interface), the mean SUS Questionnaire score was 58.778, with a standard error of 1.834. For Timeline 2 (Density Graph), the mean SUS Questionnaire score was 61.534, with a standard error of 1.794. For Timeline 3 (Event Blocks), the mean SUS Questionnaire score was 67.358, with a standard error of 1.877. For Timeline 4 (Density

Graph + Event Blocks), the mean SUS Questionnaire score was 66.364, with a standard error of 1.739. For Timeline 5 (Event Thumbnails), the mean SUS Questionnaire score was 70.114, with a standard error of 1.902. All five of the mean SUS Questionnaire scores are in the OK range when interpreting the SUS Questionnaire scores through adjectives [10]. The SUS Questionnaire score can also be interpreted as a "letter grade", with percentage ratings between 0-100 similar to the traditional academic grading scale [3]. Using the letter grade interpretation, the letter grades for each respective timeline are Timeline 1 (Base Interface) and Timeline 2 (Density Graph): D, Timeline 3 (Event Blocks), Timeline 4 (Density Graph + Event Blocks), and Timeline 5 (Event Thumbnails): C. Running a Kruskal Wallis test against the SUS Questionnaire scores for all timelines, we see our SUS Questionnaire scores are statistically significant ($p = 0.0004135$). Thus, we can conclude that our SUS Questionnaire scores show that the timeline visualization greatly affects the user's usability while searching for events in the video footage. For our post hoc analysis of the SUS Questionnaire scores, we ran a Dunn's Multiple Comparison test which showed that Timeline 3 (Event Blocks), Timeline 4 (Density Graph + Event Blocks), and Timeline 5 (Event Thumbnails) led to higher SUS Questionnaire scores than Timeline 1 (Base Interface) and Timeline 5 (Event Thumbnails) led to higher SUS Questionnaire scores than Timeline 2 (Density Graph).

6.6 Timeline Visualization Ranking Results

For the participant's personal ranking of the five timeline visualizations, we created a stacked bar chart to show the participant's rankings of the timeline visualizations which is shown in Figure 6.4.

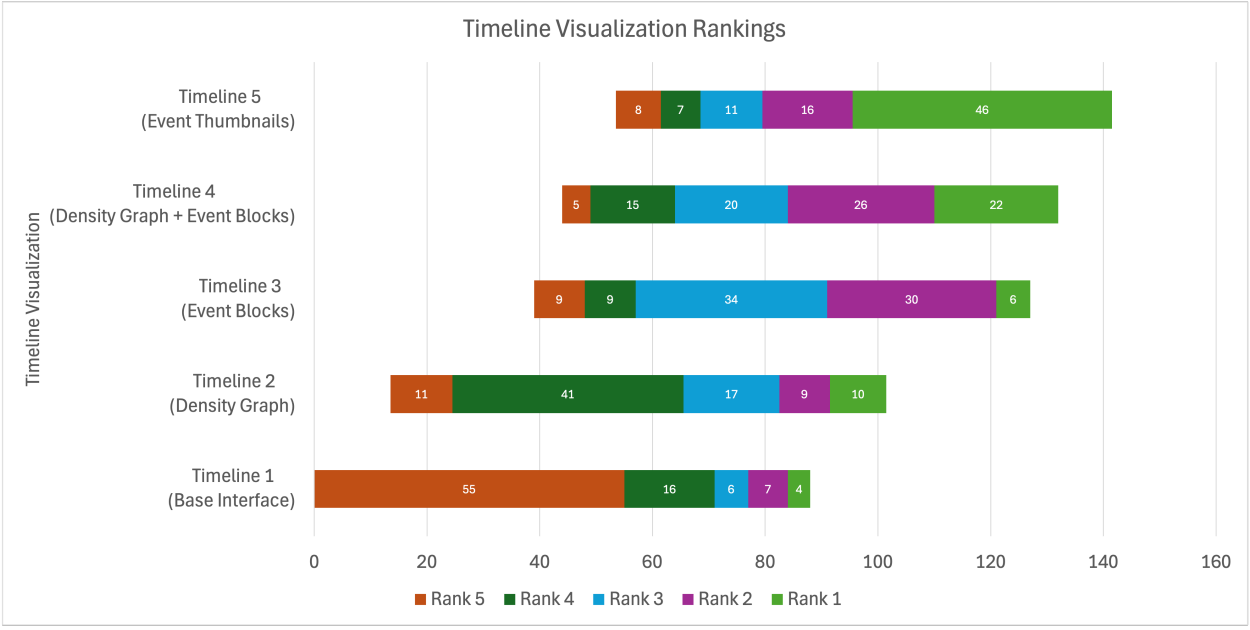


Figure 6.4: Rankings Graph

For Timeline 1 (Base Interface), 4 participants ranked it as their first choice, 7 participants ranked it as their second choice, 6 participants ranked it as their third choice, 16 participants ranked it as their fourth choice, and 55 participants ranked it as their fifth choice. For Timeline 2 (Density Graph), 10 participants ranked it as their first choice, 9 participants ranked it as their second choice, 17 participants ranked it as their third choice, 41 participants ranked it as their fourth choice, and 11 participants ranked it as their fifth choice. For Timeline 3 (Event Blocks), 6 participants ranked it as their first choice, 30 participants ranked it as their second choice, 34 participants ranked it as their third choice, 9 participants ranked it as their fourth choice, and 9 participants ranked it as their fifth choice. For Timeline 4 (Density Graph + Event Blocks), 22 participants ranked it as their first choice, 26 participants ranked it as their second choice, 20 participants ranked it as their third choice, 15 participants ranked it as their fourth choice, and 5 participants ranked it as their fifth choice. For Timeline 5 (Event Thumbnails), 46 participants ranked it as their first choice, 16 participants ranked it as their second choice, 11 participants ranked it as their third choice,

7 participants ranked it as their fourth choice, and 8 participants ranked it as their fifth choice. When asked to explain their preference for the rankings, participants mentioned that the Event Thumbnails were easy to understand and gave the participants immediate visual feedback on what events were occurring at that time. Although we saw that even though participants preferred the Event Thumbnails due to their immediate visuals, their preferences did not lead to better performance gain in regards to accuracy in completing the questions and time spent in the block. Some participants mentioned that they preferred Timeline 4 because of the amount of detail that it offered when searching for events since it combined the features of the Density Graph and the Event Blocks. However, other participants liked either the Density Graph or the Event Blocks on their own since they thought having both displayed together was too much information at once. This split in opinions clearly emphasizes that future timeline visualizations and interfaces should allow users to toggle which features they want on or off so that users do not get overwhelmed with too much information. We also saw that the Density Graph on its own was not preferred in the rankings and in terms of performance gain, it did not differ significantly than Timelines 3-5. The previewed usability was also significantly lower than the Event Thumbnails and on par with Timeline 3 and Timeline 4. Lastly, some saw the Base Interface as an easy-to-use interface but found it very difficult to find accurate information promptly.

Chapter 7

Discussion

Our study highlights some of the improvements that could be made to traditional timeline visualizations to help users navigate through long videos more seamlessly. We built upon existing concepts such as the traditional linear timeline, the density graph, and using thumbnails as well as implemented new methods for a user to navigate through events using the event blocks. Thus, we believe that the DUET system can introduce more user-friendly methods of timeline navigation for events that happen in a video regardless of the length of the video. We also believe that DUET can be expanded to other domains aside from security camera footage, for example, the system could be used to search for different events that happen in movies which could help screenwriters find scenes quicker that they want to take inspiration from when creating new films. DUET could help students who are taking asynchronous classes who have to review long lecture videos since the students could skip directly to certain areas in the video that they want to review. We believe many different domains could utilize our system to make their video analyzing experience faster.

Of course, to expand our interface in the future, there would need to be more features implemented into the system. One of the recommended features from participants was showing a preview when you hover over the scroll/playback bar, which would be very beneficial for providing more immediate visual feedback to the user. Another recommended feature from participants was the ability to pin or bookmark timestamps and events in the video, which would help a lot for users who want to quickly access certain events later on. We believe

that any future timeline visualizations should consider these features.

Aside from the features mentioned above, there was one proposed feature that highlights one primary limitation of the system, most video footage does not have the events labeled. However, although we did not create the DUET system to automatically detect and label events, we believe that event recognition and labeling is a perfect task for an emerging technology in the computer science sphere which is AI and LLMs. We believe that in a future iteration of DUET, we could allow the user to upload a large video to the interface and then using LLMs and computer vision, we could parse the video to find and create filters for different types of events that occur in the video which then could be stored and used by the user to find different events in the video they just uploaded. However, using AI and LLMs to automatically detect and label events comes with its own struggles as well since this approach can lead to inaccurate labels or categorizations. There may need to be some form of user review or some other form of human review of the generated labels before they can be used which could hinder the initial process of using the platform for any type of video. To help aid in the review of the generated labels, we plan to create a way to edit, archive and delete different event types and video categorizations so that users can handle the inaccuracies.

Another future direction for the project could focus on how the density of events can impact the perceived usability of timeline visualizations. We believe that in the future, we can adapt the timeline visualization based on the event density since we could allow the user to zoom in to create a more sparse view of the events since it would only show the ones that are happening in that timeframe, additionally when the user zooms-out the timeline visualization would become dense again to show all of the events. We believe that this improvement would help address another limitation of our study, which is that, in practice, the videos that are being reviewed can be much longer than the ones that we used in our

study.

Chapter 8

Conclusion

We believe our study highlights and evaluates the main design limitations of using a linear timeline to search through specific events in long videos and shows the promise of different methods to help visualize the events in the video to users. Our studies focused on the feasibility of implementing and combining different timeline visualizations for events to help aid in navigating through long videos and the usability of those visualizations. While some of the findings reaffirm existing results such as that users prefer the visual aspects of the event thumbnails, leading to its higher usability scores as well as that a linear timeline with no addons is not well suited for finding events which is shown by the poor performance of the Base Interface, other findings do suggest that existing video players could benefit with some additions. For example, the Event Blocks allowed users to take the least time completing the tasks. Additionally, the combination of the Density Graph and the Event Blocks led to the highest number of correct answers when completing our tasks, showing the benefits of these two visualizations in aiding how users search for events. We believe DUET's timeline visualizations could be used to inform how to improve the usability of existing video players and timeline visualizations implemented in existing and future systems. If you would like to check out the project and the data, it is available at this GitHub repository: https://github.com/echo-lab/dual-timeline.

Appendices

Appendix A

Pilot Study Coding Session

A.1 Codes Not Related to Comments About Our Timeline Visualizations

- The context of reviewing camera footage
  - Crime-related
    - C01:Reviewing camera footage is a primary task of the user:P01, P02, P03
    - C02:Crime Reports may involve camera footage available:P01, P02
    - C44:Home and community protection of neighbors involves security camera footage:P03
  - Long Time Frame
    - C08:Case where the user does not know the timeframe of the incident of interest, video review can take a while:Video length can be up to two weeks:P01, P02
    - Negative:C07:Case when the user does know the timeframe of the incident of interest, requests short 5-10 min video snippets, video review will be shorter:Significantly reduces video length needed for review:P01, P02
    - C60:Some companies are obligated to keep the footage for a certain amount

of time:P05

- Video Software
  - C16:Specialized software that users use to review footage:Security Desk, WatchGuard, Axon, ADT, Ring, GoPro Software, Axis:P02
  - C17:Some users do not use specialized software to view the security footage, they just use the standard video player that their device comes with:P01, P02
- Pre-Edited Footage
  - C27:Loss Prevention Agents help the officers filter out the large videos for information they find useful:Liasion to law enforcement:P01
- Alertness
  - C45:Home security systems have an armed mode and unarmed mode:Can set it to be armed at away times:P03
  - C62:For tracking users wearing an object, watch them for a little bit, if they are going offscreen, move on::P05
- Research/Observation Related
  - C51:New use case for security camera systems:Home monitoring, checking for wild animals eating plants, office security, checking for suspicous clients:P03, P04
  - C54:Security camera system used to help aid in research/data collection:Have to reivew a lot more of the footage to make sure accurate data was being collected about the plants:P04

- What Users seek when reviewing camera footage
  - Reviewing not knowing what they will learn

        * C04:Video Footage may not involve main incident but auxiluary information of suspect / vehicles:Footage may not show their face but can show two people breaking into a home:P01, P02, P03
        * C08:Case where the user does not know the timeframe of the incident of interest, video review can take a while:Video length can be up to two weeks:P01, P02
    - Auxiliary information that are not the target incident
        * C04:Video Footage may not involve main incident but auxiluary information of suspect / vehicles:Footage may not show their face but can show two people breaking into a home:P01, P02, P03
- Useful Video Reviewing Software Features
    - Motion Detection Indicators
        * C10:Motion is visually annotated on the timeline / video or alerted to the user through notifications:Colored bar, red box:P01, P02, P03, P04, P05
        * C46:Auxillary equipment to aid the security system:Mobile application with motion notifications and chime for motion alerts:P03
        * C61:Camera is not recording all the time, usually only recording when it senses activity:P05
    - No Intelligence
        * C11:Security Camera System does not have any advanced intelligent features:P02
        * C17:Some users do not use specialized software to view the security footage, they just use the standard video player that their device comes with:P01, P02

  - Annotations/Bookmarks
    - C18:Being able to mark/annotate timestamps in the footage for later review/storage and be able to navigate through them without temporal navigation:P01, P02
  - Multi-Window
    - C30:Basic video players have multiple windows you can open:P01
    - Negative:C31:Multiple views can be helpful but often views you can negate:P01
  - Playback speed
    - C32:Some players have the option to fast forward through the footage:P01
  - Audio/Transcript
    - C34:Feature to transcribe the audio in the security camera footage:Axon, transcribes audio and picks up on police warnings for body cams:P01
  - Timeline Zoom In / Out
    - C42:Feature to zoom in or zoom out of the timeline bar using the scroll wheel:P01
  - Offline Video Saving
    - C52:Able to download the data from the systems to store for offline use:P04
- Information given to users before reviewing footage
  - Visual Cues
    - C05:Users are given basic visual information about suspect before looking at the footage:biometric information, eye/hair color:P02
    - C28:Users are given basic information about shoplifting before looking at the footage:What was stolen, what time it occured:P01

        * C06:Users are given basic information about vehicle appearance before looking at the footage:Vehicle make/model/color:P02
        * Negative:C64:Motion alerts can obstruct view of user and prevent identification of obscure visuals:P01
    - Temporal Cues
        * C63:Users are given temporal information about suspect's actions before looking at the footage:receipts, credit card information:P02
        * C28:Users are given basic information about shoplifting before looking at the footage:What was stolen, what time it occured:P01
        * C63:Users are given temporal information about suspect's actions before looking at the footage:receipts, credit card information:P02
        * C07:Case when the user does know the timeframe of the incident of interest, requests short 5-10 min video snippets, video review will be shorter:Significantly reduces video length needed for review:P01, P02
        * Negative:C14:Eyewitnesses do not give any information about the timeframe of the incident of interest:P02
        * Negative:C08:Case where the user does not know the timeframe of the incident of interest, video review can take a while:Video length can be up to two weeks:P01, P02
    - Sequence of Events
        * C33:Sequence of events in security footage:Kids scope out store, later return to rob store:P01
- Difficulties with reviewing camera footage
    - Challenges regarding multiple video sources

    * C03:No standard across different businesses regarding video footage format, software ease of use:Disk/USB/File Player:P01, P02, P04

- No / Limited intelligent functionality
    * C09:AI features does not distinguish types of behaviour, only detects motion in view:P02, P03, P04
    * C32:Some players have the option to fast forward through the footage:P01
    * C53:Motion detection requires larger objects for some systems:Axis:P04
    * C65:Some applications only have time based filters:P03
- Massive amount of material to review
    * C13:Cannot jump to specific instances, must continously review video footage:P02, P04
    * C29:Reviewing footage across large timeframes is basically impossible:P01
- Massive amount of material to review even with intelligence
    * C12:Videos that have lots of traffic, must review every instance of a person-/vehicle/subject:Level of computer vision must be very detailed to filter out specifics:P02
    * C47:Some camera systems are prone to false positive results:Insects that fly into the view, Amazon deliverys:P03
- Incorrect Cues Given
    * C15:Eyewitnesses do not give correct information about the timeframe of the incident of interest:P02
- Low Quality
    * C35:Camera footage is not usually high quality:P01

A.2 Codes Related to Comments About Our Timeline Visualizations

- Navigability of current system
  - Event Filtering
    - E19:Being able to have specific filters for types of events that occur in the videos to narrow down the search time using timeline 2:specific things like loading objects in unloading, uh, opening a vehicle, closing a vehicle, getting into a vehicle, getting out, carrying an object:P01, P02, P04
    - E48:Current filtering method is more informative than existing home systems::P03
  - Reduce Search Time
    - E20:Using the event blocks to click to instances of events and being able to let the events run to see if you found the right one to minimize searching:P02
    - E38:Scenario for using the density graph feature:For a carjacking, do not need to watch the periods of downtime, can just watch the peaks:P01
  - Temporal Cues
    - E21:Event Blocks and Density Graph give a better idea of where the event is and what is occuring during the time of the event:P03 initially disliked T4 but used it for task 1 and said this:P02, P03
  - User Interface
    - E36:UI is user friendly and simplistic:Normal players have various features which make the UI very busy:P01
  - Toggling Features

        * E39:Allow advanced features to be toggled on and off:P01, P03

- Timeline Type Preference Reasons
    - Finding Known Events
        * E23:Prefer Thumbnails for finding a person carrying an object:P02
    - Immediate Visuals
        * E56:Prefer Thumbnails for immediate visual of what is happening for an event:P04, P05
        * E22:Prefer Thumbnails over the event blocks + Density graph:P02
    - Control over intelligent features
        * E40:Prefer Density Graph over Event Blocks:P01: More confusion than the density graphs, density graphs give more control and less automation, make sure nothing is being cut out:P01
    - Confirmation
        * E55:Prefer Thumbnails for checking if an event is miscategorized or not::P04
- Suggested Video Reviewing Software Features
    - Video Traversal
        * E41:Feature to grab the slider and move it along with the video:P01
    - Temporal Cues
        * E25:Feature to change the opacity of event when a user has viewed it / Visual indicator so a user does not reclick an already viewed event if they do not want to:P02
    - Fast Forwarding

        * E37:Feature to increase the speed that the videos play at/Fast forward:P01, P03

    - Advanced Filters
        * E26:Advanced Identification / Filtering:Look for a Red car, or Blue truck instead of just person / vehicle:P02

- Limitations with current Features
    - Thumbnails misguiding users
        * E59:Thumbnails may lead to users assuming events are not the ones they are looking for:P04
        * E24:May make a mistake from relying on the image since one frame of a thumbnail is not enough for a whole event:P02
        * E50:Thumbnails being only the first frame are not as helpful:P03
    - Navigation
        * E43:Need to be able to move timeline in small chunks to notice subtle changes:P01
    - Massive amount of material to review
        * E58:Density graph may be overwhelming with lots of events:P04
        * E49:Event Blocks and Density Graph give too much information/repetitive:P03
    - Temporal Cues
        * Negative:E57:Do not prefer the event blocks because it is harder to see when events start and end:P04

Appendix B

Question Pro Survey

DUET User Study Survey

Title of research study: DUET: Dual United Event-Based Timelines (IRB # 24-559)

Principal Investigator:

Name: Sang Won Lee
Phone: 540-231-4857
Email: sangwonlee@vt.edu

Other study contact(s):

Name: Andy Luu
Phone: 571-332-0057
Email: andyluu@vt.edu

Key Information: The following is a short summary of this study to help you decide whether or not to be a part of this study. More detailed information is listed later on in this form.

The purpose of this study is to get a better understanding of how security camera operators currently search for specific events over long periods of time and how the new system that we designed can assist them in the search process. This study is a one-hour-long session where we will simply ask you about your experiences and opinions with viewing long videos or surveillance video footage and then have you answer some questions by completing tasks with the interface that we developed. The following sections contain more detailed information.

Detailed Information: The following is more detailed information about this study in addition to the information listed above.

Who can I talk to?

If you have questions, concerns, or complaints, or think the research has hurt you, talk to the research team at andyluu@vt.edu

This research has been reviewed and approved by the Virginia Tech Institutional Review Board (IRB). You may communicate with them at 540-231-3732 or irb@vt.edu if:

- You have questions about your rights as a research subject
- Your questions, concerns, or complaints are not being answered by the research team
- You cannot reach the research team
- You want to talk to someone besides the research team to provide feedback about this research

How many people will be studied?

We plan to include about 15-30 people in this research study.

What happens if I say yes, I want to be in this research?

- The researchers will provide a survey that should take about one hour to complete.
- During the survey, we will ask you about your experiences and opinions with viewing long videos or surveillance video footage.
- Afterward, you will be asked to answer questions by completing a set of tasks using a system that we created that has publicly available security camera footage with events throughout that footage and five different timeline visualizations to navigate through those events in the video.

What happens if I say yes, but I change my mind later?

You can leave the research at any time, for any reason, and it will not be held against you. If you decide to leave the research, contact the investigator so that the investigator can delete any of the information about you that we collected.

Is there any way being in this study could be bad for me? (Detailed Risks)
There are no known risks to participating in this study.

What happens to the information collected for the research?
We will make every effort to limit the use and disclosure of your personal information, including research study and medical records, only to people who have a need to review this information. We cannot promise complete confidentiality. Organizations that may inspect and copy your information include the IRB, Human Research Protection Program, and other authorized representatives of Virginia Tech.The results of this research study may be presented in summary form at conferences, in presentations, reports to the sponsor, academic papers, and as part of a thesis/dissertation.

Can I be removed from the research without my OK?
The person in charge of the research study or the sponsor can remove you from the research study without your approval. Possible reasons for removal include: if we have reason to believe you do not meet the eligibility requirements (i.e., are not a security camera operator or related occupation that views long hours of video footage with event-based data)

What else do I need to know?
If you agree to take part in this research study, the research team will enter the email addresses into a random drawing and compensation will be given to approximately ten percent of participants as a $10 gift card. The odds of being selected will be maintained as 1 in 10. The participants will be contacted separately.

---

* Do you consent to participate in this research study?

- ◯ I consent to participate in this research.

- ◯ I do NOT consent to participate in this research.

* Please enter your email address.

* What gender do you identify as?

- ◯ Male
- ◯ Female
- ◯ Non-Binary
- ◯ Prefer not to disclose
- ◯ Other

* How old are you?

- ◯ Under 18
- ◯ 18-24
- ◯ 25-34
- ◯ 35-44
- ◯ 45-54
- ◯ 55-64
- ◯ Above 64

* What race or ethnicity describes you?

- ◯ American Indian or Alaskan Native

- Asian/Pacific Islander
- Black or African American
- Hispanic
- White/Caucasian
- Other

* Have you ever viewed surveillance video footage, if so, how often do you view surveillance video footage (Examples: footage of a site (home, street, retail stores, building, bird watchers, etc.))?

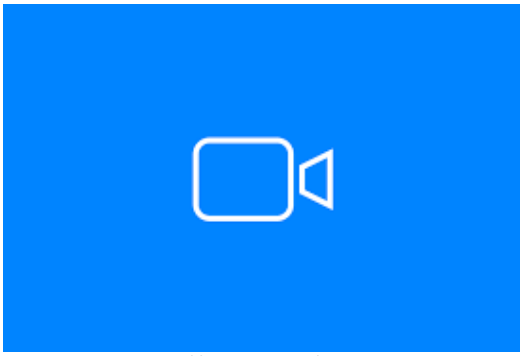
- Daily
- Weekly
- Monthly
- Yearly
- Never

If you have previously viewed surveillance video footage, what are the typical events or incidents are the ones that you look for when you view the footage?

If you have previously viewed surveillance video footage, what information do you typically use and are given as clues to spot the events? (time window, profile, or nothing)?

If you have previously viewed surveillance video footage, what software/products/systems did you use to view the surveillance video footage?

In this study, you will be asked to use a video player with five different timeline visualizations to complete tasks in a random order. The following tutorial describes the five timeline visualizations (Timeline 1, 2, 3, 4, and 5). Please watch the video carefully and try to understand how each timeline visualization works before you move on to the following questions.



URL : https://youtu.be/9EorQV75Uy4

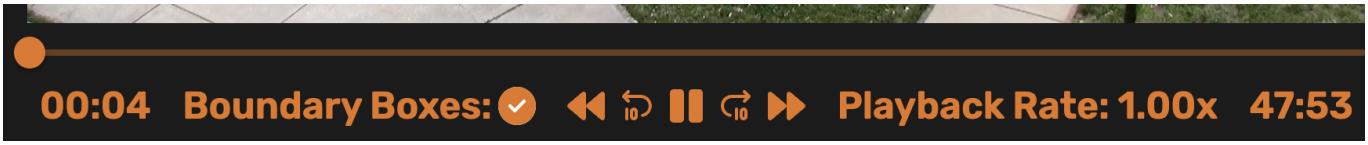
* Please watch this quick tutorial about how to use the five different timeline visualizations and confirm that you have watched the entire tutorial. Here is a link to each of the timeline visualizations if you want to explore any of them further:

Timeline 1 (Base Interface)
Timeline 2 (Density Graph)
Timeline 3 (Event Blocks)
Timeline 4 (Density Graph + Event Blocks)
Timeline 5 (Event Thumbnails)

If you have any troubles viewing the videos here is a link to the YouTube Channel: https://youtube.com/@andyluu8958

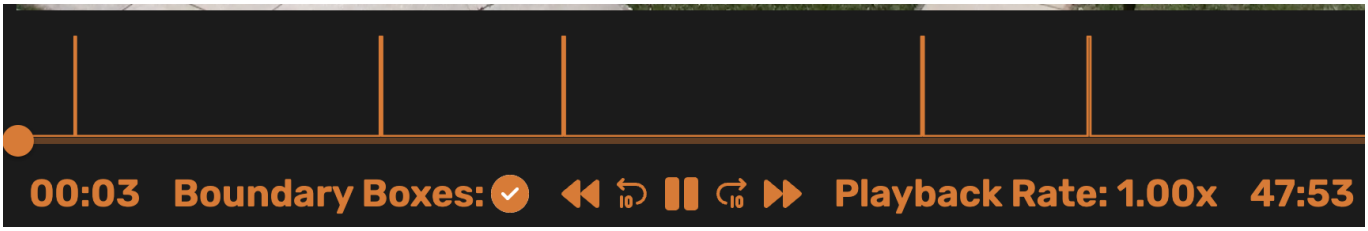
- ◯ I have watched the entire tutorial.

* Please select the timeline visualization that is shown in the image below.



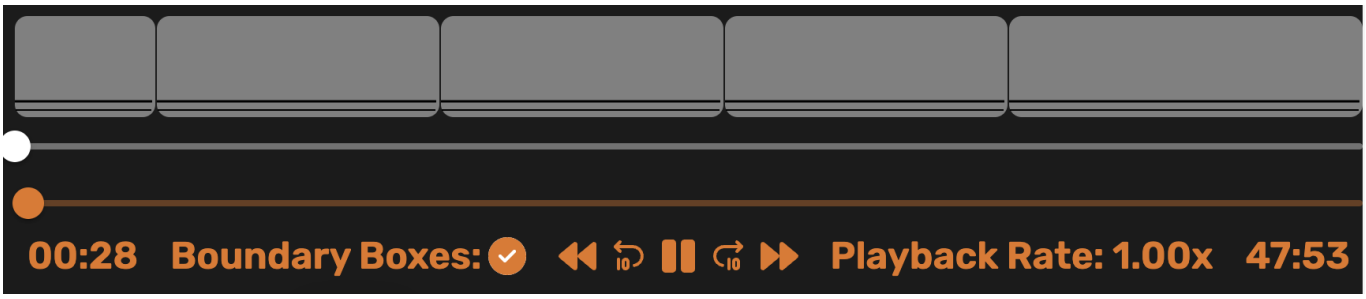
- Timeline 1 (Base Interface)
- Timeline 2 (Density Graph)
- Timeline 3 (Event Blocks)
- Timeline 4 (Density Graph + Event Blocks)
- Timeline 5 (Event Thumbnails)

* Please select the timeline visualization that is shown in the image below.



- Timeline 1 (Base Interface)
- Timeline 2 (Density Graph)
- Timeline 3 (Event Blocks)
- Timeline 4 (Density Graph + Event Blocks)
- Timeline 5 (Event Thumbnails)

* Please select the timeline visualization that is shown in the image below.

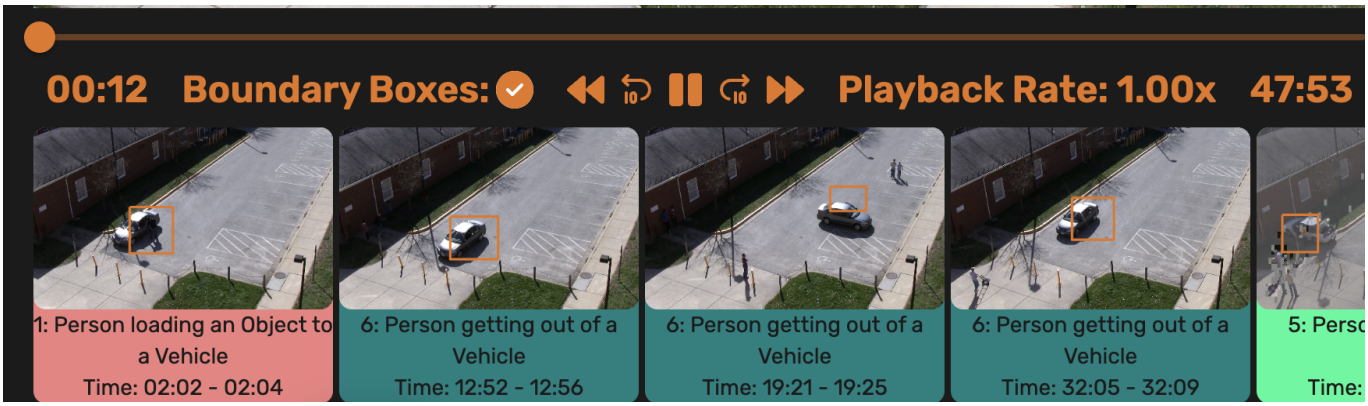


- Timeline 1 (Base Interface)
- Timeline 2 (Density Graph)
- Timeline 3 (Event Blocks)
- Timeline 4 (Density Graph + Event Blocks)
- Timeline 5 (Event Thumbnails)

* Please select the timeline visualization that is shown in the image below.

- Timeline 1 (Base Interface)
- Timeline 2 (Density Graph)
- Timeline 3 (Event Blocks)
- Timeline 4 (Density Graph + Event Blocks)
- Timeline 5 (Event Thumbnails)

* Please select the timeline visualization that is shown in the image below.



- ◯ Timeline 1 (Base Interface)
- ◯ Timeline 2 (Density Graph)
- ◯ Timeline 3 (Event Blocks)
- ◯ Timeline 4 (Density Graph + Event Blocks)
- ◯ Timeline 5 (Event Thumbnails)

Please open this link in a different tab:

https://duetimelines.vercel.app/?videotype=VIRAT_S_0002&password=${timelinesUsed}

If the interface seems a little small, please resize your window until it is as large as possible. Once you open the web page and briefly play with the timeline visualization, please move to the next section to complete the tasks.

* Please enter the password displayed on the webpage.

* Task: The driver of this car gets out of the car and has a conversation with another person.

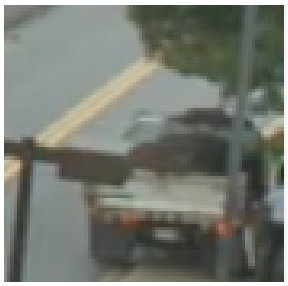
What is the car that the other person gets into later?

Car 1

Car 3

Car 4

Car 5



Car 2

I could not find the answer within five minutes.

---

* Task: During which timeframe do you see this man unloading a blue shoulder bag from his car?

- ◯ 11:00 – 14:00
- ◯ 18:00 – 20:00
- ◯ 20:00 – 22:00
- ◯ 32:00 – 35:00
- ◯ 43:00 – 45:00
- ◯ 13:00 – 15:00
- ◯ 16:00 – 18:00
- ◯ 22:00 – 23:00
- ◯ 34:00 – 35:00
- ◯ 45:00 – 47:00

- ◯ I could not find the answer within five minutes.

* Please select the timeline visualization that you used to complete the previous tasks.

- ◯ Timeline 1 (Base Interface)
- ◯ Timeline 2 (Density Graph)
- ◯ Timeline 3 (Event Blocks)
- ◯ Timeline 4 (Density Graph + Event Blocks)
- ◯ Timeline 5 (Event Thumbnails)

* System Usability Scale (SUS)

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
I think that I would like to use this UI frequently.	◯	◯	◯	◯	◯
I found the UI unnecessarily complex.	◯	◯	◯	◯	◯
I thought the UI was easy to use.	◯	◯	◯	◯	◯
I think that I would need the support of a technical person to be able to use this UI.	◯	◯	◯	◯	◯
I found the various functions in this UI were well integrated.	◯	◯	◯	◯	◯
I thought there was too much inconsistency in this UI.	◯	◯	◯	◯	◯
I would imagine that most people would learn to use this UI very quickly.	◯	◯	◯	◯	◯
I found the UI very cumbersome to use.	◯	◯	◯	◯	◯
I felt very confident using the UI.	◯	◯	◯	◯	◯
I needed to learn a lot of things before I could get going with this UI.	◯	◯	◯	◯	◯

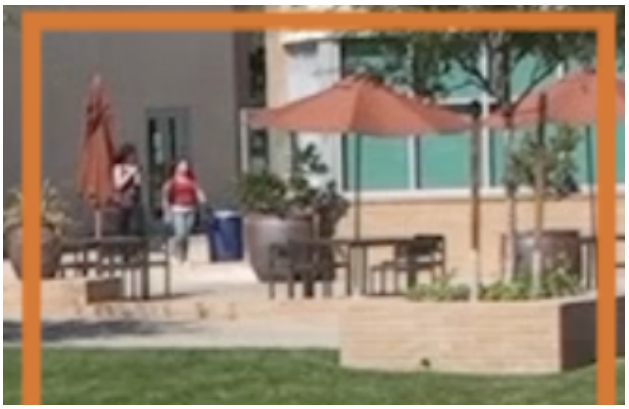
Please open this link in a different tab:

https://duetimelines.vercel.app/?videotype=VIRAT_S_0100&password=${timelinesUsed}

If the interface seems a little small, please resize your window until it is as large as possible. Once you open the web page and briefly play with the timeline visualization, please move to the next section to complete the tasks.

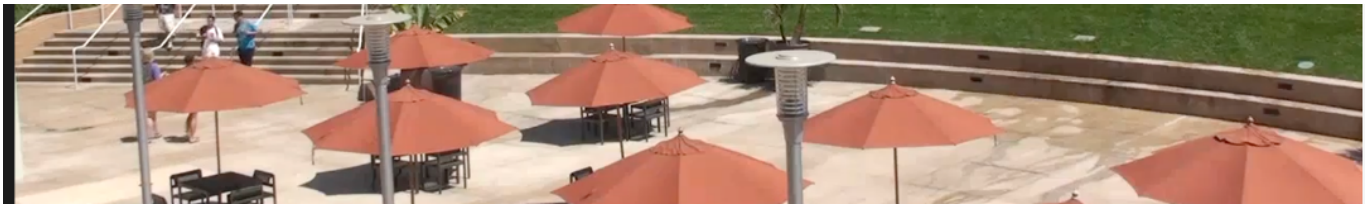
* Please enter the password displayed on the webpage.

* Task: How long do these two people stay in the building?



- ◯ Less than eight minutes
- ◯ Less than fourteen minutes
- ◯ Less than sixteen minutes
- ◯ Less than two minutes
- ◯ Less than twelve minutes
- ◯ I could not find the answer within five minutes.

* Task: During which timeframe do you see a man in a white shirt run across the camera in this area?



- ◯ 48:20 – 48:30
- ◯ 28:20 – 28:30
- ◯ 49:20 – 49:30
- ◯ 29:20 – 29:30
- ◯ 38:20 – 38:30
- ◯ 47:20 – 47:30
- ◯ 27:20 – 27:30
- ◯ 46:20 – 47:30
- ◯ 26:20 – 27:30
- ◯ 34:20 – 35:30
- ◯ I could not find the answer within five minutes.

* Please select the timeline visualization that you used to complete the previous tasks.

- ◯ Timeline 1 (Base Interface)
- ◯ Timeline 2 (Density Graph)
- ◯ Timeline 3 (Event Blocks)
- ◯ Timeline 4 (Density Graph + Event Blocks)
- ◯ Timeline 5 (Event Thumbnails)

* System Usability Scale (SUS)

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
I think that I would like to use this UI frequently.	◯	◯	◯	◯	◯
I found the UI unnecessarily complex.	◯	◯	◯	◯	◯
I thought the UI was easy to use.	◯	◯	◯	◯	◯
I think that I would need the support of a technical person to be able to use this UI.	◯	◯	◯	◯	◯
I found the various functions in this UI were well integrated.	◯	◯	◯	◯	◯
I thought there was too much inconsistency in this UI.	◯	◯	◯	◯	◯
I would imagine that most people would learn to use this UI very quickly.	◯	◯	◯	◯	◯
I found the UI very cumbersome to use.	◯	◯	◯	◯	◯
I felt very confident using the UI.	◯	◯	◯	◯	◯
I needed to learn a lot of things before I could get going with this UI.	◯	◯	◯	◯	◯

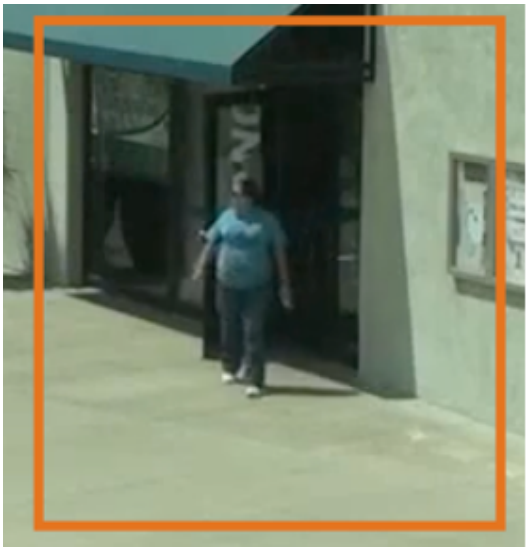
Please open this link in a different tab:

https://duetimelines.vercel.app/?videotype=VIRAT_S_0102&password=${timelinesUsed}

If the interface seems a little small, please resize your window until it is as large as possible. Once you open the web page and briefly play with the timeline visualization, please move to the next section to complete the tasks.

* Please enter the password displayed on the webpage.

* Task: How long does this CHICANO employee take a break for?



- ◯ About one minute
- ◯ About two minutes
- ◯ About four minutes
- ◯ About three minutes
- ◯ About five minutes
- ◯ I could not find the answer within five minutes.

* Task: During which timeframe does the owner of this golf cart return and drive away?

- ◯ 21:00 – 22:00
- ◯ 22:00 – 23:00
- ◯ 19:00 – 20:00
- ◯ 20:00 – 21:00
- ◯ 40:00 – 41:00
- ◯ 26:00 – 27:00
- ◯ 27:00 – 28:00
- ◯ 24:00 – 25:00
- ◯ 25:00 – 26:00
- ◯ 47:00 – 48:00
- ◯ I could not find the answer within five minutes.

---

* Please select the timeline visualization that you used to complete the previous tasks.

- ◯ Timeline 1 (Base Interface)
- ◯ Timeline 2 (Density Graph)
- ◯ Timeline 3 (Event Blocks)

- ◯ Timeline 4 (Density Graph + Event Blocks)
- ◯ Timeline 5 (Event Thumbnails)

* System Usability Scale (SUS)

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
I think that I would like to use this UI frequently.	◯	◯	◯	◯	◯
I found the UI unnecessarily complex.	◯	◯	◯	◯	◯
I thought the UI was easy to use.	◯	◯	◯	◯	◯
I think that I would need the support of a technical person to be able to use this UI.	◯	◯	◯	◯	◯
I found the various functions in this UI were well integrated.	◯	◯	◯	◯	◯
I thought there was too much inconsistency in this UI.	◯	◯	◯	◯	◯
I would imagine that most people would learn to use this UI very quickly.	◯	◯	◯	◯	◯
I found the UI very cumbersome to use.	◯	◯	◯	◯	◯
I felt very confident using the UI.	◯	◯	◯	◯	◯
I needed to learn a lot of things before I could get going with this UI.	◯	◯	◯	◯	◯

Please open this link in a different tab:

https://duetimelines.vercel.app/?videotype=VIRAT_S_0400&password=${timelinesUsed}

If the interface seems a little small, please resize your window until it is as

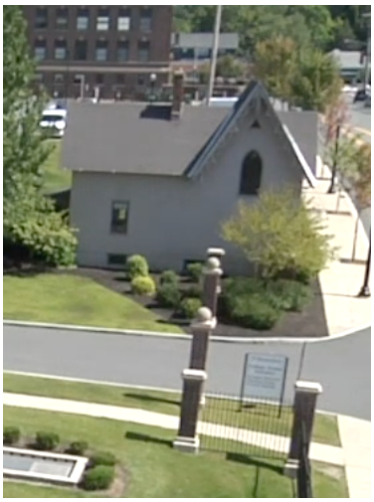
large as possible. Once you open the web page and briefly play with the timeline visualization, please move to the next section to complete the tasks.

* Please enter the password displayed on the webpage.

* Task: How long does the blue pickup truck remain parked after two men in white shirts begin unloading objects from the trunk of the blue pickup truck?

- ◯ About six minutes
- ◯ About nine minutes
- ◯ About twelve minutes
- ◯ About twenty minutes
- ◯ About thirty minutes
- ◯ I could not find the answer within five minutes.

* Task: During which timeframe does a couple run across the screen starting from this house carrying a pizza?



- ◯ 50:00 – 51:00
- ◯ 20:00 – 21:00
- ◯ 30:00 – 31:00
- ◯ 40:00 – 41:00
- ◯ 10:00 – 11:00
- ◯ 55:00 – 56:00
- ◯ 25:00 – 26:00
- ◯ 35:00 – 36:00
- ◯ 45:00 – 46:00
- ◯ 15:00 – 16:00
- ◯ I could not find the answer within five minutes.

---

* Please select the timeline visualization that you used to complete the previous tasks.

- ◯ Timeline 1 (Base Interface)
- ◯ Timeline 2 (Density Graph)
- ◯ Timeline 3 (Event Blocks)
- ◯ Timeline 4 (Density Graph + Event Blocks)

◯ Timeline 5 (Event Thumbnails)

* System Usability Scale (SUS)

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
I think that I would like to use this UI frequently.	◯	◯	◯	◯	◯
I found the UI unnecessarily complex.	◯	◯	◯	◯	◯
I thought the UI was easy to use.	◯	◯	◯	◯	◯
I think that I would need the support of a technical person to be able to use this UI.	◯	◯	◯	◯	◯
I found the various functions in this UI were well integrated.	◯	◯	◯	◯	◯
I thought there was too much inconsistency in this UI.	◯	◯	◯	◯	◯
I would imagine that most people would learn to use this UI very quickly.	◯	◯	◯	◯	◯
I found the UI very cumbersome to use.	◯	◯	◯	◯	◯
I felt very confident using the UI.	◯	◯	◯	◯	◯
I needed to learn a lot of things before I could get going with this UI.	◯	◯	◯	◯	◯

Please open this link in a different tab:

https://duetimelines.vercel.app/?videotype=VIRAT_S_0500&password=${timelinesUsed}

If the interface seems a little small, please resize your window until it is as large as possible. Once you open the web page and briefly play with the timeline visualization, please move to the next section to complete the tasks.

* Please enter the password displayed on the webpage.

* Task: Order the following three events:

Drag your choices here to rank them

First Event

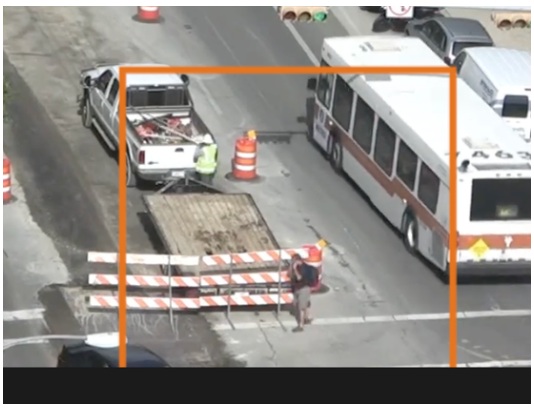
Last Event

A person carrying a backpack rode a bike on the pedestrian sidewalk.

A person wearing pants ran to catch the green light while crossing the street.

A construction worker waved their hands to their colleagues.

* Task: What is the object unloaded from the white truck with a flatboard trailer?



- A slow-down sign
- An orange construction traffic barrel (cylinder-shaped sign)

- ◯ A backpack
- ◯ A safety helmet
- ◯ A safety vest
- ◯ I could not find the answer within five minutes.

---

* Please select the timeline visualization that you used to complete the previous tasks.

- ◯ Timeline 1 (Base Interface)
- ◯ Timeline 2 (Density Graph)
- ◯ Timeline 3 (Event Blocks)
- ◯ Timeline 4 (Density Graph + Event Blocks)
- ◯ Timeline 5 (Event Thumbnails)

* System Usability Scale (SUS)

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
I think that I would like to use this UI frequently.	◯	◯	◯	◯	◯
I found the UI unnecessarily complex.	◯	◯	◯	◯	◯
I thought the UI was easy to use.	◯	◯	◯	◯	◯
I think that I would need the support of a technical person to be able to use this UI.	◯	◯	◯	◯	◯
I found the various functions in this UI were well integrated.	◯	◯	◯	◯	◯
I thought there was too much inconsistency in this UI.	◯	◯	◯	◯	◯
I would imagine that most people would learn to use this UI very quickly.	◯	◯	◯	◯	◯
I found the UI very cumbersome to use.	◯	◯	◯	◯	◯

I felt very confident using the UI.	○	○	○	○	○
I needed to learn a lot of things before I could get going with this UI.	○	○	○	○	○

* Please rank (1-5) the following in order of which timeline visualization you liked the most, if you need a refresher, here are links for each of the timeline visualizations:

Timeline 1 (Base Interface)
Timeline 2 (Density Graph)
Timeline 3 (Event Blocks)
Timeline 4 (Density Graph + Event Blocks)
Timeline 5 (Event Thumbnails)

Drag your choices here to rank them

Timeline Visualization I liked the Most ↑

Timeline Visualization I liked the Least ↓

- Timeline 1 (Base Interface)
- Timeline 2 (Density Graph)
- Timeline 3 (Event Blocks)
- Timeline 4 (Density Graph + Event Blocks)
- Timeline 5 (Event Thumbnails)

* Please explain your preference for the ranked choices you made in the previous question.

* Are there any other features you think would be nice to implement in our current system?

* Do you believe the tasks we had you perform were realistic to those you look for when you browse for events in surveillence camera footage?

* Do you have any questions for us?

Bibliography

[1] Bryam Abril, Juan Diego Jara, Patricio Cuzco, and Pablo Gallegos. 2020. Development and Design of a Unified Remote Video Surveillance System for Homes, Using Free Software Tools. In *Proceedings of the 3rd International Conference on Vision, Image and Signal Processing* (New York City, NY, USA, 2020-05-25) *(ICVISP 2019)*. Association for Computing Machinery, New York City, NY, USA, 1–6. https://doi.org/10.1145/3387168.3387194

[2] María José Abásolo and Carlos Sebastián Castañeda. 2021. Improving Usability and Intrusion Detection Alerts in a Home Video Surveillance System. In *Computer Science – CACIC 2020* (Silver Spring, MD, USA, 2021), Patricia Pesado and Jorge Eterovic (Eds.). Springer International Publishing, Silver Spring, MD, USA, 350–364. https://doi.org/10.1007/978-3-030-75836-3_24

[3] Aaron bangor, CHFP, Philip Kortum, and James Miller. 2009. *Determining What Individual SUS Scores Mean: Adding an Adjective Rating Scale - JUX*. JUX - The Journal of User Experience. https://uxpajournal.org/determining-what-individual-sus-scores-mean-adding-an-adjective-rating-scale/

[4] Matthew Brehmer, Bongshin Lee, Benjamin Bach, Nathalie Henry Riche, and Tamara Munzner. 2017. Timelines Revisited: A Design Space and Considerations for Expressive Storytelling. *IEEE Transactions on Visualization and Computer Graphics* 23, 9 (2017), 2151–2164. https://doi.org/10.1109/TVCG.2016.2614803

[5] C. Clavel, T. Ehrette, and G. Richard. 2005. Events Detection for an Audio-Based Surveillance System. In *2005 IEEE International Conference on Multimedia and Expo*

(2005-07). IEEE, New York City, NY, USA, 1306–1309. https://doi.org/10.1109/ICME.2005.1521669

[6] Kelly Gordon. 2020. *5 Principles of Visual Design in UX*. Nielsen Norman Group. https://www.nngroup.com/articles/principles-visual-design/

[7] Cathal Gurrin, Alan F. Smeaton, and Aiden R. Doherty. 2014. LifeLogging: Personal Big Data. *Foundations and Trends® in Information Retrieval* 8, 1 (2014), 1–125. https://doi.org/10.1561/1500000033

[8] A. Harma, M.F. McKinney, and J. Skowronek. 2005. Automatic Surveillance of the Acoustic Activity in Our Living Environment. In *2005 IEEE International Conference on Multimedia and Expo* (2005-07). IEEE, New York City, NY, USA, 4 pp.–. https://doi.org/10.1109/ICME.2005.1521503

[9] Orland Hoeber and Joshua Gorner. 2009. BrowseLine: 2D Timeline Visualization of Web Browsing Histories. In *2009 13th International Conference Information Visualisation* (2009-07). IEEE, New York City, NY, USA, 156–161. https://doi.org/10.1109/IV.2009.19

[10] Jeff Sauro. 2018. *5 Ways to Interpret a SUS Score – MeasuringU*. MeasuringU. https://measuringu.com/interpret-sus-score/

[11] Priti Vasant Kale and Samidha Dwivedi Sharma. 2014. A Review of Securing Home Using Video Surveillance. In *International Journal of Science and Research (IJSR) Vol 3* (2014). International Journal of Science and Research (IJSR), Ahmedabad, Gujarat, India, 1150 – 1154. https://api.semanticscholar.org/CorpusID:53309198

[12] K. A. Kelly. 2023. The Ring Video Doorbell and the Entry of Amazon into the Smart

Home: Implications for Consumer-Initiated Surveillance. *Journal of Consumer Policy* 46, 1 (2023), 95–104. https://doi.org/10.1007/s10603-022-09534-3

[13] Terra R. Kelly, Pranav S. Pandit, Nicole Carion, Devin F. Dombrowski, Krysta H. Rogers, Stella C. McMillin, Deana L. Clifford, Anthony Riberi, Michael H. Ziccardi, Erica L. Donnelly-Greenan, and Christine K. Johnson. 2021. Early Detection of Wildlife Morbidity and Mortality through an Event-Based Surveillance System. *Proceedings of the Royal Society B: Biological Sciences* 288, 1954 (2021), 20210974. https://doi.org/10.1098/rspb.2021.0974

[14] Juho Kim, Philip J. Guo, Carrie J. Cai, Shang-Wen (Daniel) Li, Krzysztof Z. Gajos, and Robert C. Miller. 2014. Data-Driven Interaction Techniques for Improving Navigation of Educational Videos. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology* (New York City, NY, USA, 2014-10-05) *(UIST '14)*. Association for Computing Machinery, New York City, NY, USA, 563–572. https://doi.org/10.1145/2642918.2647389

[15] B. Li and M. Ibrahim Sezan. 2001. Event Detection and Summarization in Sports Video. In *Proceedings IEEE Workshop on Content-Based Access of Image and Video Libraries (CBAIVL 2001)* (2001-12). IEEE, New York City, NY, USA, 132–138. https://doi.org/10.1109/IVL.2001.990867

[16] Richard Lipka. 2016. Scalable Timeline Visualization. In *2016 9th International Conference on Human System Interactions (HSI)* (2016-07). IEEE, New York City, NY, USA, 150–156. https://doi.org/10.1109/HSI.2016.7529624

[17] Andrea Macarulla Rodriguez, Christian Tiberius, Roel Van Bree, and Zeno Geradts. 2018. Google Timeline Accuracy Assessment and Error Prediction. *Forensic Sciences Research* 3, 3 (2018), 240–255. https://doi.org/10.1080/20961790.2018.1509187

[18] Brett Milash, Catherine Plaisant, and Anne Rose. 1996. LifeLines: Visualizing Personal Histories. In *Conference Companion on Human Factors in Computing Systems Common Ground - CHI '96* (Vancouver, British Columbia, Canada, 1996). ACM Press, Vancouver, British Columbia, Canada, 392–393. https://doi.org/10.1145/257089.257391

[19] Alessandra Mileo, Peng Wang, and Alan Smeaton. 2014. Semantically Enhancing Multimedia Lifelog Events. In *Advances in Multimedia Information Processing – PCM 2014* (Silver Spring, MD, USA, 2014-12-01). Springer International Publishing, Silver Spring, MD, USA, 163–172. https://doi.org/10.1007/978-3-319-13168-9_17

[20] Phong H. Nguyen, Kai Xu, Rick Walker, and B.L. William Wong. 2014. SchemaLine: Timeline Visualization for Sensemaking. In *2014 18th International Conference on Information Visualisation* (2014-07). IEEE, New York City, NY, USA, 225–233. https://doi.org/10.1109/IV.2014.14

[21] Sangmin Oh, Anthony Hoogs, Amitha Perera, Naresh Cuntoor, Chia-Chih Chen, Jong Taek Lee, Saurajit Mukherjee, J. K. Aggarwal, Hyungtae Lee, Larry Davis, Eran Swears, Xioyang Wang, Qiang Ji, Kishore Reddy, Mubarak Shah, Carl Vondrick, Hamed Pirsiavash, Deva Ramanan, Jenny Yuen, Antonio Torralba, Bi Song, Anesco Fong, Amit Roy-Chowdhury, and Mita Desai. 2011. A Large-Scale Benchmark Dataset for Event Recognition in Surveillance Video. In *CVPR 2011* (Colorado Springs, CO, USA, 2011-06). IEEE, Colorado Springs, CO, USA, 3153–3160. https://doi.org/10.1109/CVPR.2011.5995586

[22] Paulina Piasek, Kate Irving, and Alan F. Smeaton. 2014. Using lifelogging to help construct the identity of people with dementia. In *Proceedings of the Irish Human Computer Interaction Conference 2014* (2014-09-02). DORAS | DCU, Dublin, Ireland, 1–8. https://api.semanticscholar.org/CorpusID:15535291

[23] Christian Rohrer. 2022. *When to Use Which User-Experience Research Methods*. Nielsen Norman Group. https://www.nngroup.com/articles/which-ux-research-methods/

[24] Juan Carlos SanMiguel, José M. Martinez, and Álvaro Garcia. 2009. An Ontology for Event Detection and Its Application in Surveillance Video. In *2009 Sixth IEEE International Conference on Advanced Video and Signal Based Surveillance* (2009-09). IEEE, New York City, NY, USA, 220–225. https://doi.org/10.1109/AVSS.2009.28

[25] Margaret L. Satterthwaite, Enrico Bertini, Oded Nov, John Emerson, and Anshul Vikram Pandey. 2018. *Event-Based Data - Data Types - Data Visualization and Human Rights*. Data Visualization and Human Rights. http://visualizingrights.org/kit/data/event-based-data.html

[26] Seong-Wook Joo and R. Chellappa. 2006. Attribute Grammar-Based Event Recognition and Anomaly Detection. In *2006 Conference on Computer Vision and Pattern Recognition Workshop (CVPRW'06)* (New York City, NY, USA, 2006). IEEE, New York City, NY, USA, 107–107. https://doi.org/10.1109/CVPRW.2006.32

[27] Christos Tzelepis, Zhigang Ma, Vasileios Mezaris, Bogdan Ionescu, Ioannis Kompatsiaris, Giulia Boato, Nicu Sebe, and Shuicheng Yan. 2016. Event-Based Media Processing and Analysis: A Survey of the Literature. *Image and Vision Computing* 53 (2016), 3–19. https://doi.org/10.1016/j.imavis.2016.05.005

[28] Xin Yan, Kong Wah Wan, Qi Tian, Mun Kew Leong, and Ping Xiao. 2004. Two Dimensional Timeline and Its Application to Conversant Media System. In *2004 IEEE International Conference on Multimedia and Expo (ICME) (IEEE Cat. No.04TH8763)* (2004-06), Vol. 1. IEEE, New York City, NY, USA, 507–510 Vol.1. https://doi.org/10.1109/ICME.2004.1394240

[29] YouTube. 2005. *YouTube*. YouTube. https://www.youtube.com/

[30] Jianfeng Zheng, Henry X. Liu, Steve Misgen, Kevin Schwartz, Bob Green, and Mike Anderson. 2014. Use of Event-Based Traffic Data in Generating Time–Space Diagrams for Evaluation of Signal Coordination. *Transportation Research Record* 2439, 1 (2014), 94–104. https://doi.org/10.3141/2439-09