2009-12-09

Interactive Football Summarization

Brandon B. Moon
Brigham Young University - Provo

Interactive Football Summarization

Brandon Moon

A thesis submitted to the faculty of
Brigham Young University
In partial fulfillment of the requirements for the degree of

Master of Science

Dan R. Olsen, Jr., Chair
Parris K. Egbert
Daniel Zappala

Department of Computer Science

Brigham Young University

April 2010

ABSTRACT

Interactive Football Summarization

Brandon Moon

Department of Computer Science

Master of Science

Football fans do not have the time to watch every game in its entirety and need an effective solution that summarizes them the story of the game. Human-generated summaries are often too short, requiring time and resources to create. We utilize the advantages of Interactive TV to create an automatic football summarization service that is cohesive, provides context, covers the necessary plays, and is concise. First, we construct a degree of interest function that ranks each play based on detailed, play-by-play game events as well as viewing statistics collected from an interactive viewing environment. This allows us to select the plays that are important to the game as well as those that are interesting to the viewer. Second, we create a visual transition that shows the progress of the ball whenever plays are skipped, allowing the viewer to understand the context of each play within the summary. Third, we enable interactive controls that allow viewers to manipulate the summary and delve deeper into the actual game whenever they wish. We validate our solution through two user studies—one to ensure that our degree of interest function selects the plays that are most interesting to the viewer, and the other to show that our transitions and interactive controls provide a better understanding of the game. We conclude that our summary solution is effective at conveying the story of a football game.

Table of Contents

Chapter 1 – Introduction

Internet Television is a recent shift from traditional transport mechanisms to the Internet. Much as telephone, radio, and other communication technologies have improved from this change, Internet Television provides us with a number of benefits. Traditional broadcast television is rigid and inflexible. Viewers must tune in at the right time and the right channel in order to watch the show they are interested in. With Internet Television, the viewer is in control. The flexible and individualized nature of the Internet allows a more personal selection of content to be delivered on demand. He or she can then select what show to watch and when. While Internet Television is still fairly new, we see an increasing number of broadcast companies providing this option, and it is likely that this shift will continue.

There are many other benefits from using the Internet to distribute television content. Because of the dynamic, personalized nature of content delivery over the Internet, we can provide a whole new layer of interaction that enhances the experience. With Interactive Television (ITV), the viewer is no longer limited to *what* he or she can watch; rather the viewer can now choose *how* to watch the program. Cooking and other "how to" shows can be enriched with optional segments for certain topics. News can be explored, allowing viewers to dig deeper into current events and find topics more interesting to them. Many possibilities are opened up that weren't originally possible with broadcast television.

One of the more interesting possibilities made available with ITV is that of enhancing sports broadcasting. While watching traditional broadcast sports, a viewer is limited to the replays, camera angles, and information which the broadcasting company chooses to show. The viewer does not have the freedom to review particularly interesting plays, nor can the viewer decide from which camera angle to view a replay. ITV allows us to put these options into the hands of sports fans. They can customize the broadcast to match what interests them. Viewers can also watch previous games that they missed, or even summaries of these games. They are no longer limited to watching on the broadcasting company's schedule.

Of particular interest to us is that of watching football through ITV. Most football fans do not have the time to watch each football game of every week, nor are they interested in doing so. Often they would prefer to see a summary of the game that focuses on key events, scoring plays, and other important details from the game. Hand-crafted highlight reels are generally too short and do not include enough detail. In the end, they give no sense of the story and flow of a game, and viewers are left confused about what actually happened. With the recent trend towards ITV, we believe that there is a place for a dynamic, variable-length summarization service that allows the user to understand the story of the game.

For example, most football games in the NFL run on Sundays, with several games showing at the same time (See Figure 1.1). This prevents viewers from seeing many of the games that interest them. While they can easily discover the final score, many are more interested in what actually happened during the game. Through a quality summarization service, fans can watch a 5 minute summary of each game and be caught up in under an hour, or they can watch a more detailed 15 minute summary of just the games which interest them. Because the summary focuses on telling the story of the game, viewers will have a greater understanding of the outcome. The next time they run into another football fan, they will easily maintain a conversation about any game that weekend without feeling lost, as all the main points were covered in the game summary.

Week 1		
Thu, Sep 10	**Stadium**	**Time (ET)**
TEN @ PIT	Heinz Field	8:30 PM
Sun, Sep 13	**Stadium**	**Time (ET)**
MIA @ ATL	Georgia Dome	1:00 PM
KC @ BAL	M&T Bank Stadium	1:00 PM
PHI @ CAR	Bank of America Stadium	1:00 PM
MIN @ CLE	Cleveland Browns Stadium	1:00 PM
JAC @ IND	Lucas Oil Stadium	1:00 PM
DAL @ TB	Raymond James Stadium	1:00 PM
DET @ NO	Superdome	1:00 PM
NYJ @ HOU	Reliant Stadium	1:00 PM
DEN @ CIN	Paul Brown Stadium	1:00 PM
SF @ ARI	University of Phoenix Stadium	4:15 PM
WAS @ NYG	Giants Stadium	4:15 PM
STL @ SEA	Qwest Field	4:15 PM
CHI @ GB	Lambeau Field	8:20 PM
Mon, Sep 14	**Stadium**	**Time (ET)**
BUF @ NE	Gillette Stadium	7:00 PM
SD @ OAK	Oakland Coliseum	10:15 PM

Figure 1.1: Example NFL Schedule, obtained from http://www.nfl.com

Video summarization is inherently a difficult task. Because a summary reduces the amount of original content, we have to decide what to include and what to leave out. The difficulty, then, is in choosing which plays would interest the viewer and preserve the story of the original video—and which would not. This process can be extremely subjective. Should we fail to include the right content or include content that is not important to the original story, the viewer would become confused and/or disoriented and could come away with the wrong perception of the original video.

It is important, then, to determine what creates a good summary. More particularly, we want to create a good football summary that tells the story of the game, not just its highlights. In considering how to best tell the story of a game, we take into account the research of Chu and Wu [Chu, W.-T. et al. 2008], who define the difference between a summary and a highlight reel. A summary takes into account all that happens in a video and includes the scenes that are of highest interest to the user. It establishes context for future clips, and all the selected clips are related. Transitions are often included to make sure that each clip is tied to the next. Together, the selected clips tell the story of what happened in the video. A highlight reel, on the other hand, is only concerned with highlighting individual clips of interest. Little thought is given to the context of the game as a whole, as each clip is considered independently. A good summary, then, must do more than simply select individual clips. It must consider the whole story, figure out which parts are most interesting to the viewer, and decide how to best tie them together and present them.

We have based our evaluation of a good summary on the four "C's" defined by He, Sanocki, et al. [He, L. et al. 1999]: *Coverage*, *Conciseness*, *Context*, and *Coherence*. A summary with good *coverage* includes all the key points of the original video. Having a *concise* summary indicates that only key points are included in the summary. *Conciseness* and *coverage* have a natural trade-off which must be balanced correctly based on the length of the summary. A good summary also includes *context* for each of its shots (making sure each clip is understood correctly), and *coherence*, meaning all the clips tie together with logical, natural flow. Since our focus is on telling the story of a football game, we believe that by satisfying all four of these properties we have avoided creating a simple highlight reel, but instead have effectively captured the story of the game.

In the world of ITV, it is important that the summary be dynamic. The viewer should be able to select how long he or she wants the summary to last, as well as adjust the summary to fit his or her needs. This provides the flexibility that is an inherent part of ITV. As we are no longer limited by broadcast schedules, the viewer should thus be able to match the summary to his or her interest in that particular game. This provides the rich, individualized content delivery that is expected from an ITV environment.

The rest of this project is summarized as follows. First, we discuss other relevant attempts at video and sports summarization. We then introduce the Time Warp Sports system, upon which our summarization service is built. A brief overview of our solution follows, after which we then discuss the three main challenges to solving the problem of creating an effective football summarization service. The first challenge addresses how to identify which clips are of highest interest to the viewer. We discuss two sources of information that allow us to identify these clips accurately, and we explain how we were able to verify this against user ratings. Second, we discuss how to tie the clips together and present the summary in a simple and intuitive manner. Our approach is verified through a user study. Third, we discuss our addition of interactive controls to create a more dynamic summary that fits the needs of the viewer, also verified through a user study. Finally, we conclude that our summarization service is effective at creating a quality summary of football games, and discuss future possibilities for the technology.

Chapter 2 – Related Work

A major concern in summarizing any sport is deciding what to include and what not to include. Many sports naturally lend themselves to summarization techniques because of the amount of "dead time" between plays. "Dead time" is the period between the end of one play and the beginning of the next—time used for setting up the next play, time-outs, official reviews, and other unexpected interruptions. Football has a lot of dead time. Logan, Durgin, et al. [Logan, J. D. et al. 2005] assert that a football game could be reduced from 3-4 hours to just twenty minutes by removing the dead time between plays. Unfortunately, a summary produced with this technique is fixed-length, and often longer than the viewer is interested in watching. We need a better approach if we want to provide a shorter summary.

The first challenge with creating a shorter summary, then, is deciding which plays to omit and which to include. This is critical to including proper coverage while maintaining conciseness. In order to decide, we must detect which events are happening in each play (such as touchdowns, interceptions, sacks, etc.), how interested viewers are in said events, and how to use that information to compose a summary. The second challenge is putting the selected clips together in a way that both preserves context and creates a feeling of cohesion. Moreover, the entire solution must be dynamic and focus on telling the story of the game. Many attempts have been made in these areas with varying success. In the following subsections, we discuss approaches to video segmentation, event ranking, user interest measures, and summary composition, and show how they are in fact inadequate at solving our particular problem.

Video Segmentation

One of the primary techniques used for detecting the events of each play is video segmentation. It decides not only how to divide the video into coherent clips, but also how to identify what happens in each one [Lienhart, R. et al. 1997]. Tong, Liu, et al. [Tong, X. et al. 2005] use specific clues to identify key events in soccer, such as the length of a replay scene, the duration of goalmouth views, the number of audience views after a replay scene, the amount of goal-net within replay, and the inclusion of scoreboard superimpositions on long field views. Each of these can be used to infer which events happened in each play. Tjondronegoro, Chen, et al. [Tjondronegoro, D. et al. 2003] first use clues in the audio to detect exciting events. They detect whistle blows, which often mark the beginning and end of a play, then look for excited speech or crowd cheers to identify exciting plays. Unfortunately, all of these rely on assumptions that may not hold in other situations. Often specific camera shots are not available for every play; whistle noises might not be audible, etc. This means that certain key plays could be missed, while many uninteresting events might be falsely identified as key plays. These techniques fail to provide summaries which are concise and yet have good coverage.

In order to improve accuracy, some approaches combine the extracted information from different streams available in the original broadcast, such as video, audio, temporal, and closed caption data. Zheng, Zhu, et al. [Zheng, Y. et al. 2007], Ma, Lu, et al. [Ma, Y.-F. et al. 2002], and Babaguchi, Kawai, et al. [Babaguchi, N. et al. 2002] demonstrate these mixed stream approaches. Each uses both audio and video features along with additional data to improve the accuracy of their detection algorithm. Zheng, Zhu, et al. use temporal data, such as duration of audio features, Ma, Lu, et al. use video motion, speech and music features, and Babaguchi, Kawai, et al. use text provided by a closed caption stream. By combining these features, they improve the accuracy of machine learning in order to extract information about each play. While this technique is fairly good at capturing key plays, it tends to include many false positives. This fails our requirements for conciseness. It also fails to actually identify what happens during the play. This makes the summary static, as it cannot adjust to fit time requirements.

Because event detection through audio, video, and related features is still unreliable, there have also been several attempts to integrate external information from other sources. Babaguchi, Kawai, et al. [Babaguchi, N. et al. 2004] present another approach using video analysis to extract the game clock from the video. Using the current game clock for each play, they can line up the video with external game statistics obtained through the Internet. This allows them to identify significant events, replays, pre-event shots, and post-event shots. The summary is then composed of those clips. Once again, this approach is limited. Game statistics obtained over the Internet are not comprehensive, and non-scoring events which might interest the viewer might not be included. Coverage is automatically limited to the events listed in the game statistics.

Video segmentation is still not reliable enough for event detection. Every approach is prone to error, and each uses techniques which either rely on a specific method of broadcast or require some sort of manual control. The right information for a quality summary is not available by simply analyzing the video. We must rely on external data to fulfill our requirements of conciseness and coverage.

Play Ranking

In order to select the most important events in sports, we need to develop a way of ranking each play in a game. The most common approach is to rank plays based on scoring events. Both Babaguchi, Kawai, et al. [Babaguchi, N. et al. 2004] and Takahashi, Nitta, et al. [Takahashi, Y. et al. 2005] use this approach on a similar scale, putting state change events (those that change which team has the lead) at the top, followed by other scoring events, and then events related to scoring. All other events fall into a fourth, lower category. The problem with this approach is that it does not take advantage of the different types of scoring and non-scoring events. Different scoring events may be more or less important to the game, depending on how many points were scored and how difficult it was to make them, etc. Also, while the score is what determines the outcome of a game, there are several non-scoring plays which are extremely interesting to the viewer and which have a part in determining what scoring events are possible. Leaving out these plays can affect the viewer's perception of the story of the game.

Other approaches use audio and video cues to help them rank each play or event. Tong, Liu, et al. [Tong, X. et al. 2005], Babaguchi, Kawai, et al. and Takahashi, Nitta, et al. use the temporal order of the plays, explaining that those toward the end of a game are more important than those toward the beginning, since they are closer to determining the outcome. Tong, Liu, et al. also use features such as event duration and confidence level to further differentiate between events, and Takahashi, Nitta, et al. include the number of replays of each shot as a feature as well. These are important to consider, but by themselves they do not provide enough distinction between events. They are also inaccurate to some degree, and may rank plays incorrectly.

Event ranking is critical to our goal of achieving the right balance between conciseness and coverage. Because our summary is dynamic, we need to get the correct ranking for each play so that we can best capture the story of the game, regardless of the length of the summary. For this kind of ranking, we must have more than just score and audio/video features for each play, and we need a good way to rank each play based on that information.

User Interest Levels

As an alternative to discovering and ranking game events, there have been many attempts to elicit interest from user interaction or user preference. Agnihotri, Kender, et al. [Agnihotri, L. et al. 2005] describe an approach where the user enters personality information, such as gender, Myers-Briggs Type Indicator, and other similar statistics. Specific elements such as face shots, words, anchor shots, dance shots, and more are extracted from the video and classified to match specific personality features. Based on the viewer's profile, the system can then customize a summary of the video to match the interests of the user. This approach demonstrates the worth of including user information, but is not well-tailored to a sports summarization solution. We could not customize our summary to the interests of the user without compromising the integrity of the story of the game.

We want to extract information from users that will help us to identify plays that are generally interesting. Nair [Nair, R. 2004] demonstrates a Level of Interest function that allows the system to determine which parts of a presentation are the most intriguing. Listeners were given a remote they could click whenever they heard something they wanted to "bookmark." When several people found the same part of the demonstration interesting, there was a spike in bookmarks. This spike indicates a key part of the talk. By selecting these key points, the system can compose a summary of the talk. Unfortunately, since the system does not possess any other information about the talk given, there is no accurate way to determine what it is that users find interesting and where that particular section starts or ends. Also, there is little incentive for the user to make these "bookmarks." Nevertheless, it does demonstrate an effective way of extracting user interest from user interaction.

This technique of determining user interest does a good job of showing what is interesting to the user, but it does not fully satisfy coverage and conciseness. With no sound knowledge of the actual events in the video, there is no guarantee that all key points will be identified. Also, it is unwise to rely solely on the user's input, since there is the potential for false positives as well. This technique would be better used to enhance the effectiveness of event detection. Zheng, Zhu, et al. use a type of user input to train their learning algorithm, but fail to leverage the full potential of this technique.

Summary Composition

While most of the previously mentioned approaches present some method for constructing a summary, most focus on selecting what clips to include and exclude, and do not mention anything about putting the clips together. This is a problem, however, since two of the four "C's" are focused in this area. Some context is inferred from the order of the clips, but more is needed for a good understanding of each clip. Coherence is entirely focused on tying the clips together. Yet many of these solutions simply splice the plays together in a temporal order and leave it at that. Only a few address this issue. Babaguchi, Kawai, et al. [Babaguchi, N. et al. 2004] overlay text with information about the clip at the beginning of each clip. Logan, Durgin, et al. [Logan, J. D. et al. 2005] mention the use of animations or similar transitions, but they provide no information about what to do or how to do it. In order for our summary to exhibit good context and coherence, we need some sort of transition to tie each play to the next.

Chapter 3 – Time Warp Sports

In order to provide a football fan with a dynamic summary, we need a solution that will allow him to select the game, specify how long he wants the summary to last, and then interact with the summary in some useful fashion. Rather than building a new system ourselves, we have elected to build our summarization service within the Time Warp Sports (TWS) system. TWS is an interactive viewing environment for watching broadcast sports that both enhances and expands the viewer's experience. It also provides us with a variety of useful information that will help us to compose our summaries. Our solution makes use of this information as well as the framework TWS provides in order to create a novel summarization service. In order to understand our solution, then, it is important to understand how TWS works.

When a viewer watches sports using TWS, he or she is provided with a feature-rich player that allows him or her to control how he or she watches the game. Without interaction, TWS behaves much like regular TV—it simply plays the game back to you as presented by the broadcasting company. With the use of a remote control, however, the viewer can manipulate the game with controls that allow skipping to the next play, replaying the current play, skipping back to the previous play, and/or switching the camera angle. If the viewer is not watching the game live, he or she can skip from play to play, avoiding the dead time between them. He or she can also review penalty calls from different camera angles at his or her leisure. Additionally, the viewer can pull up current statistics at any time. He or she does not have to wait for the broadcasting company to include them in the broadcast; they are always available. In this way the TWS system provides the viewer with a more enjoyable, customized viewing experience.

TWS consists of two parts that are important to our service. The first is the annotation file, which contains information and cues about every play of the football game. The second is the player, which loads in the user's browser and provides interactive enhancements alongside the game. Because both parts are essential to our summarization service, we will describe the functionality of each as it relates to our solution in the following sections.

Annotation File

Central to TWS is an annotation file that contains detailed information about the game. It is created manually by two people while watching the game live. With a small delay, this annotation is provided live, and can be used by the system immediately. One person marks the beginning and end of each play and series, after which the other inputs statistical information about each play (See Figure 4.2). Also included is information about each camera angle, each team, and the sport itself. All annotations are hosted on a central annotation server. When the viewer loads a specific game, the matching annotation file is downloaded by the player. The player then makes use of the information in the file, using the offsets to allow the viewer to skip from play to play, and using the camera information to provide different angles.

Statistics		
Field Goal Attempts	**Pass Attempts**	**Rushing Yards**
Field Goal Percentage	**Pass Completions**	**Score**
Field Goals Made	**Pass Completions/Attempts**	**Total Yards**
First Down Marker	**Passing Yards**	**Yard Marker**
Fumbles	**Penalties**	**Yards / Penalty**
Interceptions	**Punts**	**Yards Per Carry**
Kickoffs	**Rush Attempts**	**Yards To Go**

Figure 3.1: Game Statistics from the TWS Annotation File

The annotation file also keeps track of running statistics throughout the game so that the TWS player can display them at any time. These statistics are stored as snapshots for each play, and they contain the start and end offsets of each play in the video, as well as identify what quarter and down, who has possession, the current score, and other running game statistics for each team. These attributes are summarized in Figure 3.1. Most of the statistics are running totals, but by computing the difference between the current and the previous play, we can determine exactly what happened from a statistical viewpoint. Three of the attributes, "Yards To Go," "First Down Marker," and "Yard Marker" allow us to determine exactly how the ball moved on the field. This information is extremely valuable when composing a summary, and can be an effective measure of how important or interesting a play is.

Unlike other automatic solutions, which try to extract information about games by analyzing the video stream, using an annotation provides us with all the information we need. We accurately know when a play starts and ends, we know when and in what order each play occurred, and we know information about what happened in each play. This is a much more effective and accurate solution to video segmentation. This also makes our job of composing a summary much easier, as this information is easy to access and process.

TWS Player

The TWS player is designed to be simple and easy to use. Its minimal interface is focused on maintaining the TV watching experience. All of its interactive elements stay out of the way when not in use, and are designed to be intuitive. Menus slide up onto the screen when users press the triggers on the remote, and they only remain visible as long as the trigger is held down. Also, each menu is designed to look like the remote, making it intuitive for the user to determine which button does what (See Figure 3.2). In addition, the main menu's button actions remain active even when the menu is not displayed. Once a user has become accustomed to the controls, he or she can avoid the overhead of summoning the menu while interacting with the game.

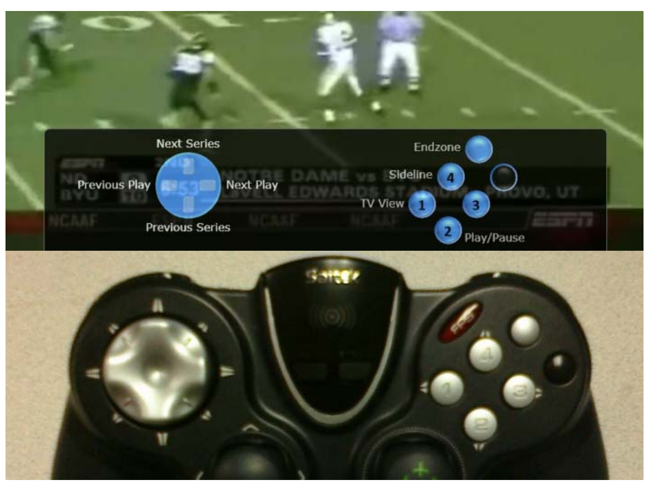


Figure 3.2: Simple menus in TWS are designed to look like the controller for intuitive use.

As previously mentioned, TWS allows the viewer to control how he or she watches the game by skipping plays, switching camera angles, and other useful actions. Each of these interactions is logged by the TWS player, along with information about when and how it was used. Currently, these actions are logged to the local machine, but the system is designed so that they may be uploaded to any server. These viewing statistics also provide useful information for building a summary. By analyzing these statistics, we can determine which plays were more interesting to viewers. This analysis will be discussed in more detail in Chapter 5.

Figure 3.3: Transitions help the viewer remain oriented while navigating through the game.

The TWS player's interface is designed as simple overlays that pop up over the video. As the viewer navigates the game, small transitional banners are used to make the viewer's actions clear to all who might be watching (See Figure 3.3). The banner lists the action performed and relevant information about the current play, such as possession and score. This provides a sense of coherency to the viewing experience and keeps the viewer oriented, even when skipping large portions of the game. We make use of this same overlay framework to create a more detailed transition that can summarize skipped plays. Using this framework helps us to fit our additions into the look and feel of TWS, making the experience more natural to the viewer. These transitions are discussed with more detail in Chapter 6.

Chapter 4 – Summarization Framework

We designed our system to work as follows: First, the viewer inputs which game he or she wants to watch and then the length of the summary desired into the TWS player. This information is sent to our summarization service on the annotation server. The service uses a Degree of Interest (DOI) function to order the plays of the game by interest, based on game statistics from the annotation and on viewing statistics from the TWS player. The service then uses the algorithm in Figure 4.1 to select which plays are included in the summary. This is a typical greedy approach, repeatedly picking the most interesting plays, adding them to the summary, and subtracting their length from the total time until the specified time is filled. Once the list of included plays is complete, the service modifies the annotation by appending the attribute "Summarize: Exclude" to all plays not included in the summary (See Figure 4.2). The modified annotation is then returned to the player.

```
Play[] sortedPlays = DOI(plays);
int totalTime = 0;
List summaryPlays = new List();

foreach (Play p in sortedPlays) {
   if (totalTime + cliplength(p) < givenTime) {
      summaryPlays.add(p);
      totalTime += cliplength(p);
   }
}
```

Figure 4.1: Algorithm for selecting plays after they have been sorted by a DOI function

```
Clip
{
  ClipID : 1024
  StartOffset : 2383.465
  EndOffset : 2389.077
  ClipType : "Play"

  Penalties_Notre_Dame : "0"
  Score_Notre_Dame : "3"
  Punts_Notre_Dame : "1"
  First_Down_Marker_BYU : "48"
  Interceptions_Notre_Dame : "0"
  Yards_vs_Penalty_Notre_Dame : "0 / 0"
  Yard_per_carry_Notre_Dame : "-6"
  Yard_Marker_Notre_Dame : "28"
  Yard_per_carry_BYU : "2.25"
  Pass_Completions_Notre_Dame : "2"
  Score_BYU : "13"
  FG_Attempts_BYU : "0"
  Yards_vs_Penalty_BYU : "5 / 1"
  Yards_to_Go_Notre_Dame : "2"
  Yard_Marker_BYU : "44"
  Transition : "BYU_Ball_-_Down:_2"
  Summarize : "Exclude"
  Interceptions_BYU : "0"
  Passing_Yards_Notre_Dame : "18"
          ....
```

Figure 4.2: Example of a clip's annotation that has been excluded from the summary.

Once the player receives the annotation, it computes a playlist of which clips are included from the original game, and then uses it to show that summary to the viewer. As each play is shown, the system determines whether there are any excluded plays between the current play and the next in the summary. If there are, the game is paused, and the viewer is shown a short transition which summarizes the excluded plays. As the transition ends, it is quickly hidden, and the summary continues. While watching the summary, the viewer can use any of the interactive methods provided by TWS, such as reviewing the play from a different camera angle or skipping back to a previous play. When he or she is done exploring the game, the viewer can resume the summary by pressing a button. At the end of the summary, the viewer is shown the final game statistics.

There are a few key challenges to making this process function. First, we need to create an effective DOI that rates plays according to how interesting they are to the viewer. Furthermore, we need to make sure our play selection considers the story of the game. It needs to find the correct balance between interesting and important plays. Second, we need to create an effective transition that will not distract or disrupt the summary viewing experience, but will rather provide the viewer with information that ties the summary together. This further helps to focus on the story of the game instead of mere highlights of key plays. Third, we want to improve the summary through interaction. We must provide viewers with means to control the summary and customize it to their preferences without changing the story of the game which our summary relates.

Chapter 5 – Creating a DOI from Viewing Trends and Game Statistics

First and most important to our solution is the creating of an effective DOI function. This function needs to return an ordered list of plays based not only on how important the play is to the game, but also on how interesting the play is to the viewer. With effective features that allow us to measure both of these qualities, we can compose a summary that tells us the story of the game instead of presenting a highlight reel. By building our system on top of TWS, we gain access to several sources of features that allow us to construct an effective DOI function. This chapter focuses on the process we used to construct our DOI function and how we validated our solution.

Figure 5.1: Screenshot of the play rater program. The user rates the play from one to ten with the stars at the bottom.

Before we can create our DOI function, we have to have a "gold standard" to which we can compare our results. Having a gold standard allows us to build our DOI function to match it as closely as possible. We created our gold standard by asking twelve volunteers to rate the plays of a game according to how interesting it was. We used the BYU vs. Notre Dame football game from 2004 as our test game. The program used to collect these ratings is shown in Figure 5.1. As each play was shown, the user was asked to rate that play from one to ten by clicking on the corresponding star. Each response was then transmitted and collected in a database. From this we computed the average rating and standard deviation of the rating for each play, shown in Figure 5.2. The average standard deviation was 1.597. Inspection of the data shows that on plays with a lower average rating, there was a more significant degree of variation. On higher rated plays, however, there was less variation. Users generally agreed on the interesting plays but were more varied on the less interesting ones. We believe this to be an adequate measure of the game since a summary naturally focuses on those interesting plays.

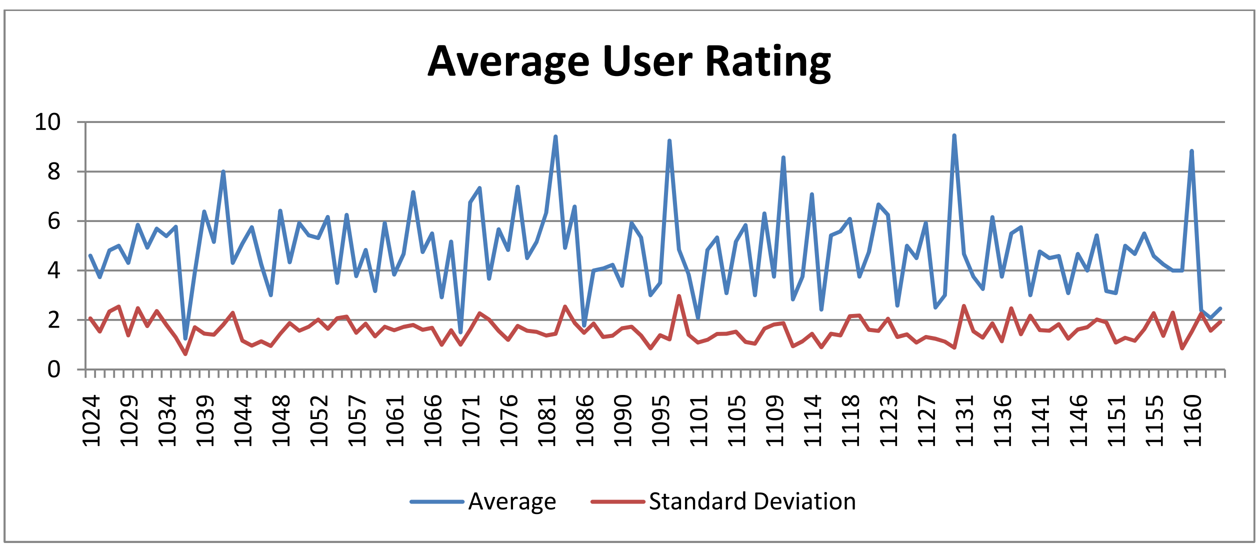


Figure 5.2: Average and standard deviation of the user ratings used for our gold standard.

Now that we have a gold standard, we must determine which features will allow us to create a matching DOI function. There are two sources for this information: the annotation file from TWS, and the log data of viewer input to the TWS player. The first of these—the annotation file—allows us to calculate game statistics for each play. These, in turn, allow us to determine how important a play is. From game statistics, we can discover which plays contain scoring and other events that are significant to the outcome of the game. Because the TWS annotation is rich with statistical information, we can consider many more aspects of a play, such as yards gained, which down it was, etc. This provides us with a tremendous advantage over just considering scoring events, as it gives us more detail with which to rank each play.

We used the following events from the game statistics to determine if a play was important to the story of the game:

- Touchdown (Offensive or Defensive)
- Field Goal
- Play after Touchdown (Point After Touchdown or Two Point Conversion)
- Penalty Yards
- Turnover (Fumble or Interception)
- Punt
- Kickoff
- First Down
- Third Down Conversion
- Fourth Down Attempt

- Yard Gain/Loss

Each game event was computed by taking the difference of the current play's statistics with the previous play's statistics. Yard gain and loss was computed with the "YardMarker" statistic. Touchdowns were determined by a score difference of six, and point-after-touchdown and two-point conversions were identified since they always follow a touchdown.

The first three features are the scoring events. These are usually crucial because they directly influence the outcome of the game. The rest are events that are statistically important to the effectiveness of a team. Turnovers and penalties are obviously bad, and events like 1st downs, 3rd down conversions, and 4th down attempts indicate the team's effectiveness at moving the ball. Also, since progress in football is primarily measured by moving the ball down the field, yard gains and losses are key to identifying the effectiveness of a team. To verify that these game events are good indicators of how important a play is to the game, we performed a linear least squares regression against the average user ratings to see how close we could match them. We used the following equation to compute the linear regression:

$$\hat{\beta} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\bar{y} = \left(\frac{1}{n}\sum x_i x_i'\right)^{-1}\left(\frac{1}{n}\sum x_i y_i\right)$$

X is a matrix composed of one row per play and one column per feature, and *y* is a vector composed of the average user ratings. We first computed the mean and standard deviation of each event, then converted each value into a z-score with the following formula:

$$z = \frac{v - \mu}{\sigma}$$

In this equation, *v* is the original value, μ is the average for that feature, and σ is the standard deviation for the feature. This normalizes all events so that they lie on the same scale. We then perform the regression against these values to determine the best coefficients for a linear equation of the features. We show the results of our regression in Figure 5.3. In all of our comparisons, we compute the cosine distance between our DOI function and the user ratings in order to measure correlation. The cosine distance is computed as follows:

$$cosine\ distance = \frac{A \cdot B}{|A| * |B|} = \frac{\sum_i x_i * y_i}{\sqrt{\sum_j x_j^2} * \sqrt{\sum_j y_j^2}}$$

A and B are vectors composed of the rating for each play from either the user ratings or the DOI function; x and y are the values from each vector. With just the game statistics alone, we achieved a 0.979 correlation between our DOI function and the user ratings with a mean squared error (MSE) of 1.075. Because our MSE is smaller than the average standard deviation of the user ratings, and because our correlation is very high, we consider this DOI function to be fairly accurate.

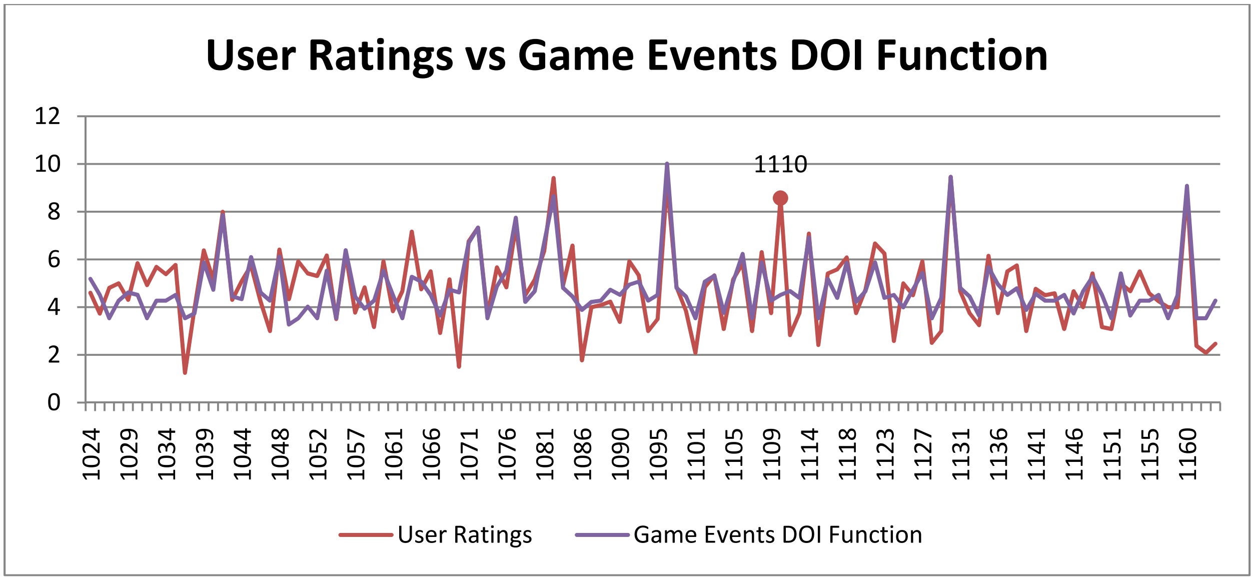


Figure 5.3: Comparison between the user ratings and our DOI function based only on game events.

Unfortunately, there are some plays that are ranked highly by users that the DOI function completely misses. For example, play 1110 (labeled in Figure 5.3) is considered very important by users, but the game events DOI function ranks it very low. When we look at the game statistics, all we see is a punt for this play. If we actually watch it, we see that the punter drops the snap, picks up the ball, runs to the right and forward for ten yards, and then punts to the opponent's two yard line. This is an exceptional and exciting play, and should be included in our summary. While just using game statistics yields a fairly accurate DOI function, we need more data that allows us to capture what actually interests the viewer.

In order to capture viewer interest, we turn to our second source of information: the TWS player's logs of viewer input for each game. From these logs, we can compute viewing statistics that tell us the average number of times each interactive command was used during each play. We have found that these viewing statistics are good indicators of how interesting a play is. For example, when a viewer sees a play he or she is interested in, he or she wants to watch the play again. In the TWS system, when a viewer changes cameras, the system automatically backs up to the beginning of the current play. This allows the viewer to quickly review those plays from different camera angles. Figure 5.4 shows a graph of camera view changes per play. It is clear that high numbers of camera view changes correlate with key plays in the game. Notably, the punt missed by our game events DOI is clearly identified. By using these viewing statistics, we can more reliably determine which plays interest viewers, even if the game statistics don't identify them as doing such.

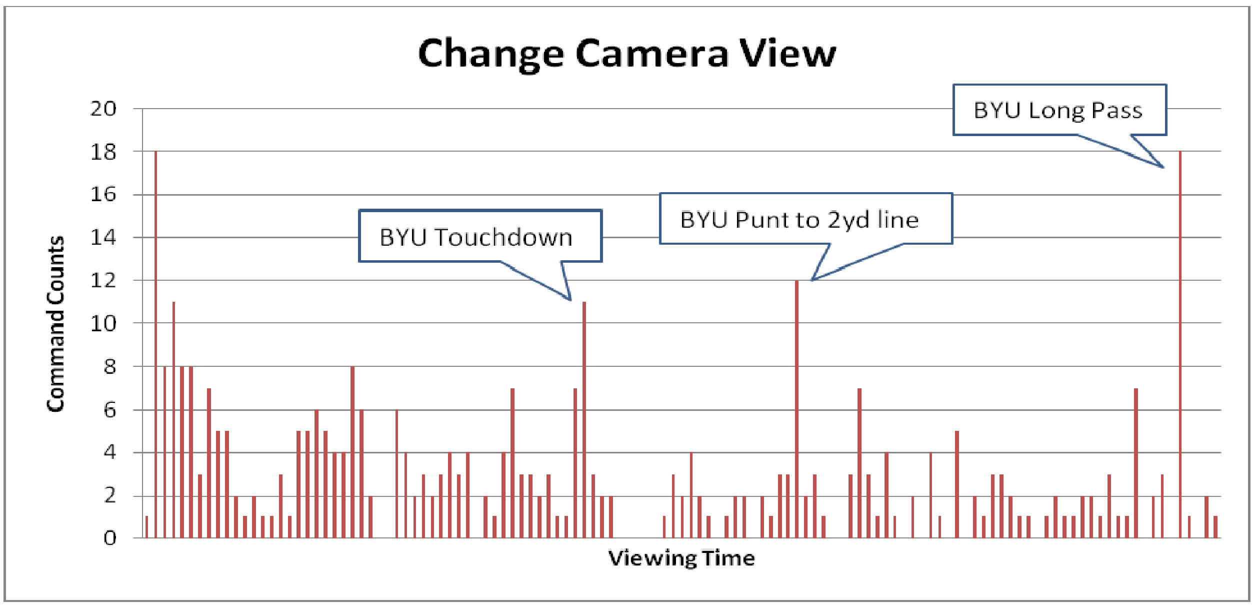


Figure 5.4: Graph of camera view switches during a user test. Example plays are indicated with a callout, showing how a high number of camera view changes relate to interesting plays in the game.

For our experiment, we used the log files from eleven separate TWS evaluations. For each evaluation, a small group of users (two to eight people) were shown the BYU vs. Notre Dame football game after being introduced to the TWS system. From the logs, we computed the average number of uses for each of the following commands:

- Next Play
- Previous Play
- Replay Down
- Change View

We did not include other commands such as “Previous Series,” “Next Series” or any of the menu commands, as these do not relate well to specific plays.

Similar to our analysis of the game events, we performed a least-squares regression against the user ratings on the viewing data and produced a similar DOI function (See Figure 5.5). Our correlation was 0.960 with a MSE of 1.9984. Although this function did not match our user ratings as well as the game events DOI function, there is still a high correlation between the two. This confirms our hypothesis that viewing statistics can tell us which plays interest viewers. We note that in this function, many of the key plays are still given high values, including play 1110 (the punt mentioned previously).

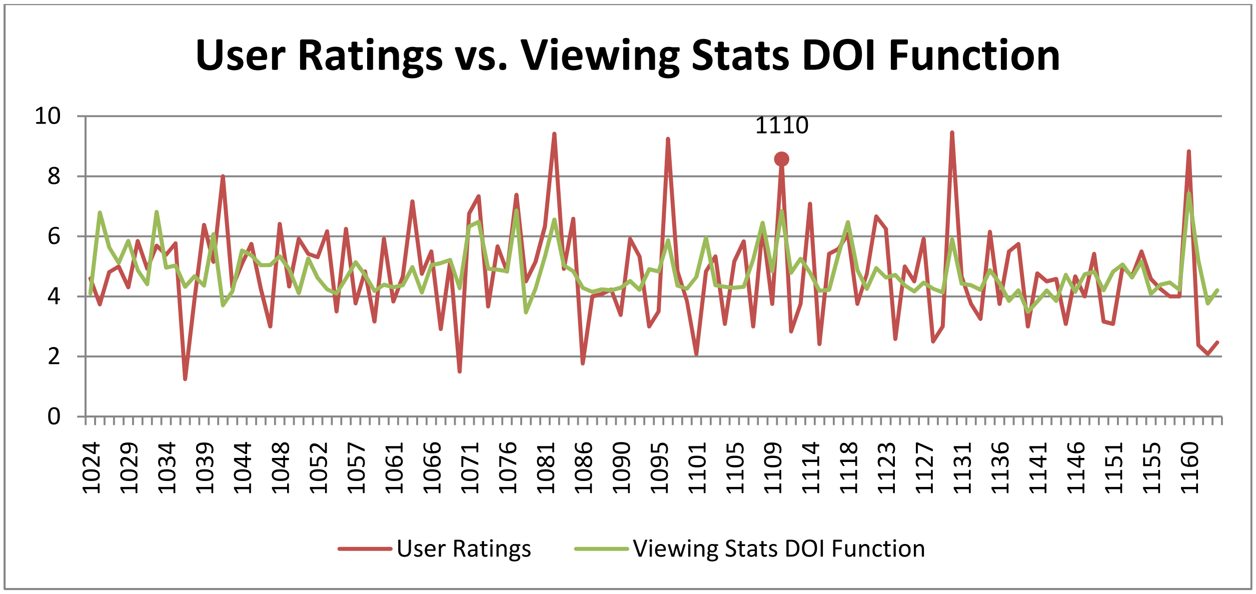


Figure 5.5: Graph of our viewing stats DOI function vs. the average user ratings

Because we expected a closer match from the viewing statistics, we investigated to determine what might have caused the mismatch. We found that the viewing statistics were slightly skewed because of the learning curve for the TWS system. During the first section of the game, most users were experimenting with the TWS player's controls, and thus were not focused on the game. Therefore, the viewing statistics for the beginning of our game don't correlate very well with the viewer's actual interest in those plays. We believe that this also accounts for some of the discrepancies in our viewing statistics DOI function. Nevertheless, it is clear from our results that there is a strong correlation between viewing statistics and viewer interest.

We also considered the possibility of viewer bias, since the majority of the participants were BYU fans. To test for this, we found the average rating from both our user ratings and our viewing statistics DOI function for each team and event. By calculating the difference between scores for each team and event, we were able to determine the relative bias. The results are shown in Figure 5.6. In spite of using mostly BYU fans in our user studies, we could find no clear bias between the two teams. Most of the events were rated about the same for each team, and were higher for Notre Dame about as often as they were for BYU. With no significant bias in either source of data, we assume that our gold standard is accurate for both BYU and Notre Dame fans, and that viewing statistics are a good source data for determining unbiased user interest.

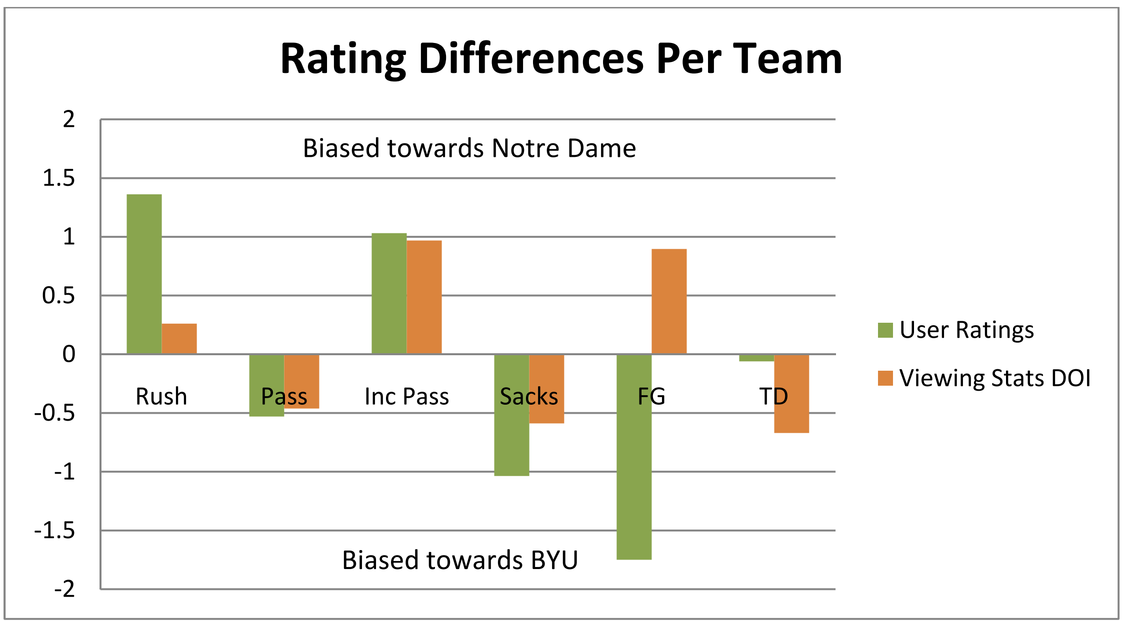


Figure 5.6: Shows the difference in average rating for each major event type for both the User Ratings and the Viewing Statistics DOI.

By themselves, both our game events DOI function and our viewing statistics DOI function would yield fairly effective summaries. By combining the two, however, we can achieve an even closer correlation with our user ratings. Figure 5.7 shows the results of a DOI function created by combining information from both game events and viewing statistics. Similar to the other DOI functions, we created this combined one by performing a least-squares regression against both the game events and the viewing statistics. With this DOI function, we computed a 0.981 correlation with a MSE of 0.971. This is a closer match than either of our previous functions. Each source of data compliments the other, allowing us to obtain a higher correlation with less error.

We believe this DOI function to be adequate at achieving our goal of finding both interesting and important plays. This in turn helps us to tell the story of the game. We use this combined DOI function to compute a summary that is dynamic and provides good coverage of the game. Because the plays are ordered by their importance to the story of the game, we will always be as concise as possible within the time allotted by the viewer. We also believe that this DOI function would perform similarly across all games, although its performance could be further optimized with more data from multiple games.

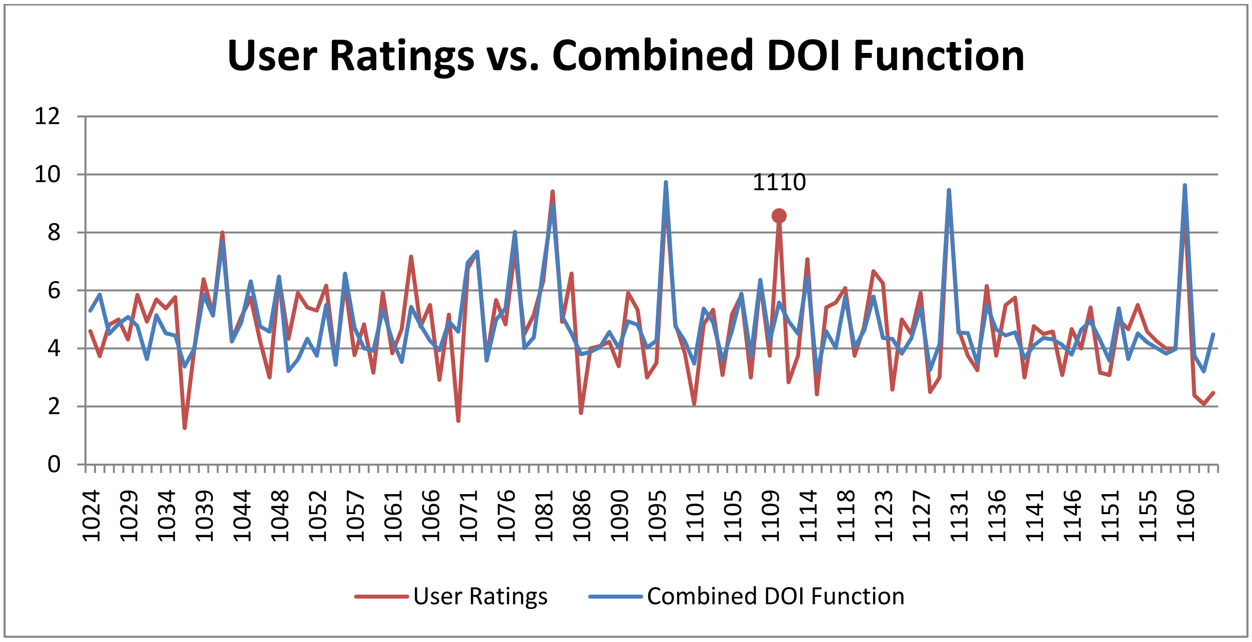


Figure 5.7: Graph of our combined DOI function vs. the average user ratings. This DOI combines both viewing statistics and game events to get a closer match.

Validation

Much of our validation has already been discussed in our description of how we constructed our DOI function. However, a summary is built by picking the most important plays and leaving out the rest, as mentioned before. Only a fairly small subset of the plays is important, being those that are ranked highest. While our DOI function may closely match the user ratings, we want to make sure that all of the highest ranked plays from the game are included. To do so, we compare the top plays from our DOI function with the top plays from the user ratings.

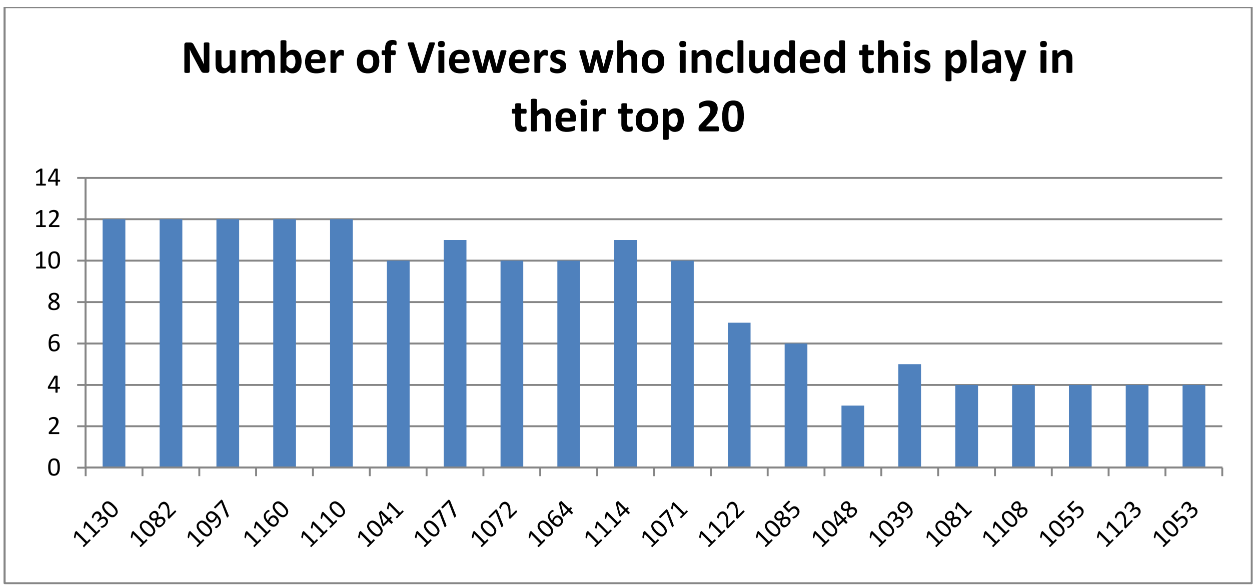


Figure 5.8: The top twenty plays from the average user ratings (ordered by average user rating) and how many users included each play in their own personal top twenty.

For our analysis, we focus on the top twenty plays from the average user ratings. To gain more insight into how important these plays are, we also compute the top twenty plays from each individual rater. For each of the top twenty average plays, we count how many of the raters included that play in their top twenty. Results are shown in Figure 5.8. From this figure, we can see that there are eleven plays which are included by the two-thirds majority of raters in their top twenty. These eleven plays, then, are key to understanding the game. For the remaining nine, we see that the raters begin to disagree. For six of the nine, either one third or fewer of the raters included that play. While we want the top twenty plays from our combined DOI function to match closely with the top twenty of the average user ratings, our focus centers on those eleven plays. Figure 5.9 contains information about each.

Play	Game Events
1130	BYU Interception Notre Dame Touchdown
1082	BYU Touchdown 42 Yard Pass
1097	Notre Dame Touchdown 54 Yard Rush
1160	BYU Pass 37 Yards
1110	BYU Punt
1041	BYU Field Goal
1077	Notre Dame Sacked 4^{th} Down Attempt
1072	BYU Fumble
1064	Notre Dame Sacked
1114	Notre Dame Pass 18 Yards
1071	BYU Pass 16 Yards

Figure 5.9: Top 11 Plays with related game events

Interestingly enough, one of the primary scoring events—Notre Dame's field goal towards the beginning of the game (play 1055)—is left out of these top eleven plays, despite being included in the top twenty. This demonstrates that not all scoring events are exciting to users. There are several other passes and sacks that rank higher. A summarization algorithm that only considers scoring plays would leave out these interesting, essential plays from its summary.

Fortunately, our DOI function does a better job at matching viewer interest through its use of viewing statistics. We computed the top twenty plays from our DOI function and compared them with the top twenty from the user ratings. There were sixteen plays in common between the two. Out of the eleven key plays from the user ratings, our DOI function included ten. Plays 1064, 1123, 1085, and 1053 were the ones left out by our DOI function from the user-rated top twenty. Figure 5.10 shows how our DOI function ranked the user-rated top twenty plays. Our DOI function ranks most of these plays with about the same score as the users. This shows that our DOI function is adept at picking out non-scoring plays which are interesting to the viewer.

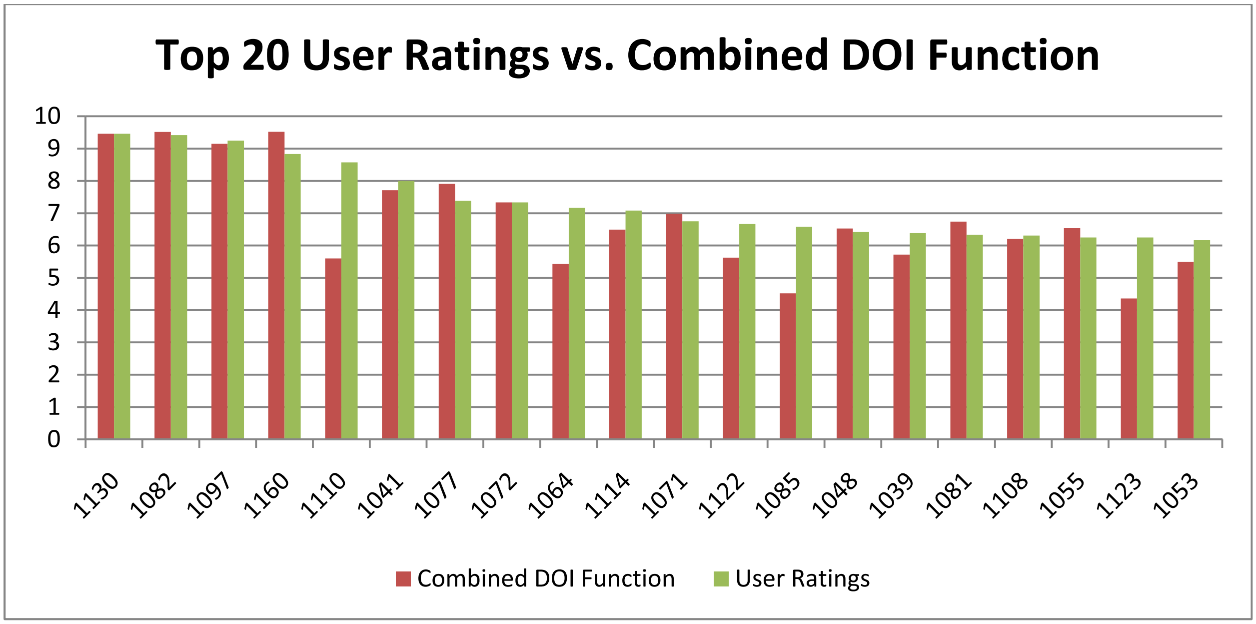


Figure 5.10: The top twenty plays from the average user ratings with their rankings, as well as the equivalent combined DOI function rating.

We would like to take a closer look at the discrepancies between our DOI function and the user ratings. We first consider play 1064, which was the only one of the eleven key plays not included by our DOI function. When we look at the game, we see that Notre Dame was sacked for five yards by BYU on a 2nd down. When we compare this with other similar plays, we see that users tended to rank sacks very high. With our set of game events, however, we are only able to track yard loss, which does not directly correlate with sacks. This is a possible cause for the mismatch.

The other plays left out by our DOI function could be the result of similar deficiencies in our set of game events. One of the three includes a kickoff where Notre Dame returns the ball for a significant gain. Our events, however, do not include how many yards the ball was returned; they only indicate where the ball stopped. If we were to modify TWS to include punt and kickoff return yardage, our system would likely catch these plays. Other possible reasons for the mismatch could be that the viewing statistics and the user ratings for these plays do not concur. All of them had a standard deviation above the average on the user ratings, indicating possible disagreement among viewers.

While there are some differences between our DOI function and the user ratings, we have shown that our DOI function is effective at ranking plays for creating a summary. We achieve our goal of creating a dynamic summary that is both concise and that provides good coverage. Because we consider viewing statistics as well as game events, we are able to discover non-scoring plays that are important to telling the story of the game. This DOI function lies at the heart of our summarization service, and allows us to create a summary that shows the viewer what he or she wishes to see. This provides a good foundation for a complete summarization service.

Chapter 6 – Providing Understanding through Transitions

Selecting the right plays is an important part of putting together a good summary. In every case, however, information is going to be left out. Anytime something is skipped over, the context is lost, and the viewer is left somewhat confused. Quick jumps in video can disorient, even distract, the viewer from the actual summary—especially in sports. Furthermore, without a sense of time the viewer may miss the significance of a play. Therefore, there needs to be something that ties each highlighted play together and provides a sense of cohesion for the entire summary. This transition between plays must provide adequate context as well as tie into the next play so that the viewer can watch without feeling lost. We accomplish this by inserting dynamic transitions when any play is skipped over.

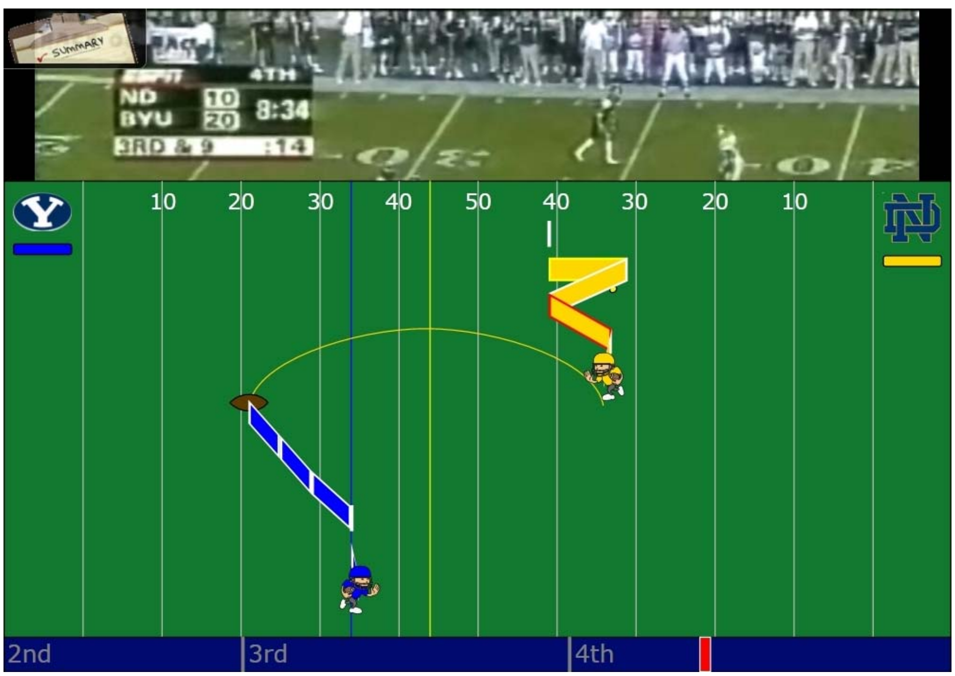
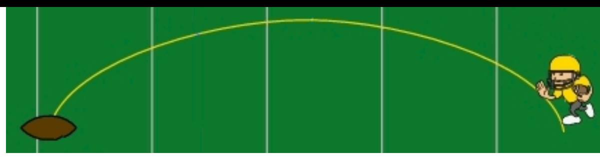
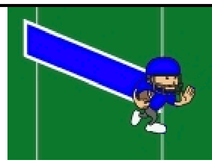
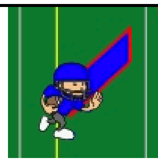
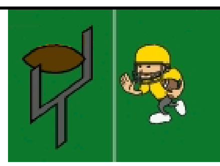
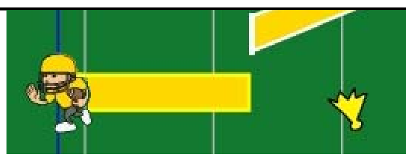
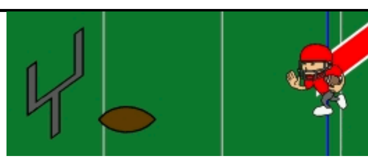


Figure 6.1: An example of a transition summary right before it disappears. A penalty, backward motion, a punt, and down and yard lines are visible. The time-bar at the bottom also shows the progress of the summary.

There are many plays in football that can be skipped over without detracting from the story of the game. An example of a common set of plays is the typical “three and out,” meaning that the offensive team fails to achieve a first down in three tries, and uses their fourth down to punt the ball. It is not uncommon for this to occur back and forth several times. While these plays may be uneventful, it is often useful to know that this occurred between two especially interesting plays. Additionally, viewers can draw conclusions about how each team did based on how the ball moved. For example, if a team shows a lot of backward motion, they can conclude that the opposing team's defense is doing well. Also, odd plays that may not have been included by the summarization service often stand out. The viewer might be interested in this, and may then use the interactive controls to review the play.

Because of the detail included in the TWS annotations, we know how the ball moved between each play. We compute the game events for each play in the same way we compute the DOI game events. Thus, we can represent every play with a small bar indicating ball movement (See Figure 6.1). A small cartoon football player follows along with the movement of the ball, helping the viewer to understand who has possession of the ball and which way they are going. Red borders are used to indicate backward movement, and yellow borders are used to indicate the result of penalties. We also included short animations to demonstrate a successful or a missed field goal (FG) or point after touchdown (PAT), punts and kickoffs, turnovers (interceptions or fumbles), and penalty calls. Figure 6.2 shows how each event looks in the transition. Team direction is switched every quarter to ensure that it matches with the camera feed.

Animation	Description
	Represents any kickoff or punt.
	Represents a normal rushing or passing play.
	Represents yard loss from being pushed back.
	Represents a successful field goal or PAT.
	Represents a penalty.
	Represents a missed field goal or PAT. The ball bounces off the post and falls down.
	Represents a turnover.

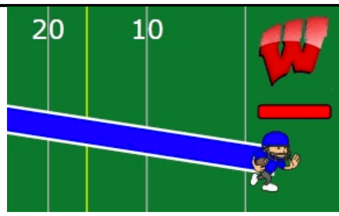
	Represents a touchdown or a two point conversion.

Figure 6.2: Transition Summary Animations

Our transitions are designed to be quick and easy to understand. We do not want them to detract from the summary, but we do want them to help provide more context, and to help the viewer understand what actually happened in the game. Between plays, the game is paused while the transition pops up quickly, displays its animation, and then hides again. The transition stays on the screen only long enough for the viewer to quickly grasp what it is showing. If only one or two plays are skipped over, they stay on screen for a shorter amount of time than they might for ten to fifteen plays skipped. The animations are fairly accelerated and generally take less than a few seconds to display. By tying the plays together with this quick but comprehensive transition, we expect to maintain a sense of cohesion for the whole summary.

Not only do we want to make sure that the viewer understands the surrounding context of a play, but we also want them to understand when it occurs within the chronology of the game. To do this, we include a persistent time bar at the bottom of the screen that shows the progress of the summary (Figure 6.3). Each quarter is marked along the time bar for reference, and a small red indicator shows where the current play is. As the summary progresses, the red indicator moves along the time bar. When a transition is visible, the indicator moves at an accelerated rate, then slows down when an actual play is being shown. With a quick glance, the viewer can determine about how far into each quarter a play occurred. This, along with the transitions, allows them to understand the context of each play, ensuring that our summary is truly just that—a summary, and not just a highlight reel.

Figure 6.3: Example of the time bar used to show the progress of the summary.

Validation

We chose to use a user study to determine the effectiveness of the transitions. Twenty users were asked to view two football summaries: one with transitions, one without. To remove any possible bias, the summary with transitions was referred to as transition style A, and the summary without transitions was referred to as transition style B. Furthermore, the twenty users were split into two groups. One group watched the summary *with* transitions before the summary *without* transitions, and the other group watched them in reverse order. Both summaries came from the same game, and were identical except for the inclusion or exclusion of the transitions. To help us understand how well the users understood and made use of the transitions, we asked the viewers to fill out a survey after each summary, as well as a final survey at the end, making a total of three surveys. For each statement, the user was asked to show how much they agreed with the statement on a Likert scale. The statements presented after each summary were:

1. I understood the main story of the game from watching the summary
2. I felt there was a lot missing from the summary

3. I was able to follow the summary easily
4. I thought the summary was boring
5. I would find this type of summary useful

For each statement, we used the following Likert scale: “strongly disagree,” “disagree,” “undecided,” “agree,” and “strongly agree.” The users circled one of these answers for each statement. It is important to note that the participants received no instruction on what the transitions would look like or how they should interpret them. We purposefully excluded this information as a test to see if the viewers would be able to figure out the meaning of each symbol on their own.

In addition to the surveys, we asked all of the participants, in groups including about four people, to answer several open ended questions about their experience. These were recorded on video tape. The following questions were asked:

- What did you like about the transition summary?
- What didn’t you like about it?
- Was there any part of the transition summary you didn’t understand?
- Did anything distract you? What?
- What could be done differently?
- Would you find this kind of summary useful? Why or why not?

Additional questions about the interactive controls were also asked, but we leave the details of those questions for the next chapter. We use the responses to the open-ended questions when considering the survey questions in order to provide some explanation as to why users answered the way they did. We also use these responses to support our conclusions.

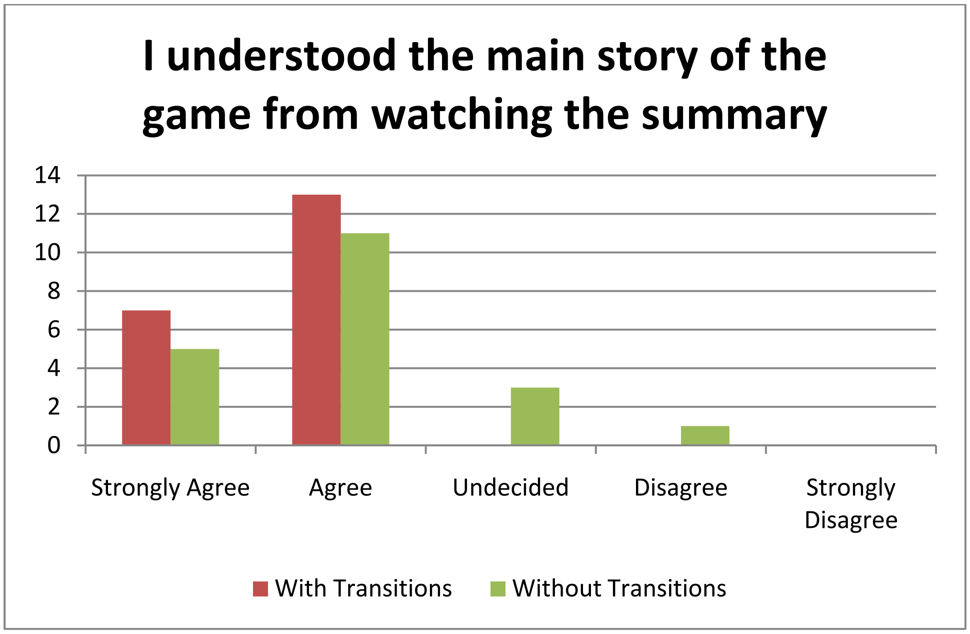


Figure 6.4: This graph shows that viewers were able to understand the main story of the game better with transitions.

Figure 6.4 shows the results for the statement "I understood the main story of the game from watching the summary," from the two surveys administered after each summary. With transitions, all users agreed that they understood the main story of the game from the summary. This is significant, since every one of our participants agreed with this statement. Without transitions, there were fewer who agreed, and a few who were undecided or disagreed. Moreover, during the open-ended questions, fifteen of the twenty participants replied that the transitions helped them understand the game better than a summary with no transitions. Five of them stated that they found the summary confusing without them. We consider this to be strong evidence that our transitions were effective in telling the story of the game instead of simply presenting the viewer with a highlight reel.

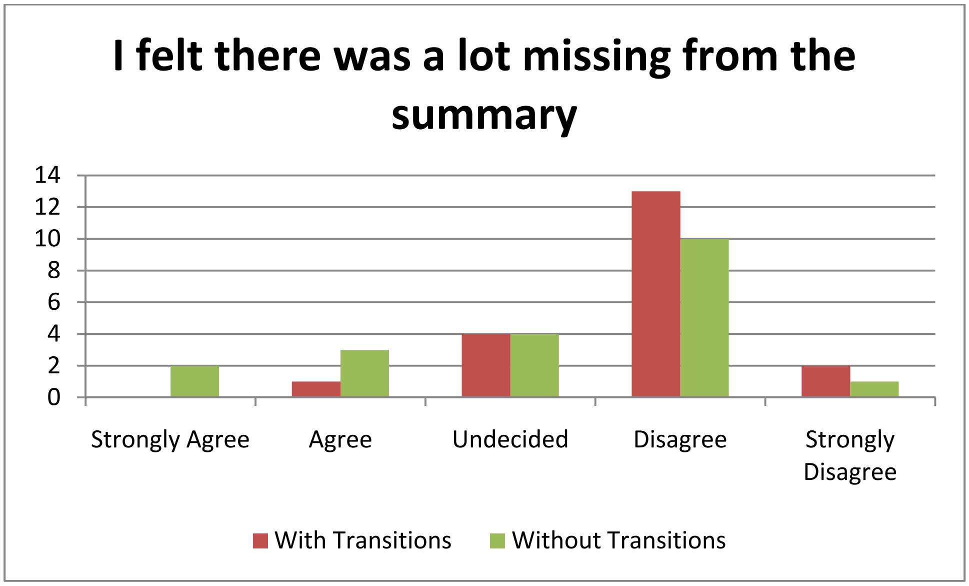


Figure 6.5: This graph shows that viewers felt they saw more of the game with the transitions.

We show the results of the next statement, "I felt there was a lot missing from the summary," in Figure 6.5. Our transitions provide insight into what plays were skipped over as well as the context for each play shown in the summary, so we expect that users will not feel like they have missed anything. From this graph, our hypothesis is confirmed. More users disagreed with this statement when watching the summary with transitions, indicating that they felt they saw a lot more of the game with transitions rather than without. This strongly supports our claim that the summary helps users to understand the story of the game.

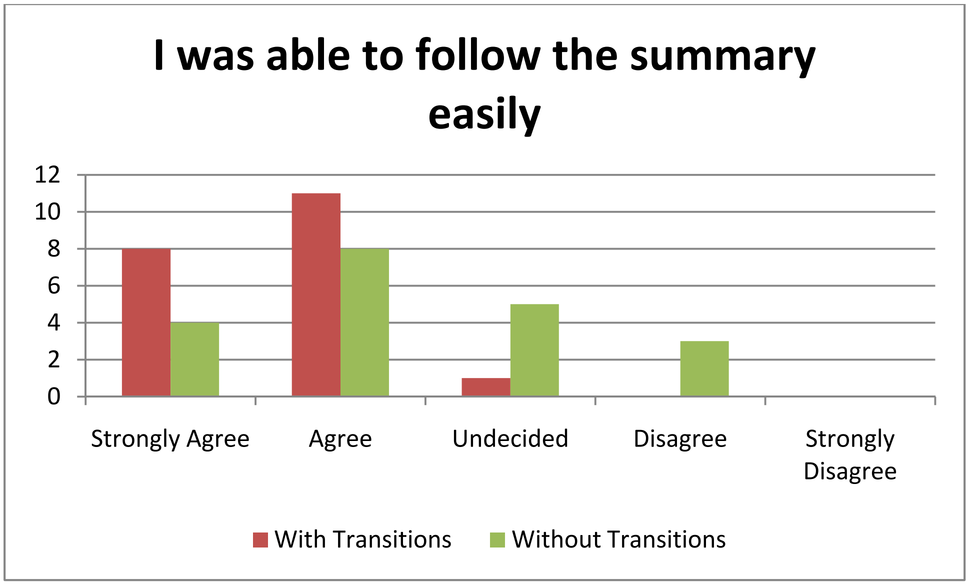


Figure 6.6: This graph shows that viewers followed the summary a lot easier with the aid of our transitions.

We are also concerned with how well users understood the transitions, and whether or not they provided a sense of cohesion. In Figure 6.6, we see that almost all of our viewers agreed with the statement "I was able to follow the summary easily" after watching the summary with transitions. Compared to our summary without transitions, we see that adding transitions greatly increased the viewer's ability to follow the summary. From this we can infer that the transitions were easy to follow and provided coherence for the entire summary.

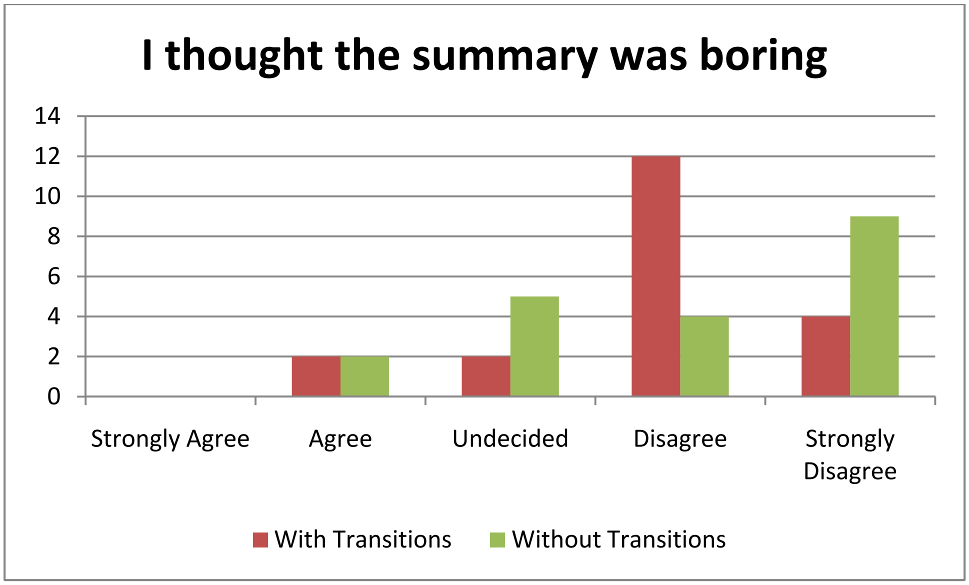


Figure 6.7: This figure shows that our transition summary wasn't too distracting.

Interestingly enough, when given the statement "I thought the summary was boring," results from the user survey indicated that viewers found the summary with transitions slightly more boring than the summary without them (See Figure 6.7). We believe this is because of the small interruption caused by the transition repeatedly popping up. During the transition, the video is paused, causing a gap in the audio as well. Users might have found this to be somewhat distracting. This gap could be reduced or eliminated in a more professional version of this type of transition. Whether or not this gap is the cause of the difference in results, it is significant to note that the difference is small. We can conclude that in spite of the small distraction caused by including the transitions, the transition summary provides a greater benefit overall.

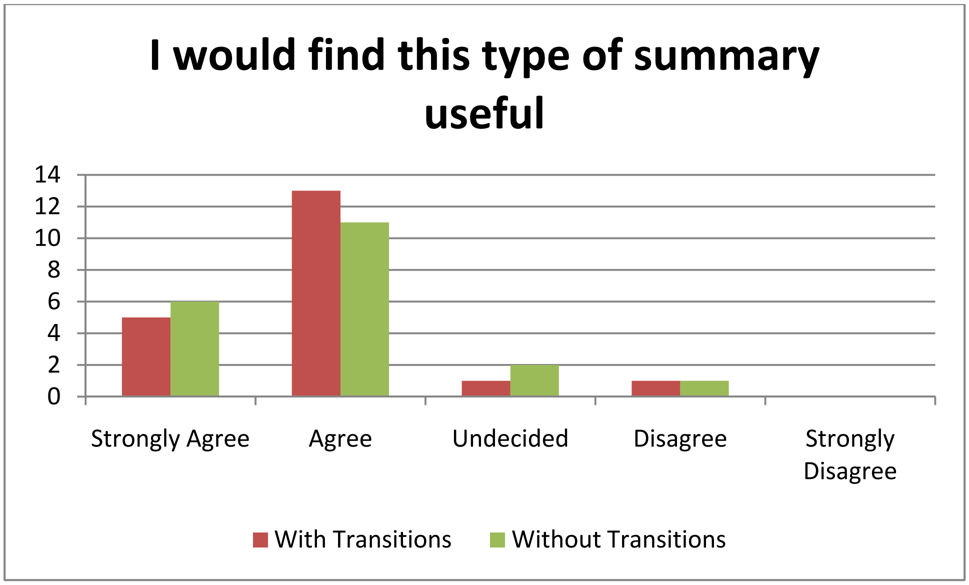


Figure 6.8: This graph shows that viewers found summaries with and without transitions almost equally useful.

Our last statement, “I would find this type of summary useful,” was included to see how much the users appreciated the transitions. Unexpectedly, Figure 6.8 shows that viewers felt about the same for both summaries. Most users found both summaries useful, with only two or three who were undecided or disagreed. Because several of our participants were not football fans, this is not surprising. We expected that more of the true football fans would find the summary with transitions useful. We discovered, however, that including transitions did not seem to influence how useful the viewers perceived the summary. Regardless, this graph does show that viewers are extremely interested in an on-demand summarization service.

The final survey consisted of two parts, one dealing with the transition summary, the other dealing with the interactive controls. Only the statements related to the transition summary are listed here, the interactivity questions are listed in the next chapter. The transition summary statements presented were:

1. The play-by-play transitions helped me follow the game better
2. I had difficulty understanding parts of the transition
3. I understood clearly what the transitions represented
4. I found the transitions confusing and distracting

Each of these statements used the same Likert scale as the other survey.

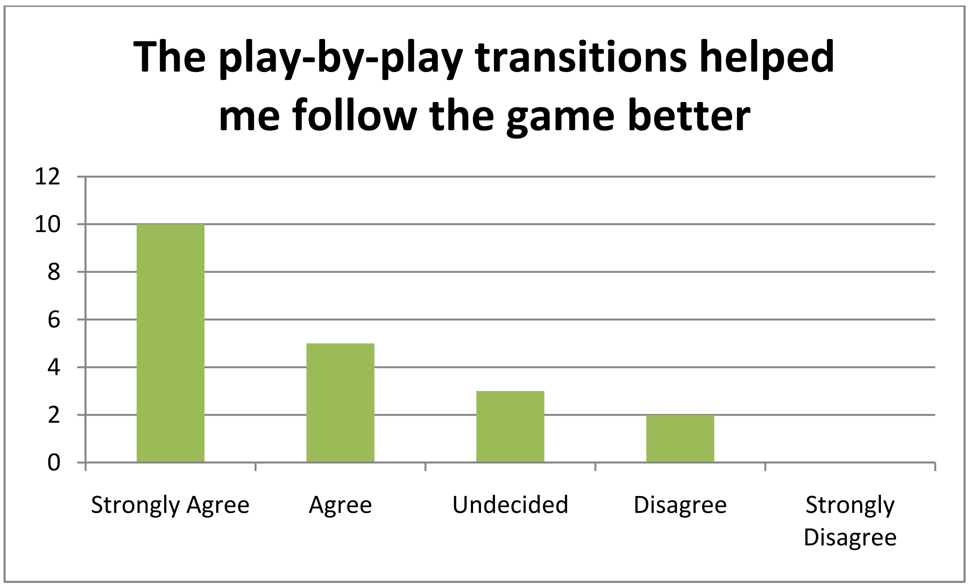


Figure 6.9: This graph shows that viewers liked the information the transitions provided.

Figure 6.9 shows the results of the statement "The play-by-play transitions helped me follow the game better." Over half of the participants felt very strongly that the transitions helped them to follow the game. This strongly supports our claim that the transitions help viewers to maintain context, and that they provide coherence for the summary. During the open-ended questions, three of the participants identified the difference between a highlight reel and a summary, and expressed how the transitions helped to create a better summary. Four participants directly compared the summary with ESPN SportsCenter and pointed out that our summary was much more in-depth.

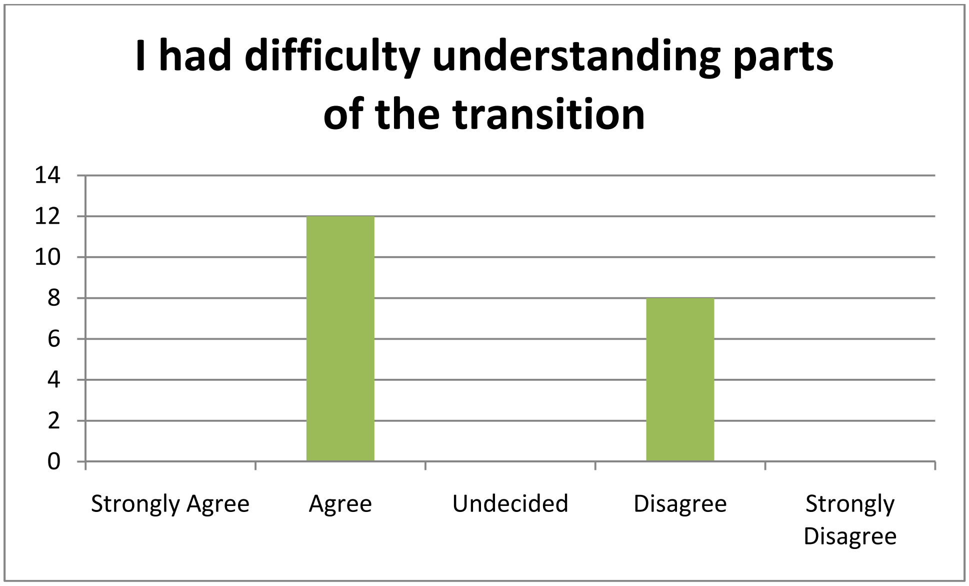


Figure 6.10: This graph shows the split in users over understanding of the transitions.

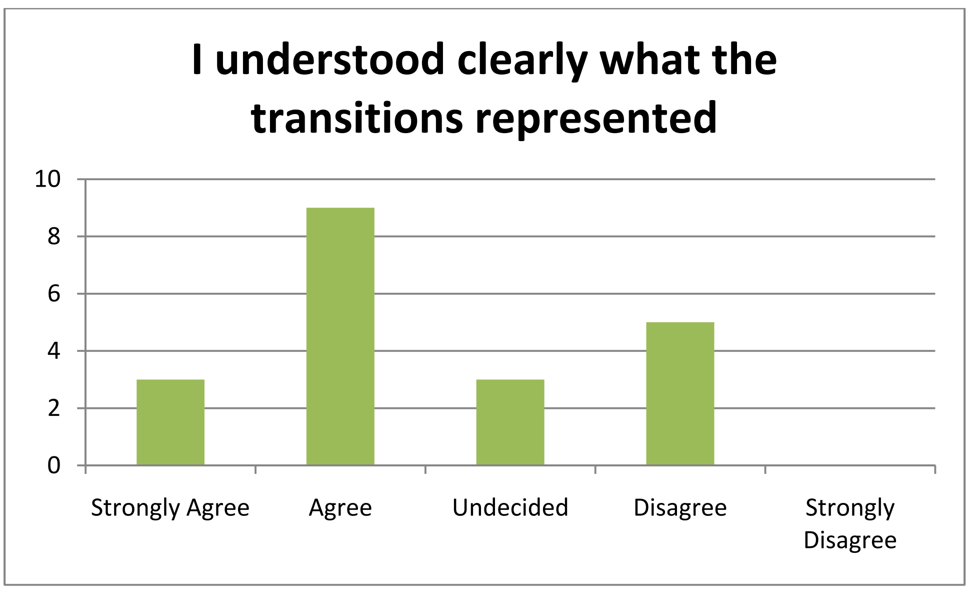


Figure 6.11: This graph shows that most users felt the transitions were pretty clear.

We see some major disagreement on the statement “I had difficulty understanding parts of the transition” in Figure 6.10. Over half of the participants agreed, showing that many of them were confused by the transition. Oddly enough, on the next statement, “I understood clearly what the transitions represented,” most people agreed (See Figure 6.11). This seems unusual until we look at some responses to the related open-ended questions. When asked directly, eighteen of the twenty participants responded that they “eventually figured out” what the transitions meant. As previously mentioned, participants were given no direction as to what the transitions would look like, or what they meant. It is natural that most participants would be confused at first. Our open-ended questions, however, reveal that almost all of them understood the transitions in the end, with only two that remained slightly confused.

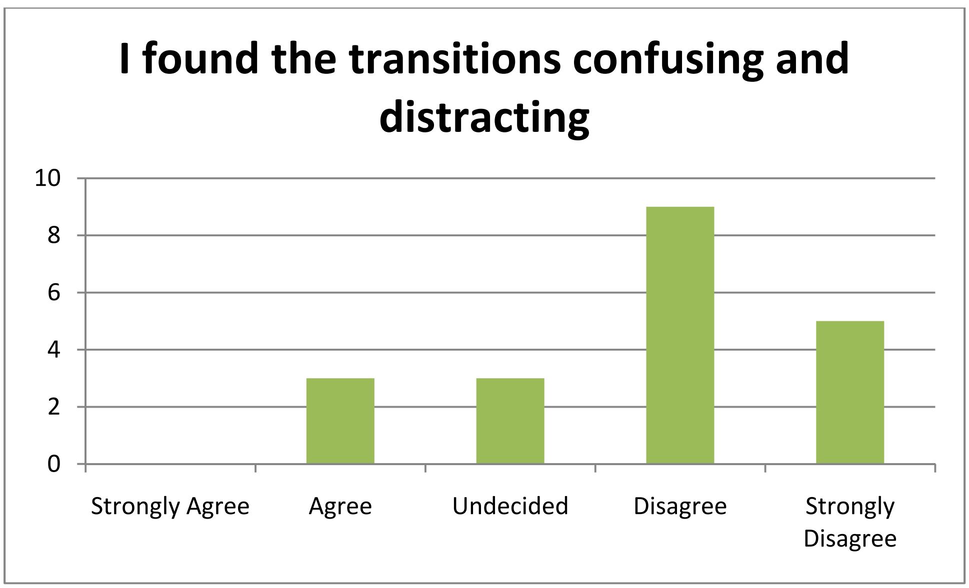


Figure 6.12: This graph shows that few people felt the transitions were distracting and confusing.

Our last statement, “I found the transitions confusing and distracting,” is shown with its results in Figure 6.12. We see that there were a few who found the transitions confusing and distracting, but most did not agree. This is similar to our results from the statement “I thought the summary was boring," and is consistent with our hypothesis. In response to our open-ended questions, four of the participants replied that the transitions were distracting and three mentioned that they found them unnecessary. Included in the four that found them distracting were the two who didn't quite understand the transitions. This suggests that while generally appreciated, there were a few that did not find the transitions useful. We consider this an acceptable trade-off. While the transitions can be somewhat distracting to some, the information they provide for most viewers easily counterbalances this problem.

Of the other comments made during the open-ended questions, we found that the main reason viewers found the transitions distracting was that the animations were "too cute." We found this significant because it indicates that the biggest disagreement among the participants was an aesthetic one. Many of the other disagreements were also matters of preference. Five of the participants suggested including the results of a penalty call in the transition, and six pointed out that the actual transition between one play to another was choppy, referring to the skip in audio and video. From this we conclude that almost all of the participants found the purpose of the transitions clear, in spite of having disagreements with style decisions.

We believe that the results of the user study clearly indicate that our transitions aided the viewers in understanding the true story of the game. Most participants felt that they saw more of the game with the transitions, and were able to follow the summary better. Furthermore, even with the minimal exposure they received to the transitions during a ten minute summary, viewers were able to quickly interpret each symbol and extract the relevant information. We conclude that our transitions provide the right context to help viewers follow the game better, as well as provide a sense of coherence to the entire summary by telling more about how the game progressed.

Chapter 7 – Improving Personalization with Interactivity

No summary will please everybody. All football fans are unique, and are interested in different parts of the game. One advantage we have with ITV is that our solution does not have to be static. By making our summary dynamic, we can better meet the needs of every football fan. In doing this, however, we want to be careful to preserve the primary goal of telling the story of the game. Customizing content to match a viewer's interests can lead to a misinterpretation of how the game actually played out. Instead, we want to provide the viewer with the option to take control of the summary. He or she can drop out and interact with the game without affecting which plays are summarized. This allows us to maintain the story of the game while still providing the viewer with the ability to change that summary.

Figure 7.1: Viewers can use the Summary command to stop and start the summary at any time.

To provide that control, we extend the interactive controls of the TWS player to work with our summarization service as well. Our methodology is simple: When the viewer is watching a summary, he or she is already in "summary mode." At any time while watching the game, the viewer can perform an action on that play, in much the same way he or she would do so while watching the game normally with TWS. At that point, the viewer drops into a regular viewing mode. He or she can skip forward and backward, change the camera angle, and/or view statistics. Once the viewer is satisfied, he or she can press another button to resume viewing the summary (See Figure 7.1). The system is put into summary mode once again, and that summary continues from the viewer's current position in the game.

For example, while watching our summary, a fan may notice something in the transitions that catches his or her attention. By simply pressing “Previous Play,” he or she can drop out of summary mode and quickly skip back to that particular play. If a specific highlighted play catches the viewer’s attention, he or she can review it from several different camera angles. The freedom to explore means that the viewer will not be left unsatisfied with the summary. They do not have to wait for the end of the whole summary to review the part of the game they were interested in; they can do so immediately. Resuming the summary is just as easy.

Figure 7.2: A summary icon in the upper left corner of the screen helps the user know if they are in "Summary Mode."

When the summary begins, a small icon in the upper left indicates that the user is currently in summary mode (See Figure 7.2). This provides context for the user, allowing them to determine whether the summary is currently playing or whether they are in a regular viewing mode. The icon disappears when they exit summary mode and reappears when they resume the summary. It also flashes momentarily when summary mode is reactivated, providing further visual cues for the viewer.

This simple but useful method of interacting with the summary helps to make our summary more dynamic. In addition, we also allow the viewer to select the length of the summary beforehand. This creates a more dynamic summary, since the viewer can indicate how much of the game he or she is actually interested in. If the viewer just wants to see the big plays of the game, he or she can specify a small amount of time. Alternatively, he or she can specify a large amount and watch some of the more interesting—but less important—plays. Thus, the viewer can take control of the summary in two unique ways. A more involved fan might request a short summary, then spend a lot of time exploring by using the interactive controls. Other, more casual fans might simply enter a longer time and trust the summary to show them the most interesting plays of the game. Allowing the viewer to specify the duration as well as take control via the TWS system puts control into the viewer's hands.

Validation

We evaluated our interactive solution with the same user study that was used to evaluate the transition summary. Before actually testing our system, each user was given 15 minutes to familiarize themselves with the TWS controls. Explicit directions were not given on how to use the system; rather, they were shown how to access the menus of the TWS player, and were then allowed to explore the capabilities. We assumed that this would teach them how to use the different menus in order to figure out which options were available to them, and that this would help them learn how to use the interactive controls with the summary. After familiarizing themselves with TWS, the users were shown both summaries, and told that the interactive controls could still be used during both of them.

To create a more realistic situation, participants were told that they had 15 minutes before work to watch the summary of a game they had missed the previous night. Since they were shown a 10 minute summary, they had room to experiment with the interactive controls. They had to keep themselves on track, however, in order to see the whole summary within the allotted time. During the final survey (after both summaries) they were presented with an additional four statements, which were:

1. I liked being able to drop out of the summary at any time.
2. I was confused when trying to interact with the game during a summary (i.e. switch camera angles, rewind or skip, etc.).
3. I felt that being able to interrupt the summary made the summary more difficult to follow.
4. In the future, I would be more likely to use the interactive controls during a summary.

Like the surveys presented before, these questions were answered on the same Likert scale. Additionally, we included open-ended questions related to the interactive controls, which were:

- When and why did you use the interactive controls during the summary? If not, why didn't you?
- What difficulties did you find using the interactive controls?
- What did you like about the interactive controls?

Discussion of common answers to these questions is included in our analysis of the survey results.

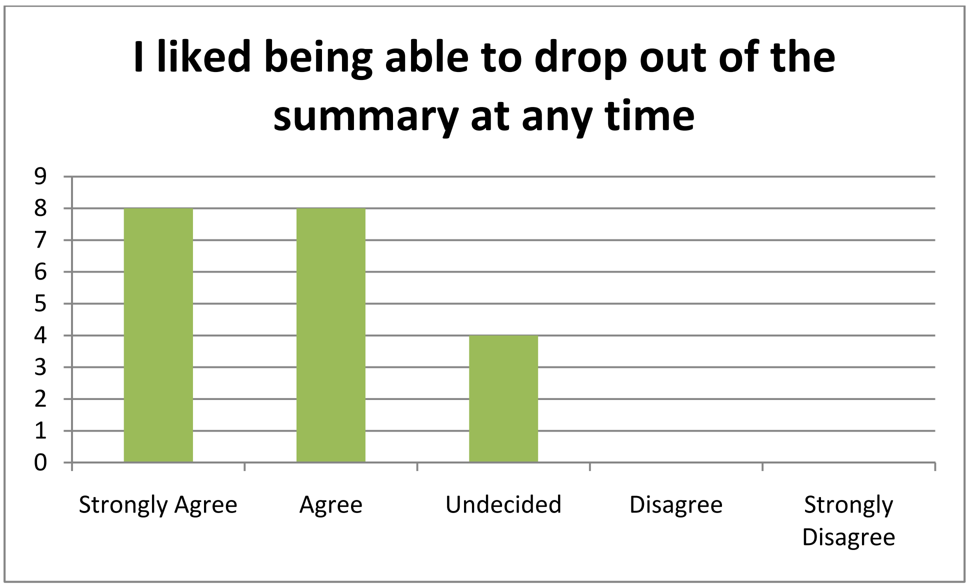


Figure 7.3: This graph shows that almost all users liked being able to interact with the summary.

Figure 7.3 shows the response of viewers to the statement "I liked being able to drop out of the summary at any time." As we expected, viewers enjoyed being able to take control of the summary. Only a few of the participants were unsure. In response to the open-ended questions, twelve participants mentioned that they used the interactive controls to review a play, either from the same or a different camera angle. Six mentioned that they used the controls to skip ahead in the summary, and two mentioned that they used the controls to skip back. Five users mentioned that they really liked being in control. We consider this to be strong evidence that interactive controls help to improve the summary.

To better understand how the interactive controls were used, we also logged every action that the participants made while watching both summaries. We modified the logging system previously mentioned to include an entry whenever the viewer entered or exited summary mode. We then used those log entries to compute how often users dropped out, and then how long they spent out of summary mode. During the two summaries, users dropped out 6 times on average with a standard deviation of 4. This shows that users were indeed interested in interacting with the summary. When dropping out, users stayed out of summary mode for about 41 seconds on average, with a standard deviation of 22.5 seconds. The average play length was 10.7 seconds with a standard deviation of 3.9, so users may have watched a play 2 or 3 times, or may also have backed up to watch a skipped play.

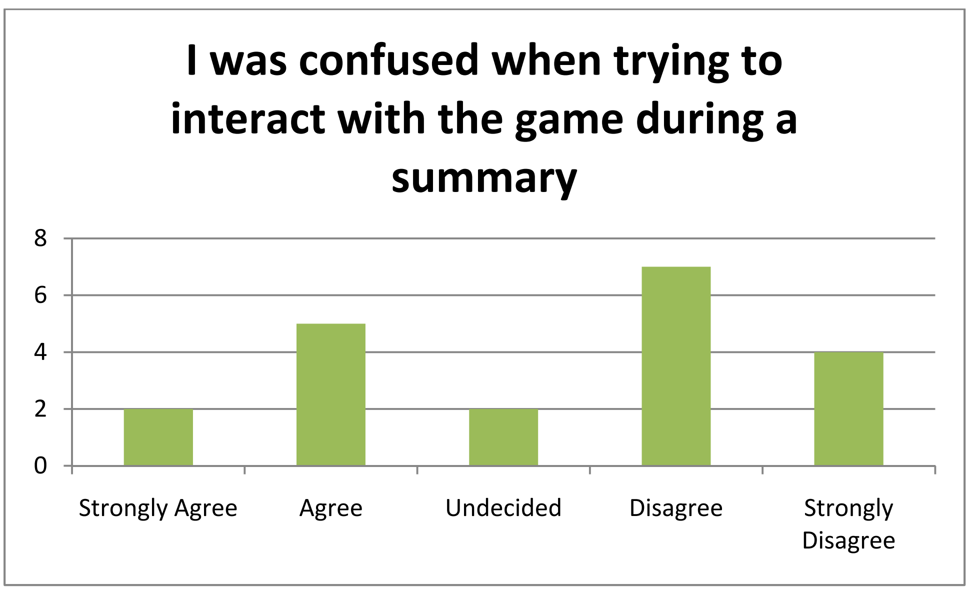


Figure 7.4: This figure shows that there was considerable confusion with the interactive controls.

In Figure 7.4, we see that many of the participants agreed with the statement "I was confused when trying to interact with the game during a summary." While there was still a large part that disagreed with this statement, we see that many users were not very clear on how to use the interactive controls. During the open-ended questions, users indicated that the biggest source of confusion arose from not being aware of whether they were in summary mode or not. In fact, four of the twenty users had to be reminded that they were not in summary mode during the study. Two users also responded that they were not aware that the summary drove itself, nor that they did not have to skip ahead.

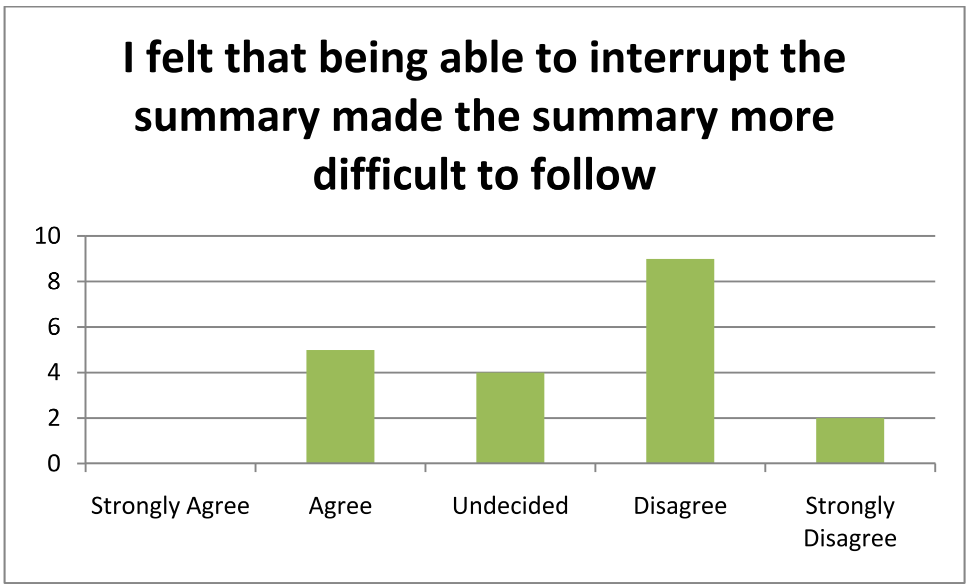


Figure 7.5: This graph shows that some users had a harder time following the summary when interacting with it.

There were a few users who had a harder time following the summary when interacting with it. Figure 7.5 shows users' responses to the statement "I felt that being able to interrupt the summary made the summary more difficult to follow." It is possible that those who agreed were expressing their frustration with the TWS interactive controls in general. Most users, however, did not agree that the controls made the summary difficult to follow.

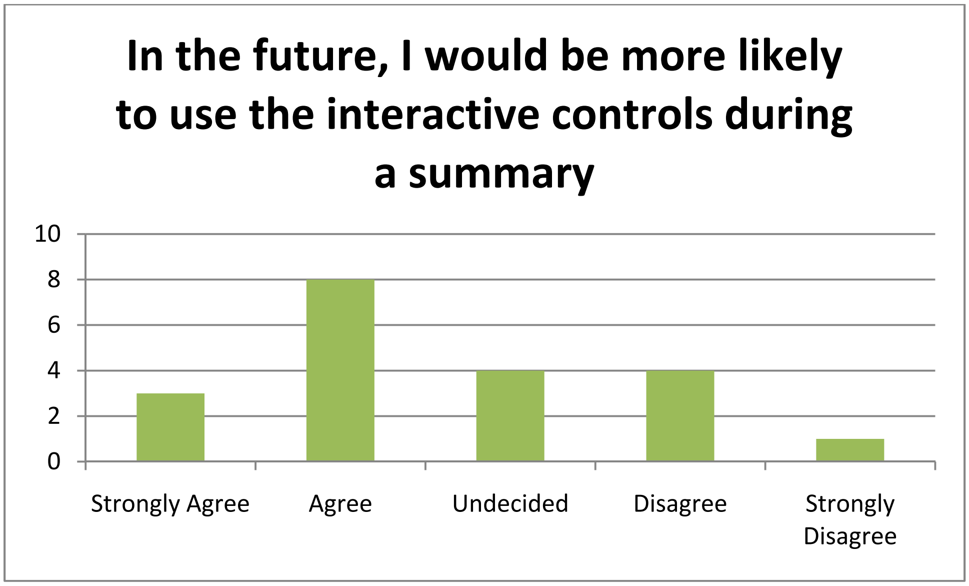


Figure 7.6: This figure shows how likely viewers would be to use the interactive controls in the future.

Figure 7.6 shows that although many users agreed with the statement "In the future, I would be more likely to use the interactive controls during a summary," there were some who disagreed. This seems to conflict with the response to our first statement, in which most users agreed that they liked being able to drop out of the summary at any given time. In response to the open-ended questions, three users mentioned that they would not use the interactive controls because they felt the summary was adequate. We assume this means that while people like having the option of dropping out of a summary to explore the game deeply, many people trust the summarization service and would not frequently use the interactive controls. It is also possible that the few who were confused by the controls felt that they would be less likely to use them in the future.

The results of the user study show that viewers enjoy being able to take control of the summary. As we expected, viewers reported that they used the controls mainly to review a play from the summary at different angles, or to review skipped plays that they thought might be interesting to them. There was some confusion concerning the controls, but we feel that they could be enhanced so that they become more intuitive to the user. In spite of the confusion, however, participants expressed a lot of interest in an interactive summary. This style of interaction creates a flexible, dynamic summary solution that puts the power in the viewer's hands.

Chapter 8 – Conclusion

With the rapid growth of Internet Television, there is a need for an on-demand dynamic football summarization service. We have demonstrated an effective solution that presents the viewer with an interactive summary that does more than just find highlights—it tells the story of the game. By integrating game statistics that have been extracted from an annotation file with viewing statistics drawn from the logs of how users interacted with the game, we generated a DOI that effectively ranked plays both by importance and by interest. This allowed us to capture those "cool plays" which interest viewers, even if they do not result in a score. Once the key plays have been selected, we introduce short animations to ease the transition between plays. These transitions help the viewer to understand what is happening in the segments of the game they do not see. We allow the viewer to input how long a summary he or she wants to see, and also allow the viewer to interact with that summary. He or she can drop out and explore the game at his or her leisure.

We have shown how our solution satisfies the four "C's" defined by He, Sanocki, et al. [He, L. et al. 1999]: *Coverage*, *Conciseness*, *Context*, and *Coherence*. Our use of a DOI function that uses both game events and viewing statistics to rank plays gives us a good balance between *coverage* and *conciseness*. By ordering plays according to viewer interest and game statistics, we make sure we cover the plays that viewers want to see. This also ensures that we are *concise*, since each play added to the summary adds the next most important play in the game. The transitions help the viewer establish the proper *context* by showing what happened before each play. They also add a sense of *coherence* to the summary, taking what might otherwise be a selection of disjointed plays and tying them together.

We also ensure our summary is dynamic by allowing the viewer to determine the length of the summary. This gives the viewer the option of watching the most important plays in a shorter summary, or watching more of the details in a longer summary. We also allow the viewer to control his or her version of the summary by providing interactive controls. The viewer can skip backwards, forwards, and review any play of the game, however they prefer. This creates a dynamic environment that gives the viewer the freedom to watch the summary in the way he or she wants to. Having a dynamic summary allows us to support the innovation of ITV.

Our solution opens several options for future work. First, we believe this style of summary can be expanded to work with any sport. Since TWS has been designed to work for all sports, we believe that we can collect the same type of information for each one; and with the addition of customized transitions, we can provide an efficient summary. Second, we believe this summarization technique can be expanded to work with live broadcast games. In this way, viewers who come to the game late can "catch up" with live time through the summarization service. This idea could also be used as an effective way to keep up with two games simultaneously.

References

Agnihotri, L., Kender, J., Dimitrova, N. and Zimmerman, J. (2005). Framework for Personalized Multimedia Summarization. Proceedings of the Seventh ACM SIGMM International Workshop on Multimedia Information Retrieval. Hilton, Singapore, ACM**:** 81-88.

Babaguchi, N., Kawai, Y. and Kitahashi, T. (2002). Event Based Indexing of Broadcasted Sports Video by Intermodal Collaboration. IEEE Transactions on Multimedia **4**(1): 68-75.

Babaguchi, N., Kawai, Y., Ogura, T. and Kitahashi, T. (2004). Personalized Abstraction of Broadcasted American Football Video by Highlight Selection. IEEE Transactions on Multimedia **6**(4): 575-586.

Chu, W.-T. and Wu, J.-L. (2008). Explicit Semantic Events Detection and Development of Realistic Applications for Broadcasting Baseball Videos. Multimedia Tools and Applications **38**(1): 27-50.

He, L., Sanocki, E., Gupta, A. and Grudin, J. (1999). Auto-summarization of Audio-video Presentations. Proceedings of the Seventh ACM International Conference on Multimedia (Part 1). Orlando, Florida, United States, ACM**:** 489-496.

Logan, J. D., Durgin, S. A. and Burd, D. (2005). Methods and Apparatus for Recording and Replaying Sports Broadcasts. U. S. Government. United States, Gotuit Video, Inc. **US Patent 2005/0005308 A1**.

Ma, Y.-F., Lu, L., Zhang, H.-J. and Li, M. (2002). A User Attention Model for Video Summarization. Proceedings of the Tenth ACM International Conference on Multimedia. Juan-les-Pins, France, ACM**:** 533-542.

Nair, R. (2004). Calculation of an Aggregated Level of Interest Function for Recorded Events. Twelfth Annual ACM International Conference on Multimedia. New York, New York, United States, ACM**:** 272-275.

Takahashi, Y., Nitta, N. and Babaguchi, N. (2005). Automatic Video Summarization of Sports Videos Using Metadata. Advances in Multimedia Information Processing - PCM 2004, Springer Berlin / Heidelberg. **Volume 3332/2005:** 272-280.

Tjondronegoro, D., Chen, Y.-P. P. and Pham, B. (2003). Sports Video Summarization Using Highlights and Play-breaks. Proceedings of the Fifth ACM SIGMM International Workshop on Multimedia Information Retrieval. Berkeley, California, United States, ACM**:** 201-208.

Tong, X., Liu, Q., Zhang, Y. and Lu, H. (2005). Highlight Ranking for Sports Video Browsing. Proceedings of the Thirteenth Annual ACM International Conference on Multimedia. Hilton, Singapore, ACM**:** 519-522.

Truong, B. T. and Venkatesh, S. (2007). Video Abstraction: A Systematic Review and Classification. ACM Transactions on Multimedia Computing, Communications, and Applications **3**(1): 3.

Zheng, Y., Zhu, G., Jiang, S., Huang, Q. and Gao, W. (2007). Highlight Ranking for Racquet Sports Video in User Attention Subspaces Based on Relevance Feedback. IEEE International Conference on Multimedia and Expo. Beijing, China**:** 104-107.