Utilizing Statistics and Machine Learning to Detect Events and Summarize Basketball Game Footage

Kenna Zimmerman, B.S. Computer Science

- The NBA yields billions of dollars each year
- Games average 2+ hours, with up to 13 games a day
- Busy fans can struggle to catch all their own teams' games, let alone all 30 teams
- Much of the game footage is irrelevant or uneventful
- Highlights are entertaining but don't provide entire summaries of games

Figure 1: Closeups of players and coaches are often shown but are not relevant to events in the game.

Goals

- Create a computer model that distills games into only the most exciting and pertinent events
- Maintain the essence of games by playing clips in sequential order and retaining all important events

Methods

- My proposed model takes in a video file of a full-length basketball game. Every 10th frame is then analyzed, and the following data is recorded:
    - Timestamp of frame
    - Quarter and time left (via OCR model)
    - Camera angle type (close-up or profile, via machine learning model)
    - Volume (in dBFS, closer to zero = louder)
    - Camera movement

Figure 2: Example of a frame being processed

Methods, cont.

- Certain events are more correlated with sound and frame differences than others, as shown in table 1.
- Then, the most exciting clips are chosen by ranking the sound and movement measurements
- These clips are then played sequentially, without including any frames classified as closeups

Event	Movement p-val	Sound p-val
Free Throw	**2.01E-08**	**6.28E-17**
Foul	**4.96E-04**	**4.96E-12**
Enters Game	**1.65E-03**	**1.00E-03**
Misses 2-pt	**2.00E-06**	**1.25E-03**
Misses 3-pt	**2.50E-02**	**5.08E-03**
Makes 2-pt	**1.00E-05**	**6.15E-03**
Timeout	2.62E-01	5.19E-02
Makes 3-pt	**1.10E-03**	1.10E-01
Turnover	**2.70E-02**	1.46E-01
Violation	7.98E-01	2.53E-01
Rebound	**3.80E-08**	9.63E-01

Table 1: Results of statistical analysis on the difference in pixel difference and sound level between events and non-events. Green cells indicate $p \leq 0.05$

Preliminary Results

- Summary created where clips selected on sound and movement data: youtu.be/qKqVldvmsdM
    - Includes many irrelevant clips and clips that run for too long, yet also excludes many relevant clips
- Summary created where clips selected using "play-by-play" data on the game: youtu.be/cJcVgqv1cOQ
    - Much more relevant selections than sound and movement alone, depends on data outside video

Conclusion / Next Steps

- First results are promising for creating comprehensive summaries of many NBA games.
- Further work includes:
    - Improving selection of beginnings and endings of clips by incorporating sound/movement
    - Automating the entire process (including downloading video and play-by-play data)
    - Adding explanations of what event is taking place in the clip and displaying relevant statistics
- Potential real-world applications of this model are:
    - Website or mobile app where users can watch these summaries
    - Partnership with NBA or other sports associations. Would allow for use of copyrighted footage
    - Potential use of model in NCAA basketball or different sports

References

Li, B., & Sezan, M. I. (2001). *Event detection and summarization in sports video.*

Saur, D., Tan, Y.-P., Kulkarni, S., & Ramadge, P. (1997). *Automated analysis and annotation of basketball video* (Vol. 3022).

Xu, M., Duan, L.-Y., Xu, C., Kankanhalli, M., & Tian, Q. (2003). *Event detection in basketball video using multiple modalities.*