Multi-modal Tools for Keeping Statistics of Ultimate

Hartti Suomela
210 Panama Street, Stanford
hartti@csli.stanford.edu

ABSTRACT
Statistical analysis of plays and players has become more and more important for team sports. In professional sports leagues coaches and managers use detailed statistics to gain competitive edge in player recruitment and strategy development. Statistics can also benefit media and sports fans, as well as teams among amateur sports. Unfortunately amateur sports teams – including all ultimate teams – lack the resources and tools to gather data during games. This paper studies the challenges and technologies in real-time event data gathering from ultimate games, and describes tests of an initial prototype in the process of implementing such a system.

Author Keywords
Team sports, multi-modal, speech, pen-input, real-time, event recording

ACM Classification Keywords
H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

ULTIMATE [1]

Ultimate is a team sport played by two seven-player squads with a plastic flying disc on a field similar to football. The field is 120 yards long and 40 yards wide with 25-yar deep end zones at both ends of the field (see Figure 2).

The object of the game is to score by catching a pass in the opponent's end zone.

Figure 1: Ultimate is team sport played with plastic disc.

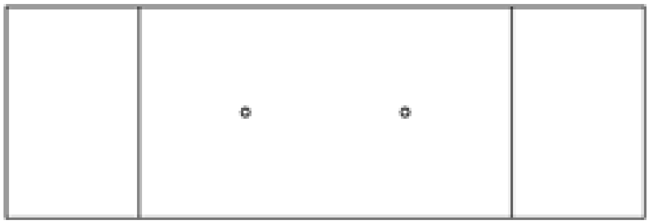

Figure 2: Field layout for ultimate.

A player must stop running while in possession of the disc, but she may pivot and pass to any of the other receivers on the field. When pass to teammate is incomplete (a dropped pass, an interception, a pass out of bounds) or when a player is caught holding the disc for more than ten seconds, a change of possession occurs and the team previously on offence is now on defense and vice versa. Ultimate is governed by Spirit of the Game; a tradition of sportsmanship that places the responsibility for fair play on the players rather than referees.

Physical contact should be avoided. If such contact happens, the fouled player will notify this with "Foul" call, and the disrupted event is replayed. There are also other calls, which will lead to replays.

Player substitutions are allowed only after point is scored and before the next point is started with a pull (similar to kick-off in football). Between each point the offense directions are reversed.

Requirements for a event recording system

The rules of the game, the conditions during game (weather, audio environment), statistical and analysis needs of various stakeholders, and the amateur status of the sport impose restrictions and requirements for a system (not including the requirements for system architecture and interfaces to other systems):

- Reliable / no crashes or data loss (must)
- Fast event input (must)
- Easy to learn (must)
- Easy to use, including on-site setup (must)
- Separate people has to be able to reproduce the same resulting data from a single game (barring a human error, e.g. lapse of concentration)

- Accuracy independent of person managing the event recording[1] (optimal accuracy ±1 yards, acceptable accuracy ±2.5 yards)
- Should support event recoding also from video material (taped games)
- Need for undo capability (must)
- System must not require consent from the players to allow for example opponent scouting (must)
- System should be able to produce data for analysis during a game, although not necessarily all statistical analysis needs to be possible. An example of real-time data need for a coach is "points and minutes" played by a player, which allows him to select fresh legs on the field.
- System should be weatherproof, or at least it should be easy to protect it against weather.
- System must allow people to record events without having a table to rest the devices on.
- System must allow people to record events while standing (one-handed input, other hand might be needed to hold the devices).
- Two or more people should be able to input event data synchronously to a single data store (nice to have)
- Battery-life for the system longer than 2 hours.

Past and Existing Solutions for Ultimate

A paper and pen based statistics system RUFUS (Refined Ultimate Frisbee Uniform Score) was developed in late 80s and it was adopted by the UPA for maintaining a permanent record of all Nationals level competition. It required one person to speak out the events on the field while the other person was writing down or entering the data directly to a computer using a keyboard.

UltiStats [2] as seen in Figure 3 is a Palm OS application for gathering event data from ultimate games for statistics. The event data is inputted to the application using various forms (on-screen buttons and like) using the stylus. As the screen is small and pointing the buttons requires accuracy (one has to remember the real-time requirements) it helps having another person speaking out loud the events on the field. The system also provides some summary screens of the data. The application is Open Source.

[1] In here, accuracy is a relative requirement. The field dimensions are usually measured using strides and not with tape measure resulting in approximations. Also the sidelines are seldom marked with lines, but only intersections are marked with cones (providing less visual cues for the camera)

Figure 3: UltiStats is a form-based Palm OS application for ultimate statistics.

There are also a couple of electronic score keeping systems for ultimate, which do not allow one to record other events than scoring. ScoreKeeper [3] was developed for WFDF world championships using J2ME application running on a Symbian phone and the Score-O-Matic [4] uses touch tone input of scores analyzed at a central server.

Events to be recorded during Ultimate Game

The following list provides an overview of the primary events to be recorded from an ultimate game to allow analyze the game in detail. The examples provided for each case contain ultimate jargon, and can be therefore hard to understand for a non-player.

- Identifying players (lineup on field, throwers, markers)
- Thrower/catcher locations (optionally including receiver's run direction during catch)
- Pass length (e.g. short, medium, long/huck); optional, to be recorded if thrower locations are not recorded
- Pass type (e.g. dump, swing, attacking); optional, to be recorded if thrower locations are not recorded
- Throw type or the pass (e.g. backhand, sidearm, hammer, other)
- Catch type (really optional, as does not provide any relevant information for the analysis)
- Pass result (e.g. throw-away, drop, half-drop-half-throwaway, interception, point block, stalled)
- Force direction of the marker (e.g. angle or word definitions like straight-up, force-flick, force-backhand, force-middle, no-dump)
- Stall count at pass-time (might be possible to estimate this based on catch and throw timings)
- Foul-calls including contest or not-contest information (also a categorization might be useful, e.g. throwing, receiving, away-from-the-disc)

- Pick-calls
- Travel-calls
- Wind direction
- Defense formation (e.g. simple categories like Man, Junk/Clam, Zone, Zone-to-Man)
- Offense formation (e.g. simple categories like Stack, Horizontal, Spread)

DISCUSSION OF PROSPECTIVE TECHNOLOGIES

Because the system should allow event recording also without consent from the tracked team to enable opponent scouting, sensor-based solutions are not suitable.

Vision based solutions (see for example [5] and [6]) can be used to pinpoint the thrower locations on the field. The on-site setup of this kind of system is laborious – either there are multiple cameras, which need to be accurately located for the system, or then a single camera needs to be installed high above the playing ground. This is not usually possible in tournaments much less during practices. Additionally the lack of marking lines on the field (like available in football) makes the calibration of the camera(s) harder.

Many professional football and basketball teams as well as some NCAA teams use manual video analysis in their opponent scouting (also called advance scouting). This process includes someone to manually annotate the video of the game with player information, offence and defense formation information, etc. The marked clips can then be sorted and viewed in various ways, but no statistical data is gathered during this process. Such solutions are provided by for example Pinnacle Systems [7] and Sports Tec International [8].

Keyboard based solutions usually require two hands to operate the keyboard efficiently and a place to rest the keyboard on. Also keyboards are suitable for inputting quantified data, but fast, accurate input of field positions is not possible in real time. One solution is to create a grid for various field locations and devote one key for each grid. As an example, a really rough grid could be:

- Own end zone
- From end zone to own brick mark
- From brick mark to mid-field
- From Mid-field to opponent brick mark
- From brick mark to opponent end zone
- Opponent end zone.

However this does not fulfill the accuracy requirements describe above. Also memorizing tens of different shortcuts takes time (even with specialized keyboard templates), which is in conflict with the easy-to-learn requirements.

Even with these limitations a leading sports statistics company Stats Inc uses keyboard input to record events from various sports. Usually the keyboard input is done while monitoring a recorded video feed of the game, allowing slower pace, and correction of errors.

Pen-input alone might not be adequate or fast enough to record all needed events during an ultimate game. For example marking menus are not suitable for small screens (if the application is to run on PDAs) and identifying players through these menus is challenging (20+ players on the team, and on field line-up changes constantly). If separate on-screen buttons are used (like in UltiStats), the user has to spend a lot of time finding the correct button and point to it, simultaneously losing the vision of the field.

Adding a voice channel to provide supporting data for the pen-input data, the same way voice is used to enhance the interaction with QuickSet [9] is a valid option. These multi-modal interfaces make it possible to input data more efficiently, but the existing research has not studied the limitations of multi-modal interfaces while inputting a flow of past paced events and having to divide concentration between the screen, speaking and the game events on the field. Also the background noise usual to sporting events might cause problems, unless a highly directional microphone is used

There is one multi-modal statistics software for basketball supporting data input also through voice recognition from Dominion Software called Basketball Statistical Software for Team Athletics. [10]

CALCULATIONS: GOMS AND FITT'S LAW

With GOMS some rough estimates for the time needed to jot down a player position on a screen and returning the focus to the events on the field can be estimated. This should not exceed the expected minimal time interval between events in an ultimate game. Both these times need to be calculated.

Another area of calculations is to use Fitt's law to estimate if the time available for the positioning of the player allows the PDA-sized screens, or does one need to use Tablet-size screens to achieve required input accuracy.

INITIAL HYPOTHESES AND PROTOTYPE

Multi-modal interaction (pen and speech) was selected as target technology for the event input. A paper prototype was develop to test the initial hypotheses:

1. Record keepers (people) prefer bird-eye field of he field to perspective view to mark the player locations
2. People can position players on the field pretty accurately and the accuracy does not degrade much even if they are speaking out loud throw-by-throw commentary.

3. People can position the thrower and draw the marker alignment for every throw, even while speaking.
4. The time between pointing the thrower location and saying out loud the catcher name in pretty constant, so synchronizing the drawing and the speech is straightforward (also this would allow to distinguish the high stall-count situations).

As the testing situation did not allow placing video camera high above the field of play to automatically position the players the second hypothesis was slightly modified:

2. Different people position the players on the field at same locations and the accuracy does not degrade much even if they are speaking out loud throw-by-throw commentary.

The prototype created to test these hypotheses was a paper prototype. Field templates were printed on paper. Stacks of these templates were then placed into a binder.

TESTING THE PROTOTYPE

A test with two test subjects was carried out during a scrimmage tournament at Stanford, where the fields were marked with cones only. The test subjects had knowledge of the game and also knew the names of most of the players in the recorded team. The test setup consisted of: two video cameras, a tripod, two binders with various field templates (one page per each possession), and two pens.

One video camera standing on a tripod was placed at the one end of the playing field to allow capturing the events on the field without operator.

The other video camera was running all the time during the test but operated manually by the evaluator to capture the actions of the test subjects and also to capture the audio (part of the test required the test subject to speak out loud the events on the field while drawing the throwing locations). The videos were later synchronized to allow studying the response times of the test subjects.

Findings

The first hypothesis was not supported by the test subjects' comments. One of the test subjects liked better the perspective view, while the other preferred the bird-eye view. The test needs to be re-evaluated with larger test group, and also the drawing area of the fields need to be equal in that test. In the test described here, the fields were equal in length, not in area, which might cause test subjects to prefer the field with more room for markings (bird-eye field).

The second hypothesis was supported to some extent. The positions as marked by the test subjects were not identical (±5 yards for both X and Y), but the accuracy did not decrease substantially when voice input was added. Unfortunately also the standard deviation of the positions was quite big, meaning that the differences were not constant. This hypothesis needs to be revisited with next prototype and also study situations, where fields are marked using sidelines also.

Also the third hypothesis was shown to be problematic, as the markers are not stationary, but tend to move around the thrower. This makes it hard to align the marker correctly, even if the test subjects were informed to draw the marker location during the pass release. A better approach might be to quantify the marker positions (as described earlier), and the record keeper could then speak out the locations, allowing also temporal adjustment.

Because of problems in the test setup, realized only after testing, we were not able to test the fourth hypothesis. Again this needs to be evaluated with the computerized prototype.

FURTHER WORK

The next step in evaluating multi-modal interfaces in gathering event data from ultimate is to implement a working prototype of such system. The plan is to use Windows Tablet PC and Java environment as development environment.

REFERENCES

1. Ultimate Players Association. http://www.upa.org/.
2. UltiStats. http://www.moshpit.org/pilot/ and http://sourceforge.net/projects/ultistats/.
3. Suomela, H. : ScoreKeeper Abstract in review for HCI-International 2005.
4. Score-O-Matic http://www.scoreomatic.com/
5. Duan, L., Xu, M., Chua, T., Tian, Q. Xu, C. A mid-level representation framework for semantic sports video analysis. *Proc. ACM Multimedia 2003*
6. Müller, B., Anido, R. O. Distributed RealTime Soccer Tracking. *Proc ACM Video surveillance & sensor networks 2004.*
7. Pinnacle Systems Team Sports http://www.pinnacleteamsports.com/
8. SportsCode by Sports Tec International http://www.sportstecinternational.com/
9. Cohen, P., Johnston, M., McGee, D., Oviatt, S., Pittman, J., Smith, I., Chen, L., Clow, J. QuickSet: multimodal interaction for simulation set-up and control. *Proc. SIGCHI 1999.*
10. Basketball Statistical Software for Team Athletics http://dominionsoftwareinc.com/prod01.htm