

# Muithu: A Touch-Based Annotation Interface for Activity Logging in the Norwegian Premier League

Magnus Stenhaug<sup>1</sup>, Yang Yang<sup>2</sup>, Cathal Gurrin<sup>2</sup>, and Dag Johansen<sup>1</sup>

<sup>1</sup> University of Tromsø, Norway

<sup>2</sup> INSIGHT Centre for Data Analytics, Dublin City University, Ireland

**Abstract.** Annotation of content is a key enabling technology for multimedia systems development. In this demonstration, we present a real-time activity annotation interface designed to be intuitive while imposing minimum effort on the user. Our solution is to use a smartphone and implement a tile-based touch interface. The interface was developed as a part of a larger project in collaboration with Tromsø IL, one of the top soccer teams in Norway. In this demonstration submission we present and evaluate the annotation interface of Muithu.

## 1 Introduction and Description

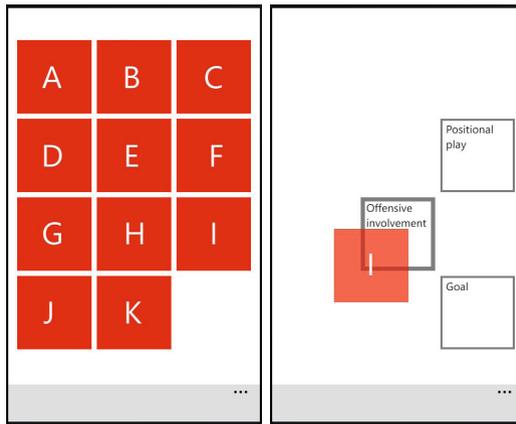
Manual and automatic annotation of content are both key enabling technologies for multimedia processing and multimedia information retrieval. In multimedia processing, annotations have been used extensively for both training and evaluation. Meanwhile, the field of information retrieval has a long history of well designed comparative evaluations using document datasets and annotated relevance judgements. The method of gathering annotations has mainly been an ad-hoc activity carried out after the data has been gathered. Outside of the field of Quantified Self analysis using technologies such as Daytum [1], the concept of real-time annotation of activities have received little consideration, yet they are likely to be more accurate than post-hoc annotations. The main contribution of this work is to demonstrate a low-cost approach for annotating real-world content in real-time on smartphones, or any handheld device.

Called Muithu [4] and part of a larger project (Bagadus [3]), the annotation interface we present in this work was designed to support easy, real-time annotations of player activities on the field. Muithu was evaluated with football coaches, but could be applied to any real-world annotation task. For this implementation, we found that a typical annotation session yielded 16 events of importance on average. Video sequences corresponding to each annotation were captured from different angles, and presented on a website which enables coaches and players to engage in a constructive dialogue concerning each annotated event.

Our approach to designing the annotation interface was to use the (recently popular) tile interface similar to that of Windows Phone. By using our interface, a user can make high-level semantic annotations with any handheld device.

Although we present a demonstration in the area of sports, this annotation interface is suitable for any real-world annotation, such as those needed for lifelogging [2] or video processing [5].

The possible annotation options are structured as a directed acyclic graph. Each of the nodes in the graph contains a short descriptive text, with edges describing the relation between the nodes. Each of the root nodes are represented with their own tile arranged in a grid as shown in Figure 1 - left. The tiles are coloured squares or images, with the nodes description shown as text labels. Simple annotations can be done by clicking the respective tiles. By pressing and holding a tile, more detailed annotations can be made by dragging a tile into a boundary of a secondary tile (Figure 1 - right). These secondary tiles represent the children of a node in the annotations graph. The use of hierarchical annotations allows the user to enter both coarse and fine grained annotations using the same interface.



**Fig. 1.** Two screens of the Muithu interface. Left screen shows the player selection screen and the right shows the options screen for the selected player. The names have been replaced with letters for privacy reasons.

The Muithu annotation interface was designed with the following user experience properties in mind:

**Mobile Friendly Interaction.** One of our design principles is to ensure that users can accurately make annotations. On mobile devices, button size is generally limited by the size of fingers. Small touch targets will potentially lead to big problems. Thus, we adopt big tile square buttons to increase touch targets, better utilize available screen real estate, and prevent errors. We hide the fine grained annotation screen and only display it when the user employs certain gestures.

**Minimise Memory Load.** In order to minimise the user's memory load, the user should not have to remember information from one part of the application to another. For instance, when making new annotations, the user's

attention may be distracted a few times. Hence, we minimise the user’s memory load by making options visible, i.e., display activated grid’s information with transparent effect, overlaying it on top of screen (Figure 1 - right).

**Learnability.** The users interactions should not interfere with the other normal activities, and requires an interface which is easy to learn. We minimise the number of steps it takes for a user to complete an annotation task. Much of the annotation interaction is via touch-and-slide manipulation. Familiarity may come from the fact that it follows standards or that the design follows a metaphor from real world experience.

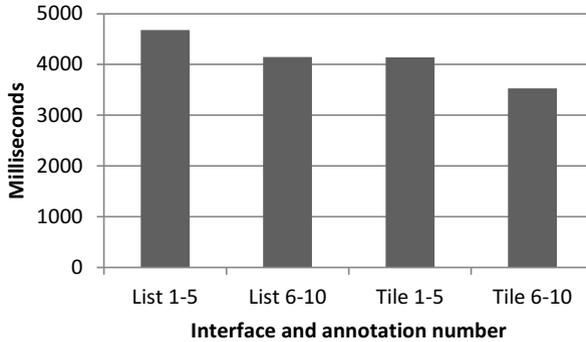
**Adaptive Interface.** We designed the Muithu annotation interface to be extensible, and hence useful across a variety of real-world deployments. Annotation can be added or removed on the fly, making this interface easily configurable and essentially use-case agnostic.

## 2 Evaluation

We present the results of a real-world experiment into the effectiveness of this annotation interface. The user testing the system included a Norwegian Premier League soccer coach (expert) and group of novice users. In [4], we put a soccer coach in charge of operating the system during their national level matches and training sessions. The coach would record high-level important events during the 2012 and 2013 Norwegian Premier League. By facing a camera towards the coach, we measured the time the coach had his attention away from the match in order to annotate. Our experiment showed that the coach would spend around 3.1 seconds on annotating a single event on average.

It is our conjecture that as a result of being able to make an annotation by a single swipe with the thumb, the attention of the user can be maintained on real-world activities. Therefore, we will show that Muithu is 15% faster at annotating than a conventional list-based annotation interface. To evaluate the interface, we measured the time a group of people spent to complete a set of prescribed annotations. A multiple choice drop-down list (two layer hierarchy) provides the baseline for our evaluation. Using a list interface allows us to compare Muithu with a more familiar baseline. Each of the subjects completed a series annotations in succession, altering between the two interfaces, in a real-life environment. Eight users without any prior experience with using the interface completed a series of 20 annotations each, totalling 80 annotations for each interface.

Figure 2 shows the result of our evaluation. The graph shows the average time for annotations one through five and six through ten. The average time used to make an annotation using the list based interface is 4.4 seconds and 3.8 seconds using Muithu’s interface. This is a 15% improvement over the list-based interface. This evaluation does not measure the impact of the learnability and attention-maintaining aspects of Muithu, though we are confident that these are improved also.



**Fig. 2.** The average measured time for the tile and list interfaces. The measurements are grouped into the first five and the last five annotations for each interface.

### 3 Discussions and Conclusions

We have presented our interface for real time annotations. We also compared Muithu with a conventional list-based interface, and showed that our interface improved on the time spent by both experts and normal users to annotate an event. Our experiments shows that the time spent to annotate is around 3.4 seconds from start to finish and that the system was easy to learn. We suggest that this style of user annotation is applicable to many real-world use cases. Aside from use within the sports domain, the use of this type of interface might prove successful in other teaching environments such as driver's education.

### References

1. Daytum: Collect, categorise and communicate everyday data, <http://daytum.com>
2. Aizawa, K.: Digitizing personal experiences: Capture and retrieval of life log. In: Proceedings of the 11th International Multimedia Modelling Conference, MMM 2005, pp. 10–15 (2005)
3. Halvorsen, P., Sægrov, S., Mortensen, A., Kristensen, D.K.C., Eichhorn, A., Stenhaug, M., Dahl, S., Stensland, H.K., Gaddam, V.R., Griwodz, C., Johansen, D.: Bagadus: an integrated system for arena sports analytics: a soccer case study. In: Proceedings of the 4th ACM Multimedia Systems Conference, MMSys 2013, pp. 48–59. ACM, New York (2013)
4. Johansen, D., Stenhaug, M., Hansen, R.B.A., Christensen, A., Høgmo, P.-M.: Muithu: Smaller footprint, potentially larger imprint. In: ICDIM, pp. 205–214. IEEE (2012)
5. Silva, J.A., Cabral, D., Fernandes, C., Correia, N.: Real-time annotation of video objects on tablet computers. In: Proceedings of the 11th International Conference on Mobile and Ubiquitous Multimedia, MUM 2012, pp. 19:1–19:9. ACM, New York (2012)