Kapture

Gesture-based Whiteboard Capture Using the Microsoft Kinect

CS4099 Major Software Project 2012-13

Produced by Craig Garrigan

Supervised by Dr. Tristan Henderson

University of St Andrews

Motivation

Meetings are an everyday part of the busy schedule for staff at the School of Computer Science in which they will commonly use a whiteboard to write down notes and generate ideas along with those in attendance.

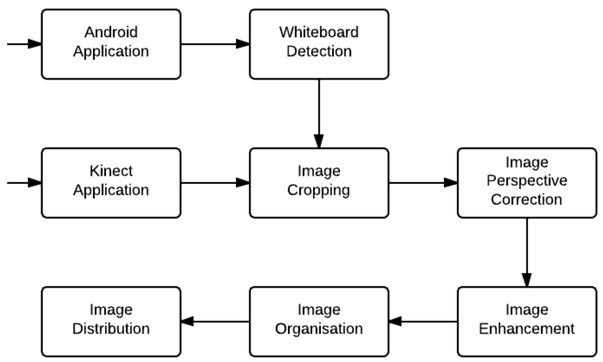
Following each meeting, it is often desirable to keep a record of what was produced by taking a photograph of the whiteboard content for later reference. This might also be followed by sending the image to the other meeting participants for their own record keeping.

Using the wonders of technology, it should be possible to automate the entire process, freeing up the time it would normally take to capture the whiteboard, organise the image and share it with others.

Aim

The aim of this project is to develop a system which uses the tracking capabilities of the Microsoft Kinect sensor to recognise a gesture performed by a user in front of a whiteboard. The Kinect's camera can then be used to capture the contents of the board and make it available in a suitable format for the user to view at a later date. The system should also provide a means of identifying those present when the capture was made and give the user the option to share it with them.

System Overview



- A whiteboard capture is initiated by performing a gesture in front of the Kinect
- Alternatively, an Android app has been produced which allows whiteboard images taken using a camera phone to be uploaded into the system
  - Both feed into the same web application processing pipeline
- All whiteboard captures can be browsed at a later date using a web interface
- The user can choose to have a captured image emailed to others present at an ongoing meeting
  - Meeting participants' email addresses are discovered by querying the user's Google Calendar

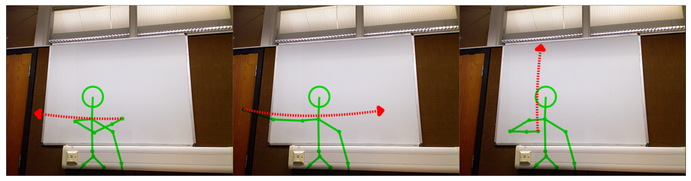
Gestures

Gesture detection is implemented using the Microsoft Kinect SDK which provides programmatic access to skeletal tracking data about users in front of the Kinect.

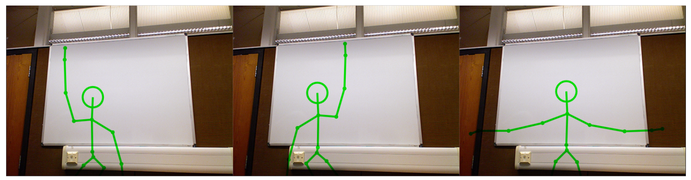
Two different classes of gesture were considered, with three specific gestures implemented from each class which correspond to possible actions the application can perform:

1. Capture whiteboard and upload to web application
2. As above, but also email image to the user
3. As above, but also email image to any meeting participants

Swipes: a capture is initiated if the hand is quickly swiped through the air in a specific direction, e.g. left to right, right to left, or straight up.

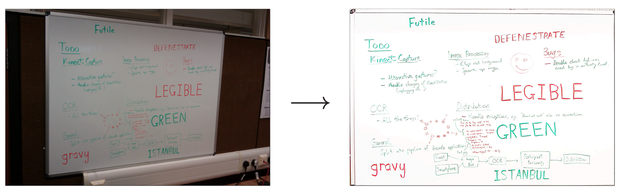


Postures: a capture is initiated when the arms are held in a specific pose for 4 seconds, e.g. left hand raised, right hand raised, or both hands outstretched.



Web Application

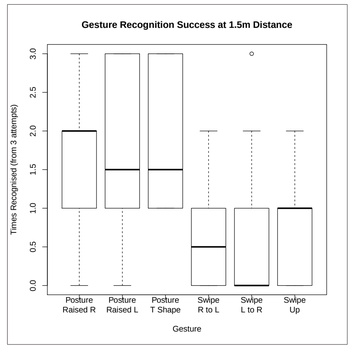
Once uploaded to the web application, the image is cropped, perspective corrected and enhanced using a combination of OpenCV and The GIMP.

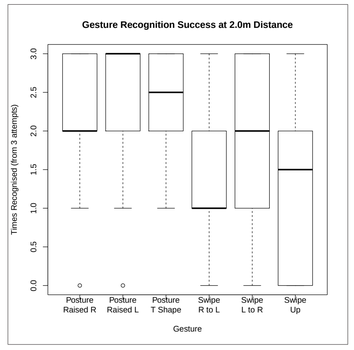


Finally, the image is distributed via email if requested and then made available to view using the web interface alongside the other captures.

Evaluation

Ten participants took part in a study to test the system and evaluate the suitability of the gestures chosen. Each of the six gestures was performed three times with the Kinect at a distance of 1.5m from the user and another three times at 2.0m.





Results showed that the postures were more reliably recognised than the swipe gestures from either distance.

It was also discovered that recognition was more successful when the Kinect was further from the user for both swipes and postures. This can be attributed to 1.5m being too near the Kinect's practical operating range where its tracking capabilities become unreliable.

The participants were recorded drawing at a whiteboard to establish whether the gestures were susceptible to accidental performance during normal whiteboard use. In addition, the system was set up to observe an hour-long tutorial where students answered questions on the whiteboard. In total, 135 minutes of whiteboard use took place while the system was active.

- The swipe gestures collectively exhibited a false positive rate of 43.1 per hour - well above what would be considered acceptable
- The postures did not experience a single false positive during the same period

Conclusions

- Swipe gestures overlap too closely with typical movements made when writing on a whiteboard, making them unsuitable for this application
- Postures have a reasonable recognition rate (limited only by the performance of the Kinect) and very low false positive rate, making them suitable as gestures to initiate a whiteboard capture
- There is a tradeoff between reliable gesture recognition and capturing the whiteboard in high enough resolution
  - The Kinect must be up close (<1.5m) for the captured image to be useful when viewing it in the future
  - The Kinect must be far enough back (2m+) for the user to be tracked reliably enough that gestures can be detected with a high success rate