Fabric Manipulation: An Eye Tracking Experiment

Peter Gibbons, Phil Culverhouse, and Guido Bugmann

***Abstract*— An eye tracking study of people engaged in a cloth sorting task is reported. Both eye fixations and hand movements were tracked during the experiments. The eye and hand movements were categorized and present alone a time line. The observations are consistence with a model of task driven covert global processing of the scene followed by overt sequential processing of fixated areas of interest for grasping. A computational model is presented.**

I. Introduction

The idea of robots helping in the home has existed ever since the original concept of a robot was conceived many years ago. Whilst inventions such as the washing machine, dishwasher and vacuum cleaner have helped to reduce the time and effort spent doing household chores we still spend a significant proportion of our lives doing household jobs [1], [2]. A common theme between many of the most dreaded household chores [3] is the ability to grasp and manipulate highly deformable objects such as clothing.
If a robot is to become a useful household appliance [4]-[7] then it must be capable of completing challenging real world tasks such as cleaning, washing & ironing [8]. This premise also holds true for the service and manufacturing industry which still rely heavily on intensive human labour to perform fabric manipulation tasks. Our main research interest is the design of a domestic robot which can perform a practical real world fabric manipulation task such as sorting clothes.

There is relatively little research into robot fabric manipulation systems in comparison to robot handling for solid objects. Ono et al. [9] were among the first to investigate robot fabric manipulation using a robot hand and computer vision to perform a simple pick and place operation for fabrics on a stacked on a flat surface. More recently they have extended their work by introducing a 3D simulator to help teach the robot to how to pick up the fabric [10]. In a more dynamic industrial setup [11] texture information was used to segment a colour image in order to select a fur from a pile of furs. A morphological analysis of the fur is then performed to determine the pick-up location for the robot arm manipulator. Using a washed mass typically found in a domestic environment [12], [13] segmentation of a colour image is also used to isolate an item of clothing. The grasp location is then found by calculating the centre of segmented region. In the same paper the identification of subsequent grasping points are discovered enabling the task of unfolding the item of clothing to be performed. This is achieved by using visual information for the hemline of the clothing, in a hung-up state, found from the outline in the image and the shadow cast by the hemline. A predefined template for the type of clothing which contains a specific number for each different hemline is then used to help select the best grasping locations needed to re-orientate the item ready for positioning onto a flat surface. The process of identification is then repeated until the desired pose is achieved. If the hemline information is inconclusive then the lowest point on the hanging item is selected as the next grasp point.

Another method of estimating the pose of a hanging item of clothing is to use a simulation of a deformable model [14], [15]. By comparing the actual pose of the item against the simulation of the model the model can be adjust until it matches the actual pose. It can then be used to determine the best locations to grasp the item in order to perform an unfolding task. One constraint of this method is the amount of prior knowledge for each item of clothing in order to initially build and simulate the model. In order to improve the accuracy of the model additional information would be needed.

One common theme between all these approaches to fabric manipulation was the use of vision to obtain the necessary information needed to decide what to select and where to grasp it. So as part of an initial investigation into our design of a domestic robot we observed humans performing everyday fabric manipulation tasks in a domestic, service and industrial environment. From the observations it was unclear exactly what role vision played in the completion the tasks. Research which has studied the role of human vision in more detail when everyday tasks such as making a cup of tea [16] or preparing a sandwich [17] are being performed suggest that it is the requirements of the task which dictate the information needed by the vision system.

We decided to investigate the use of vision by using a head mounted eye tracking system to monitor human eye

Manuscript received May 15, 2008. This work was supported in part by the Engineering and Physical Sciences Research Council (EPSRC) doctorial training account.

Peter Gibbons is with the Department of Computing, Communications and Electronics, University of Plymouth, Plymouth, Devon, UK (peter.gibbons@plymouth.ac.uk).

Dr Phil Culverhouse is with the Department of Computing, Communications and Electronics, University of Plymouth, Plymouth, Devon, UK (phil.culverhouse@plymouth.ac.uk).

Dr Guido Bugmann is with the Department of Computing, Communications and Electronics, University of Plymouth, Plymouth, Devon, UK (guido.bugmann@plymouth.ac.uk).

movements whilst performing set of fabric manipulation tasks. The tasks involved sorting a pile of clothing into two separate piles by either their size or colour. This was chosen after the initial observations into human fabric manipulation in different environments highlighted sorting as a common task.

This relatively simple task of sorting clothes provides a complex search space containing objects which immediately change shape once handled. Observing in greater detail how humans perform the task we aim to gain useful clues as to what information may be being used by the visual system. While this may not necessarily lead to the design of a robot capable to completing a simple human task it does, however, provide a measure of progress for the development of the robotic system against a successful working system.

This paper details the method used to investigate human eye movements using an eye tracking system whilst a simple fabric manipulation task is being performed. The results and analysis from a set of experiments are presented detailing the findings as well as the limitations experienced with the experiments. The final sections of the paper discusses the implications of the finding from the experiments and also what conclusions can be draw from them for the design and implementation of a robotic system.

II. Experimental Method

A. Participants

Three people voluntarily took part in the experiments. Each participant was given an instruction sheet explaining the purpose of the research and the details of the experimental procedure. They were asked to sign a consent form giving their permission to participate in the experiments and also for the subsequent use of any media as part of the research. The eyesight of each participant was sufficient that they could complete each experiment unaided, i.e. without the use of glasses or contact lenses. A fourth volunteer was unable to take part in the experiments due to calibration issues. The length and angle of their eye lashes meant that part of the eye was obstructed preventing the eye tracker from obtaining a consistent reading.

B. Tasks & Instructions

The experiments focused on two short tasks associated with the household laundry process. The tasks required sorting items of clothing, commonly found in a household, into two separate baskets, as shown in Fig. 1.by either size or by colour. The clothing was initially positioned a random pile on a desk in front of the participant. Before the start of each task, instructions were communicated verbally to the participant. For example, the participant was asked “Please could you sort the pile of clothing by placing the smaller items into one basket and larger items into the other basket”.

Each task is performed whilst sitting at a desk and required a head mounted eye tracking system to be worn. The setup is not identical to that found in a domestic situation where the process would usually be performed standing up, and typically involve a linen basket, the floor or a low surface, such as a bed, to sort items of clothing on. However it was decided that using a seated position at a desk would not affect the validity of the research and would enable the collection of better data. There were several reasons that contributed to the decision to use a seated position for conducting the experiments. The head mounted eye tracking equipment had a cable connecting in to the main processing unit restricting the range of movement. This reduction in the range of heads movements would help to minimise blurring in images captured by the scene camera and also reduce the risk of the head mount moving which could result in the loss of calibration. Finally, the eye tracking system was unable to capture the full range of eye movements. By using a seated position at a desk it was possible to reduce the range of both eye movements and keep them within the limits of the eye tracking system.

C. Equipment

The experiments used a head mounted Applied Science Laboratories 4000SL eye tracker system to capture the eye movements and produce the point-of-gaze information. The eye information is detected by using an infrared light source attached to the head mount which emits a beam of light off a two way mirror into the eye. The infrared light is reflected

Fig. 1. Experimental setup.

back off eye bounced off the two way mirror to sensor and transmitted along a cable to the eye tracking system. A scene camera attached to the head mount captured the participants view through the reflection on the opposite side of the two way mirror which the participants look through.

The point-of-gaze is created by using a method of calculating the offset between the pupil centre and corneal reflection at a frequency of 50 Hz. The lower sampling rate means the setup was suited to capturing fixation based information rather than quicker saccade based information. The system has at minimum of 40° horizontal and a 30° vertical field of view, a resolution of 0.5° and an accuracy of < 1° rising to < 2° at the periphery. The scene camera has a field of view of 50° and produced a black and white PAL image with a resolution of 768 x 576. Once the point-of-gaze was calculated it was superimposed onto the scene camera image as a set of white cross hairs, as shown in Fig. 2. The image was then captured by computer as a movie file. Due to problems recording the movie file the output image from the eye tracking system could only be captured 25Hz with a resolution of 240x180. This resulted in two sets of point-of-gaze cross hairs being superimposed onto a single image, as shown in Fig. 2. This still enabled fixations to be classified.

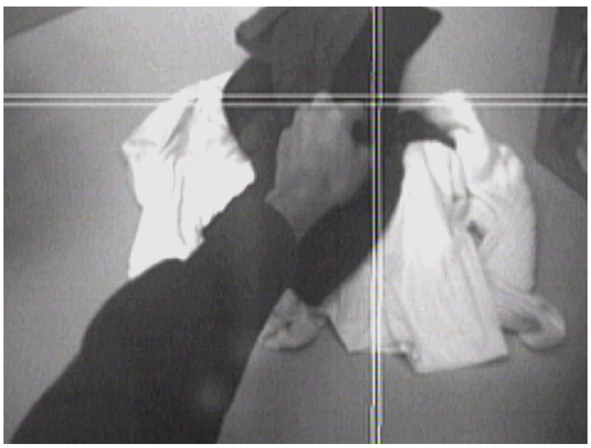

Fig. 2. Sample image captured by the eye tracker.

D. Calibration

Before a participant could perform an experiment they need to be calibrated to the eye tracking system. This involves several stages. Initially the head mounted system was secured onto the head. The position of the infrared camera and the angle of the two way mirror was adjusted to ensure a clear reflection could be captured from the eye and for the required range of eye movements. The scene camera was adjusted so that a nine point calibration sheet filled the image. The calibration sheet consisted of a large sheet of paper that covered the workspace and contained a dot at each corner, at the centre point along the edges and one in the centre. The dots were used as fixation points for the eyes during the calibration. To calibrate the participant's were require to kept their head perfectly still by biting onto a bit bar. They were then asked to fixate their eyes at each dot on the calibration sheet whilst the eye tracking system recording the position of their eyes. The accuracy of the calibration was checked by asking the participant to read the washing instructions on the label in an item of clothing. The accuracy was checked before and after each experiment was conducted to check it was correct.

TABLE I
OVERVIEW OF TYPES OF FIXATION USED TO ANALYSE DATA

Type of Fixation	Details of Fixation
Grasp point	Fixation at a subsequent grasp point
Local	Fixation close to previous fixation
Non-local	Fixation not close to previous fixation
Manipulated item	Fixation on item during manipulation
Uncovered area	Fixation at location uncovered by the manipulation of an item.
Sorting basket	Fixation at one of the sorting baskets
Out of range	Fixation outside the limit of the eye tracker

III. Data Analysis

The data analysis was performed manually using still images captured by the head mounted eye scene camera. Where clarification was required or data was missing, as seen in Fig. 2 where an arm is out of view, the image from the static scene camera was used.

Each image represented a duration of 40 milliseconds. These were grouped together into events associated to an individual item of fabric. The events were split into either eye movements or physical grasping of an item. These movement were further divided into a list of categorises through an iterative method of grounding the observations, reflecting on any anomalies then revising and re-labelling the categories until the results became self consistent. This process involved three researches and was mindful of the existing literature.

The physical grasps were categorised into either left or right hand grasps. These are displayed in Fig. 3 as blue and green to represent the smaller items and in light blue and light green to represent the larger items.

The eye movements were split into seven main categories, displayed in Table 1. The out of range category represented instances where the point-of-gaze data was not captured by the system because either the eye movements were outside the field of view or when the eyes blinked. The eye movement information for the first 6 items of clothing are not used for the eye analysis because were in full view of the participants before the experiment began.

IV. Results

A. Sorting Behavior

The sorting of items predominantly involved using both hands in alternate actions. Smaller items are initially picked up with one hand then collected together in the other hand before being dropped into the basket. For instance, items 7, 8 & 9 in Fig. 3 are collected together and dropped at the same time. The larger items are sorted individually by gathering into a single hand before being placed into the basket, as illustrated by items 11, 13 and 18 in Fig. 3. By gathering the smaller items together before placing them into the basket there is a 42% reduction in the number of movements to the baskets required to complete the overall task.

In the majority of instances participants fixated directly at the grasp point just prior to grasping the item. Item 16 was technically not fixated; the participant did fixate an area near to fixation point, but on a neighbouring item.

The fixations at the grasp point are often the first instance

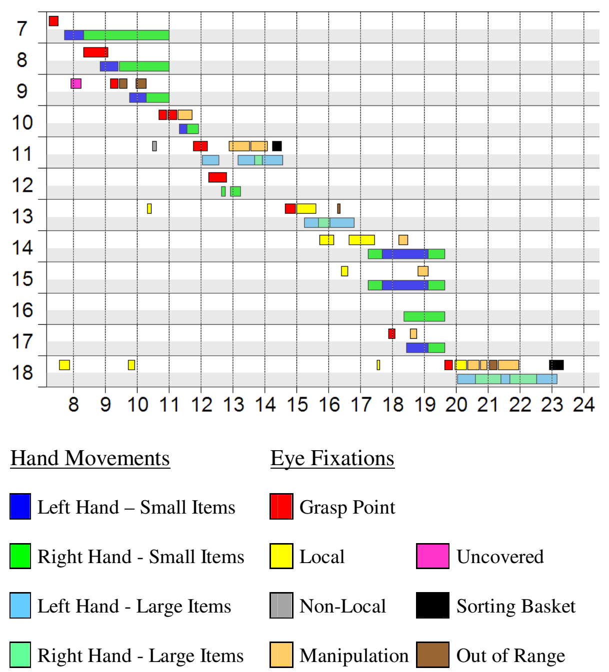

Fig. 3. Bar chart showing a time line analysis for sorting clothing items by size. The X-axis spans 17 seconds in 1 second intervals. The Y-axis represents events grouped according to which item of clothing is being looked at or manipulated. The data covers 12 items from item number 7 to item number 18. **Note –** fixations (shown in the unshaded horizontal bands) normally occur just prior to grasps (shown in the shaded horizontal bands). Fixations are displayed with the grasps and are both labelled by item.

that an item has been fixated on. Earlier fixations at items, such as items 11, 13 & 18 in Fig. 3, could be attributed to the item appearing in the line of gaze due to uncovering after moving an item above or drifts in the gaze off the edge of a neighbouring item.

Several fixations were located on the item whilst it was being manipulated. These were usually associated with large items, such as item 11 and item 18 as shown in Fig. 3. Some fixations although categorised as local were possibly a stare or the participant was waiting for new visual information, as there was no obvious visual task being conducted, either following an action or planning one, while larger items/hand were passing in the field of view.

Fixations on the sorting baskets occurred periodically prior to placement of large items, for example item 11 and 18, or a collection of smaller items.

Occasionally fixations were directed onto items which had been subsequently uncovered by the prior manipulation of an item above it. Item 9 in Fig. 3 shows an example of this type of fixation.

B. *Conceptual Model*

What stands out is the fact that the eye saccades from an item to be grasped to the next item to be grasped. There is no apparent scan of the scene. This suggests that global image processing takes place during fixations. On the other hand, what is the point of the local fixations? The only explanation we can offer is that the control/planning of the grasp requires fixation, given, some special role to the foveal region during grasp control. This is encapsulated in the design of the conceptual model shown in Fig. 4. In this model participants conduct a global covert analysis of the scene, then select a grasping point to which they saccade to, plan the grasp details and execute the grasp. This can be executed ballistically so that during execution, the saccade to the next grasps point can be planned based on previous memorized grasping points or a new analysis of the scene while fixating the current grasping point.

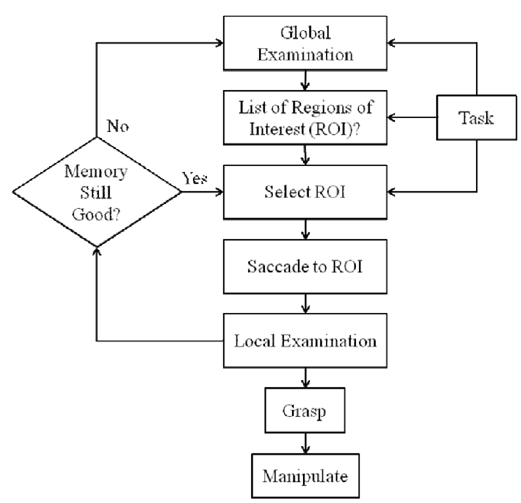

Fig. 4. Conceptual Model.

V. Discussion

The gathering of smaller items together before placing them into baskets shows that the brain is economizing on the number of movements required to accomplish the task.

From a robot design perspective this could significantly improve the energy efficiency of the robot as well as giving it human like characteristics. Having similar movement strategies could also help when cooperation between humans and robots is required in order to complete a task. Gathering similar items together also indicates that a strategy is being used to simplify the complexity of the task.

The simplification of the task indicated by physical hand movements combined with the eyes fixating direct at the required point just before the next grasp is required suggests that the visual information is task driven. This is similar to the finding of other eye tracking experiments, such as cup of tea [16] or preparing a sandwich [17] are being were they also suggest that eye fixations are task driven and occur in a just before an action happens.

VI. Conclusion

The analysis of eye movements led to the development of a general computational model for the selection of grasping points during a sorting task. A number of details of this model need to be developed or investigated further. In particular the global image processing for the selection of grasping regions of interest and also the processing to shape An interesting side effect of the implementation of such a model, which is the generation of robot eye movements quite similar to that of humans performing a similar task. For the observer this may make the robot “understandable”.

References

[1] C. Aliaga, “How is the time of women and men distributed in Europe?” *Statistics in Focus: Populations and Social Conditions*, (catalogue number: KS-NK-06-004-EN-N), Luxembourg: Eurostat. 2006, no. 2.

[2] J. Gershuny, D. Lader and S. Short, “The Time Use Survey, 2005”, Office for National Statistics (ONS), UK, 2006.

[3] CEA Publications, "5 Technologies to watch," Consumer Electronics Association, 2006.

[4] A. Gagalowicz, "Towards a vision system for a domestic robot," in *Proc. of the Int. Conf. on Systems, Man and Cybernetics, 'Systems engineering in the service of humans'*, Le Touquet, France, 1993, pp. 365-372.

[5] R. Bischoff, “Towards the development of ‘plug-and-play’ personal robots”, in *Proc. of 1st IEEE-RAS Int. Conf. on Humanoid Robot*, MIT, Cambridge, 2000.

[6] F. Pirri, I. Mentuccia, and S. Storri, "The domestic robot - A friendly cognitive system takes care of your home," In *Ambient Intelligence: Impact on Embedded System Design*, T. Basten, M. Geilen, and H. de Groot, Eds. Norwell, MA, USA: Kluwer Academic, 2003, pp. 131-158.

[7] F. Marrone and M. Strobel, "Cleaning assistant - A service robot designed for cleaning tasks," in *Proc. of IEEE/ASME Int. Conf. on Advanced Intelligent Mechatronics*, Como, Italy, 2001, pp. 1041-1046.

[8] J. S. Dai, P. M. Taylor, H. Lui, and H. Lin, "Folding algorithms and mechanisms synthesis for robotic ironing," in *Int. J. of Clothing Science and Technology,* vol. 16, no. 1-2, pp. 204-211, 2004.

[9] E. Ono, H. Ichijo, and N. Aisaka, "Flexible robotic hand for handling fabric pieces in garment manufacture," in *Int. J. of Clothing Science and Technology,* vol. 4, no. 5, pp. 16-23, 1992.

[10] E. Ono, K. Kitagaki and M. Kakikura, “Picking up a piece of fabric from layers with a hand with 3 fingers and a palm,” in *IEEE/JSR Int. Conf. on Intelligent Robots and Systems,* pp. 931-936, 2001.

[11] G. L. Foresti and F. A. Pellegrino, “Automatic visual recognition of deformable objects for grasping and manipulation,” *IEEE Trans. on Systems, Man, and Cybernetics – Part C: Applications and Reviews*, vol.34, no. 3, pp. 325-333, August 2004.

[12] K. Hamajima and M. Kakikura, "Planning strategy for task of unfolding clothes," in *Robotics and Autonomous Systems,* vol. 32, no. 2-3, pp. 145-152, 2000.

[13] M. Kaneko and M. Kakikura, "Planning strategy for putting away laundry - isolating and unfolding task," in *Proc. of the 4th IEEE Int. Symposium on Assembly and Task Planning*, Fukuoka, Japan, 2001, pp. 429-434.

[14] Y. Kita and N. Kita, "A model-driven method of estimating the state of clothes for manipulating it," in *Proc. of 6th IEEE Workshop on Applications of Computer Vision*, Orlando, USA, 2002, pp. 63-69.

[15] Y. Kita, F. Saito, and N. Kita, "A deformable model driven visual method for handling clothes," in *Proc. of the IEEE Int. Conf. on Robotics and Automation*, New Orleans, USA, 2004, pp. 3889-3895.

[16] M. F. Land, N. Mennie and J. Rusted, “The roles of vision and eye movements in the control of activities of daily living,” in *Perception* vol.28, 1999, pp. 1311–1328.

[17] M. F. Land and M. Hayhoe, “In what ways do eye movements contribute to everyday activities?” in *Vision Research*, vol. 41, no. 25-26, 2001, pp. 3559–3565.