Bilateral Behavior Media:
Visually Based Teleoperation Control with Accumulation and Support

Stephen PALM   (Ph.D. candidate)   Sato Lab, RCAST, The University of Tokyo

palm@lssl.rcast.u-tokyo.ac.jp   http://www.lssl.rcast.u-tokyo.ac.jp/~palm

Abstract:

Visually based control, accumulation and support systems are presented as an effective environment for teleoperation. We developed the Bilateral Behavior Media (BBM) paradigm and implemented systems which offer a status driven interface, collection of sampled control behavior, and operator assistance. The paradigm emphasizes a human-intuitive visual specification of task completion states without resorting to image understanding, modeling or extensive calibration. Visually based teleoperation has been effectively applied in a variety of microworld tasks where an operator's past experience in the macroworld is not applicable to the physics and scenario experienced in the microworld. Experiments of manipulating individual biological cells and microworld assembly have shown the success of the visually based BBM paradigm.

_____Conceptual Summary_____

1. INTRODUCTION

Teleoperated robots are indispensable in environments where humans cannot perform direct manipulation. In dangerous or distant locations or in the microworld, humans must manipulate the environment through a remotely controlled mechanism. Although teleoperation techniques have been extensively researched and developed, human operators still experience problems in accomplishing tasks when working through machines. Improving the teleoperating worker's situation is our underlying theme.

Recent work in visually based control methods has laid the foundation for advanced and robust master slave teleoperation. The visually based control methods offer 1) a more intuitive human machine interface and 2) allow for much simpler and economical control algorithms. [1, 2, 3, 4]

With many control techniques, appending sensors, especially visual sensors such as cameras, has been attempted in order to improve the human's understanding and control of the remote environment. However, we contend that the control method should be fundamentally based on sensing and in particular visual sensing in order to be effective in real world teleoperation applications.

We have developed the *bilateral behavior media* (BBM) paradigm based upon explicit visual communication between human operators and their teleoperated tools. The bilateral behavior media paradigm comprises three areas:

1. *Status Driven* control: A control methodology for visually specifying tasks and visually controlling machines.
2. *Behavior Sampling*: A data representation and extraction method for accumulation of visually based interactions between humans and tools.
3. *Status on Demand*: Functions for assisting and supporting humans through visual mechanisms.

Capabilities include visually navigated "redo" or "undo" based upon past visual-control sequences. Together the three areas encompass the notion of bilateral expression of behavior between humans and machines through a multiplicity of visual media.

One application of BBM techniques is cell handling in the biological world. Recent studies of aging and high fat diets have focused on analyzing Mato fluorescent granular perithelial (FGP) cells [5]. New techniques to analyze individual cells require the isolation of each cell by removing the tissue surrounding the cell. Figure 1 shows a visually based manipulator with a two micrometer wide scraper made of glass. The manipulator scrapes the undesired tissue from around the Mato FGP cell in preparation for its removal.
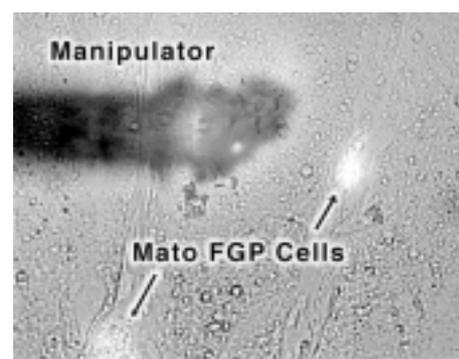

Figure 1. Cell manipulation environment.

We will first present and discuss the three main components of the BBM paradigm: status driven control, behavior media, and status on demand. This will be followed by a discussion of the implementation of the systems that perform status driven control and behavior sampling. Finally, we review the experiments performed to show the productivity of using BBM techniques.

2. THE BILATERAL BEHAVIOR MEDIA PARADIGM

2.1 Status Driven Control

Status driven is a third generation teleoperation technique where a slave manipulator is visually instructed by the master control panel. Furthering the first generation joint-angle control techniques and second generation coordinate transformation techniques, a status driven system recognizes the target status specified by the operator by extracting the task status from visual sensors.

The task environment is initially described via sensing points in the work environment and on the manipulated object. A sensing point describes a point of manipulation significance in the visual representation of the object. For example, sensing points would be used to describe the abutting surfaces in a pick and place task. The relationships between the sensing points and the change of those relationships describes the task at hand. In other words, control information is expressed in the spatial temporal relationship between corresponding sensing points associated with each object.

(a) Positioning by "Point"   (b) Place by "Center"

(c) Insert by "Arrow"

Sensing Points:
○ Target
◉ Environmental
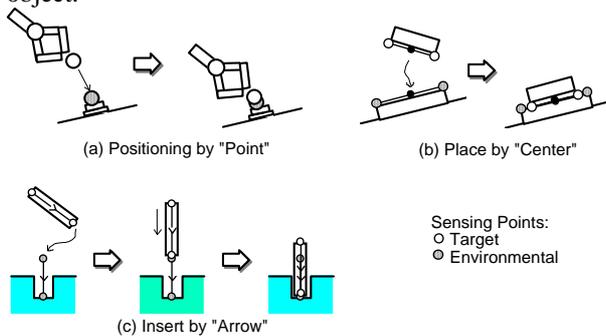
2.2 Behavior Sampling

Pure status driven control is only concerned with the immediate task of causing the sensing point to coincide to achieve task completion. Conversely, behavior sampling is concerned with the long term control aspect of visually aggregating and preserving the independent control events. Behavior sampling is:

1. A <u>data representation and storage mechanism</u> that can be used to display past control and visual sequences.

2. A means to <u>extract and accumulate the underlying raw data</u> (both object visual representation and control instructions). While some other robot control methods do accumulate past experience, they tend to be based purely on image frame data or abstract representations or models.

3. The ability to <u>syntactically organize the visual and control information</u> and allow for the addition of semantic information.

Thus behavior sampling provides the foundation for humans to review and reuse the visually based control information that was derived from status driven system which does not have an underlying concept of the objects it is manipulating.

The input to a behavior sampling system consists of the video image of the slave environment and the time stamped control information. The control information includes such items as the location of the objects in the environment and the type of control desired. The control information is typically specified by sensing points and the desired relationship of the final state of the sensing points. All of this information is processed and converted into the behavior sampling data representation. Further, if an operator wishes to input semantic information for a given node or link, the data representation is capable of annotating such information to the nodes and links.

The output of behavior sampling is an indexable, structured stream that contains both the visual and control information. The form of the stream is such that addressing of individual objects or information is readily obtainable without resorting to decoding all of the information in the stream or even in a large segment of the stream. The stream is suitable for storage (e.g., hard disk) or for transmission. Further, the output is parsable in such a manner that the form and style of the control performed on a given object in a past sequence is usable for control of a different object in a future situation. In other words, the behavior sampling output is suitable as the input to the status on demand functions.

Behavior Sampling is partially based on the concepts of hypermedia and syntactical or semiotic analysis. The syntactical or semiotic analysis method exploits the underlying structure of the real world scene in the representation. Syntactic methods extract structural information without understanding the meaning or semantics of the visual objects since the elements can be derived through low-level vision techniques. The structure of the visual and control information is extracted by observing *signs*. The first two columns of Table 1 show Gonzalez's proposed assignment of signs for the images and video domains [6]. The third column shows the control domain signs developed specifically for behavior sampling.

Table 1. Assignment of signs for various media domains.

	DOMAIN		
	Images (Spatial)	Video (Temporal)	Control (Visual)
Meta Sign	Picture	Episode	Completed Work / assembly
Signs	Objects	Scene	individual object task
SubSigns 1	Surfaces	Shot / Global Motion	sensing pairs
SubSigns 2	Lines	Objects / Local Motion	sensing points
SubSigns 3	Pixels	Stationary Change	

2.3 Status on Demand

The status on demand functionality is a visually based interface to the behavior sampled data in a status driven system. Through imagery, graphics, and text, the

status on demand system is able to display milestones in which (task) status has transitioned from one type of task to another. Thus, an operator is able to view the past sequence of events comprising tasks in an easy to comprehend and partition manner. The task status at each relevant point in time of the procedure is then available for reference and visual re-manipulation by the operator.

A status on demand system can have several levels of functionality. Lower level functions provide immediate short term support for the operator to modify a recent manipulation. Intermediate level functions including editing of manipulation parameters or annotation of semantic information. Higher level functions would allow the replay or reuse of previous manipulation procedures in new control situations.

An example low-level function is *redo*. Redo allows the operator to repeat the last style of change of state (perhaps on a different set of sensing points). This is useful with similar motions that need to be performed multiple times from (typically) different start and end points. For example, if there is a series of objects to be manipulated in a similar way, the operator would setup the sensing points for the first object and perform one or more manipulation tasks on it. For the subsequent objects, new sensing points would be used to specify the object(s) and the redo function would perform the same compound set of manipulations.

3. SYSTEMS ARCHITECTURE

The recording and display composer mechanisms of the first behavior sampling system are based upon the draft MPEG-4 framework [7]. MPEG-4 provides a toolbox of functions for video encoding such as specifying and encoding individual objects and specifying how the individual objects are composed to form a complete scene. MPEG-4 does not provide mechanisms for segregating or extracting objects from a video frame nor does it provide a mechanism for describing control relationships between objects.

Implementation of a behavior sampling system entailed developing two main components 1) an automated video segmentation method and 2) control information processing and storing. These are shown in the dashed boxed in Figure 2. The manipulator control section is similar in function to the SD-MHS control system. The object and scene encoding and decoding functions are part of the MPEG-4 framework.
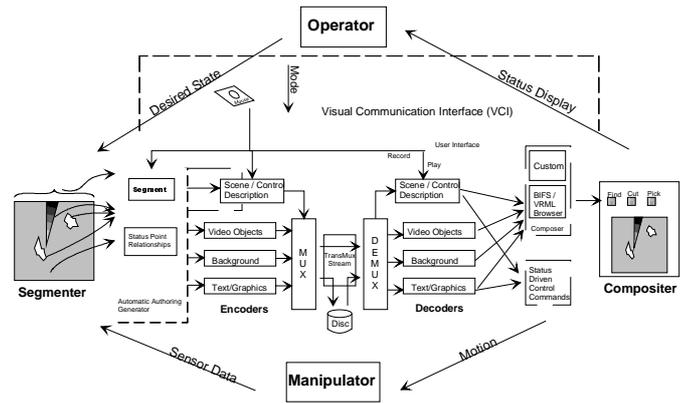

Figure 2. System Architecture

4. CONCLUSIONS

We have introduced the bilateral behavior media paradigm as an effective way of visually interacting for teleoperation. The status driven control method has been introduced and realized through the status driven micro handling system (SD-MHS). Behavior sampling allows the system to sample, structure, and store motion control sequences and their associated imagery. This behavior sampled data can be accessed to repeat or redo a recorded sequence. Experiments have shown the effectiveness of the approach in both automatic and shared control modes in microworld manipulation tasks.

References

[1] G. D. Hager, "A Modular System for Robust Positioning Using Feedback from Stereo Vision," *IEEE Trans. On Robotics and Automation*, Vol. 13, No. 4. pp. 582-595, August 1997.

[2] N. P. Papanikolopoulos, P. K. Khosla, and T. Kanade, "Visual Tracking of a Moving Target by a Camera Mounted on a Robot: a Combination of Vision and Control," *IEEE Trans. On Robotics and Automation*, Vol. 9, No. 1, pp. 14-35, February 1993.

[3] T. Shibata, Y. Matsumoto, and T. Kuwahara, "Hyper Scooter: a Mobile Robot Sharing Visual Information with a Human," *Proceedings of R&A 95*, Vol. 1, pp 1074-1079, 1995.

[4] T. Sekimoto, T. Tsubouchi, S. Yuta. "A Simple Driving Device for a Vehicle - Implementation and Evaluation," *Proceedings of IROS 97*, Vol. 1, pp. 147-154, 1997.

[5] M. Mato et al, "Involvement of Specific Macrophage-lineage Cells Surrounding Arterioles in Barrier and Scavenger Function in Brain Cortex," *Proc. Natl. Acad. Sci. USA*, Vol. 93, pp. 3269-3274, April 1996.

[6] R. Gonzalez, "Hypermedia Data Modeling, Coding, and Semiotics," *Proc. of the IEEE*, Vol. 85, No. 7, pp. 1111-1140, July 1997.

[7] R. Koenen, "Overview of the MPEG-4 Standard," ISO/IEC JTC1/SC29/WG11 N1730, Stockholm, July 1997.