METHODS **21,** 385–391 (2000)
doi:10.1006/meth.2000.1025, available online at http://www.idealibrary.com on IDEAL®

Applications of Video Mixing and Digital Overlay to Neuroethology

Mitra J. Hartmann,[1] Christopher Assad,* Brian Rasnow,†
and James M. Bower

Department of Biology, 216-76, California Institute of Technology, Pasadena, California 91125;
**Jet Propulsion Laboratory, 303-300, Pasadena, California 91109; and †Amgen, Inc.,*
1 Amgen Center Drive, Thousand Oaks, California 91320

Neuroethological experiments often require video images of animal behavior and recordings of physiological data to be acquired simultaneously, synchronized with each other, stored, and analyzed together. The use of inexpensive multimedia computers offers new possibilities for mixing video images, analog voltages, and computer data, storing these combined signals to videotape, and extracting quantitative data for analysis. In this paper, we summarize methods for mixing images from multiple video cameras and a Macintosh computer display to facilitate manipulation of data generated during our neurophysiological and behavioral research. These technologies enhance accuracy, speed, and flexibility during experiments, and facilitate selecting and extracting quantitative data from the videotape for further analysis. Three applications are presented: (A) we used an analog video mixer to synchronize neurophysiological recordings with ongoing behaviors of freely moving rats; (B) we used a chroma keyed digital overlay to generate positional data for the rat's face during drinking behavior; and (C) we combined a computer model of a rat's head and whiskers with videos of exploratory behaviors to better track and quantify movements in three dimensions. Although the applications described here are specific to our neuroethological work, these methods will be useful to anyone wishing to combine the signals from multiple video sources into a single image or to extract series of positional or movement data from video frames without frame grabbing.

Key Words: animal tracking; data synchronization; video mixer; behavioral monitoring; whisking.

Recent improvements in video technology and in the interfaces between video and computer systems can greatly enhance recording methods in experimental science. In this paper we describe methods for analog video mixing and for digital overlay on multimedia computers. These methods can be used in a variety of experimental applications, including (A) synchronization of physiological with video data (B) accurate spatial localization of features of the experimental setup (e.g., recording and stimulating electrodes), and (C) tracking movements in three dimensions by synchronizing views from different cameras. These technologies enhance accuracy, speed, and flexibility during experiments because they can provide an information-rich videotape of the subject and state of the experimental apparatus. Audio channels on the videotape can also be used to store analog voltages (e.g., physiological data) or experimenter's comments. After the experiment, these technologies facilitate selecting and extracting quantitative data from the videotape for further analysis. We begin by summarizing the methods involved, and then describe three applications in which these techniques have aided our neurophysiological research.

---

[1] To whom correspondence should be addressed.

1046-2023/00 $35.00

METHODS

1. Video Mixing

There are many instances in experimental research when it is desirable to mix the synchronized outputs of multiple video cameras. In Rasnow *et al.* (1), we presented a circuit for mixing the outputs of two video cameras (or other sources) into a single video image, which can then be recorded on standard videotape for later review and analysis. This simple, inexpensive circuit works equally well with PAL and NTSC video, and also provides an output for running a field counter useful for data synchronization and time coding. The only requirement is that the two video cameras be synchronized, or "genlocked," with each other, and therefore at least one of the cameras must have an external sync input. Professional video cameras generally have this input, as do some inexpensive cameras used for security systems. Security-system cameras are also available with extended infrared and low-light sensitivity, making them ideal for low- to no-light recording. Genlocked video signals from more than two sources can also be easily mixed by daisy-chaining multiple video mixers together (e.g., see Fig. 1).

The mixer, built from readily-available electronic components, is based on a commercial video chip that multiplexes signals from two video sources. Timing signals to switch between the two inputs are provided by monostable multivibrators that have potentiometer controlled delay times following horizontal and vertical sync pulses. This allows the user to select which portion of each camera's field of view will appear in the mixed image. Note that with this system, each partial view still appears at full resolution. This differs from the standard "picture-in-picture" convention, in which one complete image appears in a small window at reduced resolution. Finally, an absolute time code for each video field is easily generated by triggering a counter off the sync output signal. A circuit diagram and printed circuit board layout for the mixer can be found in Rasnow *et al.* (1).

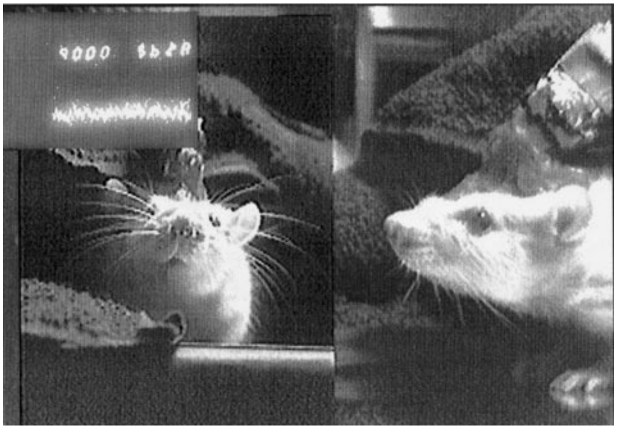


FIG. 1. Two video mixers in series were used to combine images from three cameras. Two orthogonal camera views of the rat were mixed, and then combined with the signal from a third camera pointing at an oscilloscope displaying the neural data. Poor contrast and resolution can be attributed to the low light levels necessary to establish a natural environment for the animal.

2. Chroma Keying: Viewing Video within Any Program on a Multimedia Computer

Although the video mixer allows us to compactly synchronize and store analog video signals, it is also often necessary to combine the video data with computer graphics or other digital data. In this case, multimedia computers are commonly used to "frame grab," i.e., to digitize video frames and store them to disk. However, frame grabbing is time consuming and places high demands on many of the computer's subsystems. Since each full-resolution video frame contains more than 300,000 pixels, digitally storing uncompressed video requires approximately 10 million bytes per second to be transferred to disk. Multimedia computers are thus often seriously limited in their frame capture rate, resolution, and storage capacity, and usually record video at lower quality than inexpensive VCRs. Newer digital camcorders and digital video tape recorders incorporate high-resolution frame grabbers, but the data stream must still be transferred to a computer for image processing and analysis, invoking a similar communication bottleneck. Furthermore, the resolution of digital recorders is often inferior to that of high-quality analog devices, and the images must be analyzed in software that is compatable with the image file format.

In the applications described below, instead of frame grabbing, we used chroma keying on Macintosh AV multimedia computers to overlay digital graphics on video images. Since these computers support live video, the experimental video records can be viewed, annotated, and analyzed within virtually any commercial software package. Furthermore, the screen can be recorded back to videotape, allowing the graphics overlays, visible video, and other program windows to be stored together. By exploiting the inherent high bandwidth and low cost of video without the digital storage overhead, this method can greatly enhance speed and efficiency both during the experiment and afterward, during data analysis.

To synchronously display video and computer inputs, a multimedia computer must combine two source images—the computer's normal display and the live video input—into one output image. Usually this is accomplished by spatially segregating the video and computer images: video is displayed in a rectangular window, and the computer image is displayed unaltered everywhere else. Chroma keying is an alternative for mixing the source images together on a pixel-by-pixel basis. Imagine the video and computer images displayed in two parallel planes, with the video image behind the computer image. Now imagine that all pixels of the computer image that are a specific "key" color are replaced with the underlying video pixel; i.e., the key color becomes transparent to the underlying video. The resulting output image can contain multiple and arbitrarily shaped regions of video and computer images.

Chroma keying video with the graphical output of computer programs provides many opportunities for novel data display and analysis. For example, it is straightforward to annotate or trace video images in any graphics or drawing program simply by displaying the video window underneath it and then drawing on the video image. Pixels in nonkey colors then appear to float over the underlying video (see Fig. 2A). The digital graphics can then be saved to a file containing only the drawings and annotations, and not the underlying video. This is much faster and more efficient than the more conventional method for tracing a video image: frame grabbing the video to a file; loading the image file into a graphics program; tracing the image; and finally, deleting the frame-grabbed image and its disk file. Also, the computer program's function and operation are unaffected, regardless of whether video is visible within its windows.

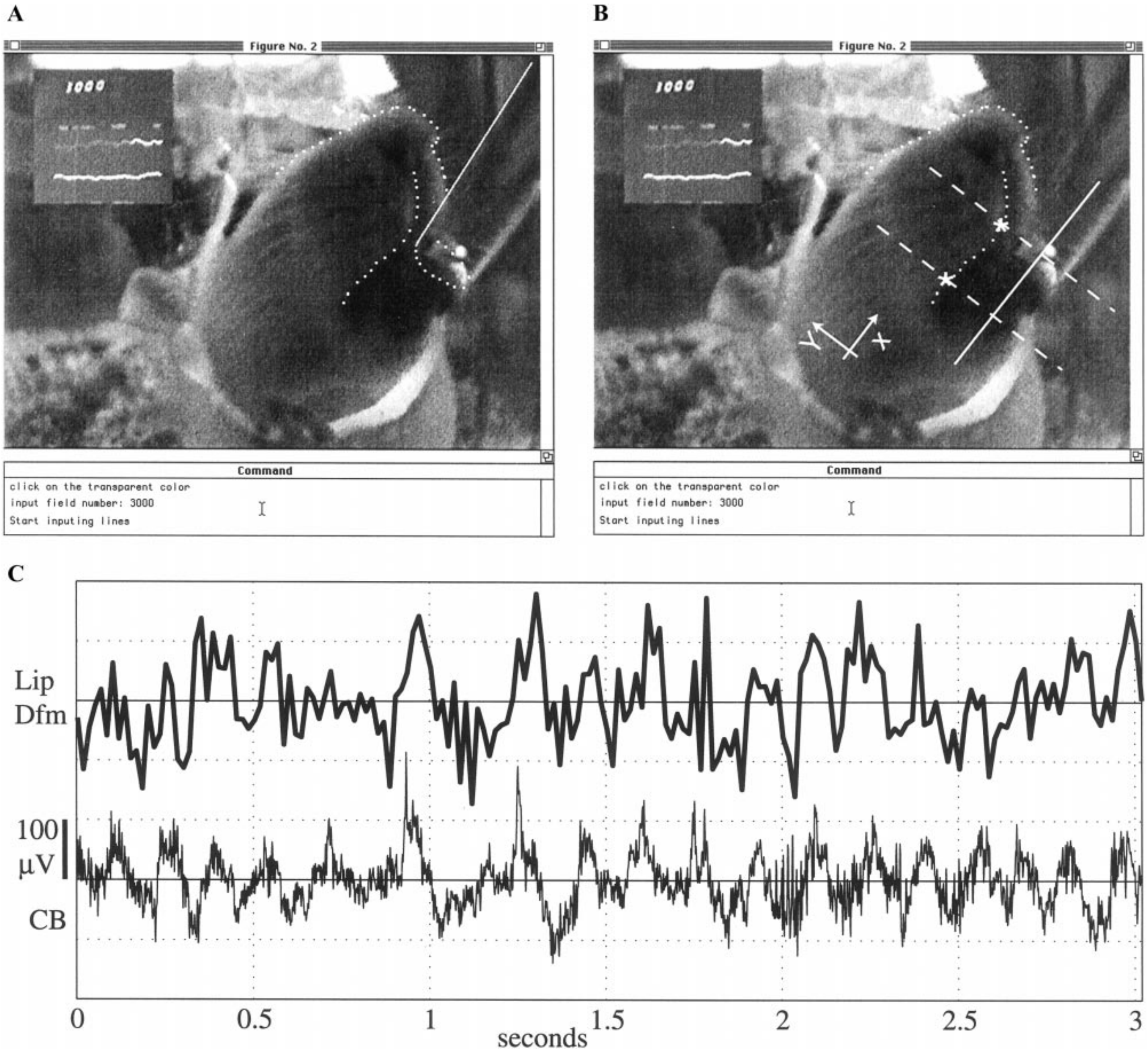
We implemented chroma keying on the Macintosh 6100AV, but the method will work equally well on the 660AV, 840AV, 7100AV, or 8100AV models. Our method for chroma keying takes advantage of a peculiarity of these computers' screen buffer implementation. With the computer displaying 256 colors, one color is always "key" or 100% transparent to video. We find this color by displaying the palette of all 256 colors over the video window, and looking for a missing color filled with the background video and surrounded by the other colors. The desired parts of the application window can then be filled with the key color to view the underlying video. These methods work even with programs that are not explicitly designed to support video. For example, using MATLAB 4.2 (The Mathworks Inc., Natick, MA), we select the key color by drawing a 256-color checkerboard in a figure window positioned over the video, and then choose the color that appears transparent (see Ref. 1 for details and a MATLAB script to implement chroma keying). We have focused on these Macintosh computers because they are inexpensive and support the methods described here with stock hardware and require no programming. They also have video outputs in addition to their video inputs, so the computer screen (including the video input displayed on it) can be recorded on any VCR. Note that the details of these methods are unlikely to work without modifications on other computer models; however, in principle any computer capable of chroma keying could be used in this manner.

APPLICATIONS

1. Video Mixing for Synchronization of Behavioral and Neural Data

We used the analog video mixing circuit to synchronize neurophysiological recordings from the rat cerebellum with the animal's ongoing behaviors. Our laboratory is particularly interested in the portions of the rat cerebellum that respond to tactile stimulation of the lips and whiskers (2), since rats preferentially use these perioral regions for exploring the environment (3–5). During exploration rats often move their large whiskers rhythmically at 6–12 Hz, in a behavior known as "whisking" (4, 5). We have proposed that the cerebellar hemispheres are specifically involved in the acquisition of sensory information and may be particularly important during exploratory behaviors (2). To test this hypothesis, we needed to accurately locate the rat's head, lip, and whisker positions in behavioral videos, and synchronize movements of these structures with neural signals recorded from the rat's cerebellum.

As shown in Fig. 1, we used the video mixer to combine images from multiple video cameras: two commercial camcorders were used to monitor the rat's behavioral activity, and a black-and-white surveillance camera in another room monitored an oscilloscope displaying the neural data. Since the video mixers permit variable window sizes and positions, we placed the oscilloscope image in a region of the picture where the animal did not explore. Two LED counters were mounted above the oscilloscope display: one counted fields from the video mixer, and the second counted pulses sent from the data acquisition system indicating the name of the open data file. Because this method continuously superimposed behavioral and physiological data, it greatly aided scanning through long periods of video looking for interesting behavioral and/or neural events.



```
click on the transparent color
input field number: 3000
Start inputing lines
```

```
click on the transparent color
input field number: 3000
Start inputing lines
```

FIG. 2. Chroma keyed digital overlay for feature extraction. (A) A MATLAB figure window is superimposed over a live-video window of a rat drinking. Positions of the water tube (solid line and dots) and of the lip and head (dotted lines) were traced using graphics functions in MATLAB. Because these digitized points could be saved without the underlying video, storage requirements were minimal. (B) Method for measuring lip deformation. The *x* axis was defined as a line tangent to the average lip position (solid line). Perpendicular deformation was then measured along *y*. (C) Neural signals recorded from the rat cerebellum, in relation to tactile input to the lip. Lip Dfm, lip deformation; CB, cerebellar recording.

2. Chroma Keyed Digital Overlay for Feature Extraction

The video mixer provides an elegant way to synchronize multiple video images, and physiological data with behavioral data, but it is still necessary to extract salient features from the video. In principle, one can use frame grabbing for this extraction, but as discussed above, frame grabbing is extremely inefficient. In contrast, chroma keying allows us to digitize and save only those portions of each video field that are most important to the data analysis. We used this technique to correlate neural activity with the position of the rat's upper lip during drinking behavior.

We have previously shown that the signals seen in the rat cerebellum during drinking behavior exhibit a rhythmicity between six and seven cycles per second that is correlated with the regularity of licking (6, 7). Because neurons in the rat cerebellar hemispheres are known to respond strongly to tactile stimulation of the lips and whiskers, we suspected that responses during drinking might be correlated with tiny deformations or deflections of the upper lip. We therefore needed to examine the behavioral variation in individual licks and correlate this with variations in the neural signal. We have used the digital overlay technique to enter the precise position of the rat's upper lip directly into MATLAB for quantitative analysis.

Figure 2A shows one frame of typical video overlay used to analyze rat drinking behavior. A MATLAB figure window is shown superimposed over (and filled with the key color, making it transparent to) a live-video window. At the beginning of analysis we traced the position of the water tube (in a nonkey color, solid outline). These digitized points were saved and plotted in every field (half a frame), so that we could be certain that the picture as a whole had not drifted between fields. Then, using the VCR pause mode, we moved field-by-field through the video, digitizing in MATLAB only the precise position and shape of the upper lip, and a reference outline, such as the nose or the eye (dotted lines). Our method for measuring lip deformation is illustrated in Fig. 2B. We determined a line tangent to the average lip position, which became our x axis (solid line). We then chose two lines, perpendicular to the x axis, that intersected distant positions along the lip (dotted lines). For each field we then measured lip deformation as the y distance between these x values on the lip (asterisks).

Figure 2C shows the neural signals from the rat cerebellum in relation to deformation of the upper lip. It is clear that both the temporal and amplitude variation in the neural signal correlate well with the tactile input resulting from the lip deformation. Only with the spatial resolution provided by field-by-field tracing were we able to determine the precise behavioral correlate to the neural data. Although we could have digitized each field in a frame grabbing application, and then imported the image files in to MATLAB to quantify lip position, this would have been tremendously resource and time consuming (approximately 3–4 min per frame). At this speed, digitization of the 182 video fields (approximately 3 s of video) would have taken more than 12 h. Using chroma keying techniques, digitization of all fields took less than 2 h, even with the precise spatial resolution required.

3. Reconstructing Three-Dimensional Head Orientation and Whisker Trajectories

In the previous application, tracing deformations of the upper lip during drinking behavior worked well because the rat's head remained relatively still. In contrast, during exploratory behaviors the rat moves both its head and whiskers simultaneously, and tracking the movements of the thin whiskers proved more difficult. This was especially apparent during whisking behaviors when the whiskers swept back and forth at close to 8 Hz, allowing only three or four video frames per whisk cycle. In addition, changes in head angle and whisker reflections made it extremely difficult to match any individual whisker in one frame with its corresponding image in the next frame.

To better study whisking behaviors, we have begun to use a computer model of the rat head and whiskers to help analyze head and whisker movements from the video. In previous studies, we had successfully used this method to analyze movements of electric fish during exploratory behaviors (1, 8). In the present study, the computer model consisted of a simple three-dimensional wire-frame mesh of the rat head. The whiskers were represented as curved lines anchored at fixed locations on the model rat's face, corresponding to the mystacial pad for the macrovibrissae (Fig. 3 shows one such whisker; see also Ref. 9). The model was constructed using measurements of the animal's head and whiskers. Aligning this model directly to the video image proved faster and more accurate than trying to track fiducial points on the rat's head.

As in the first application, we recorded the rat's behavior with two video cameras. One camera in front of the rat monitored movements in the *x–z* plane, and the other, placed alongside, monitored movements in the orthogonal *y–z* plane (Fig. 3). Coordinate system axes were first overlaid on the video using MATLAB's graphics functions. For each view, a three-dimensional coordinate axis was aligned to the actual viewing axis using an orthographic projection. The videotape was then examined for interesting sequences of whisker positions and head orientations.

The three-dimensional model of the rat head and whiskers was rendered with nonkey colors in these two coordinate systems overlaying the video. With graphical controls, we adjusted parameters to translate and rotate the model head until it was in register with the video images of the rat head (Fig. 3). The whiskers were then aligned to the video image with a control that varied the sweep angle. Our preliminary results in analyzing the actual behavior indicate that this method is more accurate than simple tracing. By overlaying the computer model on the video, we can (A) use the model to constrain the allowable whisker motions between frames; (B) fill in whiskers that are not visible in individual frames; and (C) interpolate the three-dimensional whisker positions between frames. These all lead to better estimates of the three-dimensional whisker position throughout the behavior. We hope that modeling the movements of the entire whisker field during exploratory behaviors will help elucidate the basic capabilities and neural computational principles underlying this important sensory modality.

CONCLUSIONS

The video mixing method presented here can be used to synchronize both analog and digital images with high temporal resolution, to correlate different types of data, and to efficiently track movements in three dimensions. There is no need to waste screen space with a common timing signal, nor to correlate two video tapes after the experiment. The importance of simultaneously displaying behavioral and neural data cannot be overstated, since neuroethological analysis frequently involves scanning through many hours of videotape to look for short behavioral or neural events of interest.

Chroma keying can be used to easily and accurately extract quantitative data from videotape directly into a variety of unmodified commercial programs. When combined with computer models, the

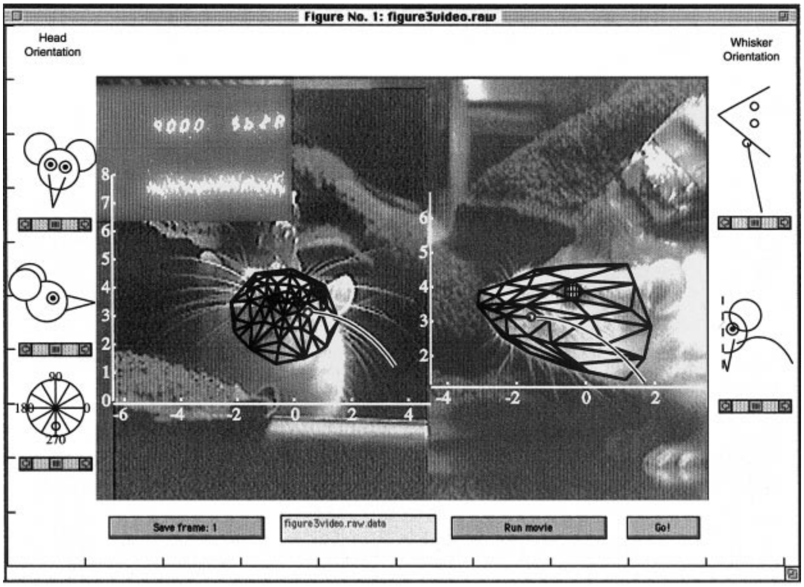


FIG. 3. A three-dimensional wire-frame model of the rat's head and one whisker is superimposed on the video of Fig. 1. The model head can be translated and rotated with the controls on the left to coincide with the rat's head in the two video views. Adjusting the fit of the model in either view automatically makes corresponding adjustments to the other view. The whisker, which is anchored to the head, can be independently rotated in the whisking plane and bent using the controls on the right. Multiple whiskers can be attached to appropriate locations on the face and rotated to angles that match the video. Mesh lines are thick to stand out above the black and white image of the rat; they can be significantly thinner when they are color-coded.

chroma keyed video overlay can be used to efficiently track and analyze movements in three dimensions. Although in the method described here the model must still be manually aligned to the video, alternative methods using automatic tracking systems are much more resource intensive and can reach the same accuracy only at much higher cost, if at all.

Mixing multiple video images, and computer graphics with video images, has greatly enhanced our experimental capabilities. Because detailed experimental records are automatically captured on videotape, we can work faster and more flexibly during experiments, resulting in more data and less stress to the animal. For data analysis after the experiment, these techniques allow us to work much more efficiently. Furthermore, the mixed videotape images are invaluable in resolving and verifying any data inconsistencies or omissions in the written notes. Within the context of the neurobiological study of behavior, these techniques allow us to begin to match the high spatial and temporal resolution of neurophysiological data with the fine details of animal behavior. Although we have described the use of multimedia and video mixer technologies in this context of neuroethology, many other experimental fields could benefit from these methods.

APPENDIX: EQUIPMENT

1. Macintosh 660AV or 840AV, or Power Macintosh 6100AV, 7100AV, or 8100AV: These inexpensive, older models are currently available at many used computer stores or by mail order.

2. Video Mixer: The circuit is described by Rasnow *et al.* (1) along with a custom-printed circuit board layout. Parts total less than U.S. $50. The circuit is designed around the MAX453, available from Maxim Integrated Circuits, http://www.maxim-ic.com.

3. MATLAB is available from the MathWorks, Natick, MA, http://www.mathworks.com.

4. Any video camera with an external sync input will work in the applications described in this paper. We used Sanyo CCD Model VDC264 black and white security cameras, which cost less than U.S. $350.

REFERENCES

1. Rasnow, B., Assad, C., Hartmann, M. J., and Bower, J. M. (1997) *J. Neurosci. Methods* **76,** 83–91.
2. Bower, J. M. (1997) *Prog. Brain Res.* **114,** 463–496.
3. Vincent, S. B. (1912) *Behav. Monog.* **1,** 7–85.
4. Welker, W. I. (1964) *Behavior* **22,** 223–244.
5. Carvell and Simons (1990) *J. Neurosci.* **10,** 2638–2648.
6. Hartmann, M. J., and Bower, J. M. (1995) *Soc. Neurosci. Abstr.* **21,** 1910.
7. Hartmann, M. J., and Bower, J. M. (1996) *Soc. Neurosci. Abstr.* **22,** 1200.
8. Assad, C. (1997) Ph.D. thesis, California Institute of Technology.
9. Brecht, M., Preilowski, B., and Merzenich, M. M. (1997) *Behav. Brain Res.* **84,** 81–97.