The Oculus Ranae

by

D. J. Collinge

and

S. M. Parkinson

School of Music, University of Victoria,
P.O.B. 1700, Victoria, B.C., CANADA, V8W 2Y2

UUCP: uw-bear@uvicctrtcollinge
BITNET: collinge@uvicctrt

Abstract

The Oculus Ranae interfaces a television camera to a
computer. It extracts movement information from a video signal
giving location and quantity of movement in real time. This is
done with one small circuit card and a computer interface. The
circuit uses no exotic components or circuitry and gives high
performance at low cost. The circuit was conceived for use in
interactive installations but has other musical interface
applications.
introduction

In our work with audience-sensitive sound installations we developed a need for an audience sensor with several special attributes. These installations are erected and left running for long periods of time, sometimes weeks or months; therefore, the sensors must be robust and reliable. Since the work is unattended for much of that time it must also be vulnerable to curiosity or maliciousness. The sensor must also be capable of detecting the position of more than one person within its range of sensitivity. Finally, it should work with minimal preparation of the installation space, since such preparation can be costly and inconvenient.

The Oculus Ramæ is the solution to these problems: it is a circuit that analyzes the output of a TV camera, detecting the portions of the image that are changing in brightness. The name is a reference to the visual system of the frog, which sees only moving objects. The circuit is small and reliable, may be remote from the TV camera for security, and does not require any special preparation of the environment.

While this circuit is well adapted to the task it was designed for, it has other applications; for example, as a dancer interface or a musical instrument interface.

Other Sensors

We investigated several other sensing techniques before choosing this one. In a previous work we had successfully used floor panels with arrays of pressure switches but these panels were extremely cumbersome and required many hours to install and test. They also required a protective covering, which made them very striking visually, misleading participants about the essentially non-visual nature of the installation. Light-beam systems have been used in these applications but they require numerous light sources and detectors to give position information and inherently can be confused by the presence of one moving object. Ultrasound and UHF systems are inauspicious and expensive, and can give good range and velocity information but they are also subject to confusion by multiple moving objects.

The Circuit

The Oculus Ramæ is composed of two circuits, the synchronization and digitization stage, and the analysis and interface stage.

Is the synchronization circuit the video signal is stabilized and sync is separated by a commercial circuit. A master oscillator is gated by the sync information and generates a pixel clock and line and pixel address. This address divides a line into 118 pixels, a field into 216 lines, and fields into not and even to account for interlacing. A six-bit "flash" A/D converts the

ICMC Proceedings 1988

16
stabilized video signal into a 64-shade luminance value for each pixel. Each pixel in a video frame is associated with a line and pixel address which is used to address a bank of 64K RAM chips. These RAMs form a frame buffer which stores the luminance of each pixel in previous frames.

In the movement analysis circuit the luminance of each pixel in a previous frame is retrieved from the frame buffer and presented with the luminance from the present frame to a comparison function in the form of a 65×65 static RAM. This RAM is initialized by the controlling computer to contain a function which compares the old and new luminance to determine if it should be considered to have changed. The output is a single bit to indicate the change, another bit to direct the frame buffer to store the current pixel (useful for hysteresis, as discussed later, and six bits in a video monitor output (also discussed later). The pixel address and the single bit output of the comparison function is presented to the analysis stage.

The analysis stage counts up the pixel changes within a 32×32 matrix of cells called the activity matrix. Since the 128×512 array of pixel change bits is partitioned into 32×32 each cell is four pixels wide and sixteen pixels tall. The analysis stage contains an 11-bit word of RAM for each cell into which the bits from the comparison function are accumulated. At the end of a frame a given cell contains a number representing the number of pixels that have changed in the corresponding portion of the image. This number is really a measure of the quantity of movement but may also be taken as the speed of movement if sampled at regular intervals. A moving object with a lot of luminance variation on its surface (e.g., boldly-patterned clothing) will generate more pixel changes than a uniformly coloured one and hence will appear to be moving faster. This apparent deficiency did not cause problems in practice.

The interface to the controlling computer can be achieved by dual-porting the activity matrix RAM, by switching the RAM between the computer and analysis circuit, or by DMA. An early version was interfaced to an Apple II using the RAM switching method, chosen because it gave the simplest circuit with this computer. The occasional loss of a bit due to contention for the RAM was not a problem in practice. The current version is interfaced to an Atari 1040ST and transmits electronic actuator activity and the monitor output via the ST DMA interface. This method gives a neat and efficient interface with no lost data. The interface allows multiple circuits to be attached to one ST, with the computer receiving data from one circuit while the others are accumulating motion data.

Comparison Functions

The most obvious function for detecting luminance change is to compare the luminance of each pixel of the previous frame with the luminance of the corresponding pixel of the present frame and
indicate change if they differ. This method is useless in practice because noise in the signal causes pixels along the boundaries between zones of differing luminance to fluctuate between the two luminance values. This function interprets such noise as scene movement. A better function requires a pixel to change by more than one shade before indicating movement. This eliminates not only noise but also some slow movements since the motion has to be rapid enough to change a pixel by more than one shade per frame. This apparent problem can actually be an advantage in certain applications.

Arbitrarily slow movements can be detected without noise by incorporating hysteresis into the comparison. Change is indicated when the luminance has changed by two shades since the last change occurred. The comparison function directs the frame buffer to store the new pixel only when a change has occurred. The degree of hysteresis can be increased to handle signals with heavy noise. Since the comparison function is a RAM which is writable by the controlling computer any such function may be implemented.

The effects of noise are greatly reduced but not entirely eliminated by these comparison functions. Our solution has been to include a threshold in the software driver which discriminates between the one-pixel changes caused by the residual noise and the multi-pixel changes caused by real scene movement. Another source of spurious change in the presence of moving scene features such as curtains moving in the wind. These objects can be masked out by extending the threshold into a threshold matrix which can be conveniently generated by measuring the response of the system to a static scene.

Applications

This device is not a "visual system" in the sense that it cannot extract features of the image beyond movement. It cannot "recognise" anything; this kind of visual processing requires more MIPS and computer vision expertise than we are prepared to afford.

Detecting only quantity and location of movement reduces complexity to a level where the analysis can be performed by cheap hardware. This system sees in two-and-a-half dimensions: x, y, and quantity of movement, which is adequate for the design application.

While there is no inherent problem in using a device of this type as a traditional interface to a musical instrument, our circuit is not ideally suited to this application. Television standards dictate a temporal resolution of no better than one sixtieth of a second and spatial resolution is similarly limited. Invisible keyboards, drumsets, etc. are possible but these limitations would make accurate performance of some musical styles impossible. Performance of music on invisible instruments may, however, be of theatrical interest.
A more practical application is as a performer interface in process improvisation. A performer improvises by controlling the value of the parameters of some automated musical process. The parameters of such a process can be directly represented in several ways:

- amount of movement in "active zone";
- location of movement in two dimensions;
- integrated movement in a zone; and so on.

Higher order relationships between areas of activity can also represent parameters of movement can be tracked and parameters controlled by the path.

We designed this device as a people-sensor for interactive installations so it is well suited to this application:

1. it is portable, rugged, and can be easily installed by anyone.
2. No preparation of the visual environment is required; thus irrelevant visual impact is avoided.
3. It is totally electronic and thus reliable.
4. It is not confused by multiple sources of motion.
5. The cameras can be installed in secure positions to avoid tampering and vandalism.
6. It is inconspicuous in comparison to other sensors.

This last point is important to us since we believe that a work can be disfigured by the potent symbolism attached to electronic technology.

Conclusion

The Oculus Raana visual motion interface offers numerous advantages over other motion sensing techniques for installation and performance situations.