METHODS & DESIGNS

A simple stereo television system with application to the measurement of three-dimensional coordinates of fish in schools

T. J. PITCHER and J. E. T. LAWRENCE

University College of North Wales, Bangor, Gwynedd, United Kingdom

Stereo television is achieved through a rapid switch circuit that alternately selects fields from a pair of matched television cameras. Images are recorded with conventional VCRs. After image positions are digitized, stereo coordinates are estimated relative to a calibration grid, which may be removed during an experiment. Necessary geometry, including calibration and iteration procedures, is detailed. The method is flexible and accurate enough to allow the use of zoom, pan, and tilt on both cameras when a calibration grid is retained in view. Examples of the system's application are given for the measurement of three-dimensional coordinates of fish in schools, both for glass-fronted aquarium tanks and in the natural environment using underwater television in sea cages.

Experimental circumstances sometimes require the use of three-dimensional coordinates: for example, in determining the spatial maps of toads (Lock & Collett, 1980) and children (Landau, Glautman, & Spelke, 1981), or in measuring the positions of birds (Major & Dill, 1978) or fish (Pitcher & Partridge, 1979). Measurements that can be transformed to three dimensions are usually made on cine or still photographs by variants of a cast-shadow technique (Cullen, Shaw, & Baldwin, 1966) or by stereo images either directly (e.g., Dill, Dunbrack, & Major, 1981) or with mirrors (e.g., Pitcher, 1973). A review of early methods applied to fish schools is given in Pitcher (1975). Recently, inexpensive video television has been employed to record behavior. Video offers low running costs, ease of rerecording, and time-lapse and slow-motion facilities. These advantages, however, are offset by its relatively low resolution compared with that of film. Stereo images from two cameras can be displayed and recorded in "split-screen" format with a video mixer, but this has the disadvantage of halving the resolution on each image and also creates problems in camera alignment. This paper reports an inexpensive and effective method of obtaining stereo television with maximum image size and resolution.

DESCRIPTION OF TELEVISION SYSTEM

The method is based upon a high-speed switch that alternately selects images from two matched television

cameras. The switched output may be recorded on any video recorder that can be synchronized from an external source and that can replay single interlace video fields [e.g., "U-matic" or four-head "VHS" video cassette recorders (VCRs)]. The television cameras, switch, and VCR must all be driven by a sync pulse generator (black and white or color), but time-base correction is not usually required.

The alternate field switch, shown in Figure 1A, consists of two identical video signal channels, each containing an electronic switch. The output from each switch is fed to a video processing amplifier, which inserts blanking and sync pulses to give a composite video signal.

Each video channel accepts a composite or noncomposite video signal from its respective television camera. The incoming signals are amplified to a suitable level and are clamped at black level to a reference potential. This ensures that the black levels in each signal are identical.

The two video switches are separately controlled by a bistable register having a Q (normal) and a $\bar{Q}$ (inverted) output. A switch is provided for manual or sequential operation. Under manual control, the bistable register may be "set" with Q at logic 1 ($\bar{Q}$ at logic 0), thus enabling camera channel 1 and disabling channel 2, or "reset" with $\bar{Q}$ at logic 1 (Q at logic 0), thus enabling camera channel 2 and disabling channel 1. The manual facility is normally used for setting-up purposes.

In alternate mode, the bistable register is clocked by the field drive signal, and, due to the data input's (D) being connected to $\bar{Q}$, the bistable register changes state at the end of each field. This results in the signals from camera 1 and camera 2 being switched on alternate fields to the video processing amplifier. This amplifier inserts, or reinserts, blanking and synchronizing signals and pro-

This work was supported by Natural Environment Research Council (U.K.) Grant GR3/3919 to T.J.P. We would like to thank DAFS for use of the The Loch Duich facilities. T.J.P. is with the School of Animal Biology, and J.E.T.L. is with the School of Electronic Engineering Science, both at the University College of North Wales, Bangor, Gwynedd LL57 2UW, United Kingdom.

Figure 1. (A) Schematic circuit diagram of the alternate field video switch. (B) Typical output video signals from the alternate field switch.

vides a composite, alternate field, video output signal, as shown in Figure 1B.

A central sync pulse generator provides line drive, field drive, mixed sync, and mixed blanking signals. These basic signals synchronize the scanning of the two cameras, drive the clamp pulse generator, and enable the composite video signal to be produced. The outgoing video signal conforms to the 625-line 50-Hz engineering specification and produces trouble-free recording on any appropriate video recorder. A 60-Hz version for North America is easily constructed using an appropriate sync pulse generator.

STEREO GEOMETRY

The procedure used to obtain three-dimensional coordinates is described in full because relatively few authors have clearly described their methods: Some have used small-angle approximations, which restrict the generality of the method (e.g., Cullen et al., 1966), whereas others

have relied on packaged computer programs in which details of the three-dimensional algorithms are not easily available (e.g., Potel, Sayre, & Mackay, 1980; Potel & Wassersug, 1981). The most accurate and robust technique is to measure parallax against a grid plane of known location (Pitcher, 1975), compensating for systematic errors resulting from lens or electronic distortion with multiple regression calibrations. If this procedure is adopted, the high-precision camera adjustment advocated by Dill et al. (1981) is not necessary, provided that the calibration grid plane is normal to the camera image plane. If the grid location is measured accurately, acceptable results may be obtained even when this does not hold.

Provided that the camera orientation does not change, the calibration grid does not have to be present throughout the experiment once its position on the left and right views has been recorded. Furthermore, if the grid remains in place, it is possible to operate zoom lenses, pan, and tilt on the two cameras to locate small objects accurately and with high resolution. In addition, wide-angle lenses

may be used if sufficient care is taken with calibration, overcoming a restriction of Potel and Wassersug's (1981) method. These advantages accrue from measuring all image positions relative to the calibration grid. The nonlinearity to which electronic television cameras are prone is also easily compensated by this method.

In summary, the procedure is:

- (1) Set up matched video cameras 1 to 2 m apart.
- (2) Align calibration grid normal to cameras as accurately as possible and measure position relative to cameras and the desired coordinate space.
- (3) Measure the grid coordinates on the image from both left and right cameras, taking care to make all measurements on the same television monitor throughout.

(4) Compensate both left and right grid images for systematic nonlinearity by multiple regression.

(5) Using appropriate geometry (see below), calculate the three-dimensional coordinates of a known three-dimensional test object.

(6) If necessary, compensate for any systematic errors on the test object by multiple regression.

(7) To use the system in an experiment, measure the positions of an object on the left and right views relative to the calibration grid (Steps 3 and 4 above) and then use these to give the three-dimensional coordinates (Steps 5 and 6 above).

The geometry for the system set up in air or entirely under water is illustrated in Figure 2A. The algorithm for

Figure 2. Geometry for the estimation of three-dimensional coordinates. (A) For use in field when camera, grid, and object are in air or water. For detailed explanation, see Appendix A. (B) For use in a water-filled aquarium, with double refraction through glass and water. For detailed explanation, see Appendix B.

obtaining three-dimensional coordinates in this instance is given in Appendix A. The cameras do not have to be set up at precisely the same distance from the grid, provided that all distances are carefully measured (cl, cxl, cyl, cr, cxx, cyr in Appendix A). Because the algorithm determines the case from the sign of the parallax, objects to be measured may be in front of or behind the grid or to either side of the camera.

When working with fish in glass-fronted tanks or when looking down into water from cameras mounted in air, an experimenter must take into account refraction, as illustrated in Figure 2B and in the equations of Appendix B. Note that the refracted angles alpha, beta, and gamma must be obtained by iteration; Appendix C sets out the case for refraction at glass and air interfaces. If the calibration grid is on the glass front of the tank, this problem is avoided, but a grid in this position is difficult to maintain in focus when zoom lenses are used. (ALGOL procedures or BASIC subroutines for the methods outlined in Appendices A and B are available from the authors.)

Interpolation

When fast-moving objects are tracked, interpolation may be necessary because successive alternate television fields are 0.02 sec apart (0.01667 sec in North America). By using the two-dimensional tracks of the object against each of the two views of the grid, a researcher may program interpolation for all pairs of fields except the first.

Digitizing the Two Images

Positions on the two television fields are most conveniently measured digitally on still-frame playback. We use Hampton Video Systems image-analyzing equipment (Model VP110 Hampton Video Ltd., Twickenham, London TW1 4BN, United Kingdom). Semiautomatic digitizing with a video disk or a digital frame store with sufficient resolution is a recently developed attractive alternative, although it is unlikely that fully automatic plotting of schooling fish images, which may appear in many different orientations, will be feasible in the near future. Alternatively, the television fields may be recorded on photographs (Partridge, Pitcher, Cullen, & Wilson, 1980) and measurement made with the devices described by Partridge and Cullen (1977) and Partridge, Dawkins, and Amlaner (1978).

Applications

The technique described in Figure 2A has been successfully used with vertically orientated photomultiplier-enhanced underwater television cameras suspended 1.5 m apart at 20-m depth in sea-cage experiments at the Department of Agriculture and Fisheries for Scotland (DAFS) Field site, Loch Duich, Ross-shire, Scotland. Despite an inadvertent mismatch of cameras, satisfactory results giving fish coordinates accurate to within 2 cm (95% limits) were obtained for saithe (*Pollachius virens*) and cod (*Gadus morhua*) schools in the sea cage (Pitcher & Lee, 1983). Interpolation was not necessary because the 30-cm-long fish were swimming at only 0.5 m/sec. The calibra-

tion grid, oriented horizontally to be normal to the camera planes, was removed for these experiments because it would have been in the way of other parts of the apparatus. A similar method has been used for caged mackerel and herring (Magurran & Pitcher, 1984; Pitcher, Magurran, & Edwards, in press). Note that when set up vertically in this way, the x and y planes on Figure 2A are horizontal, whereas the z plane is vertical.

The refraction version of Figure 2B is in use for quantitative observation of antipredator tactics in minnow (*Phoxinus phoxinus*) shoals in large $2 \times 2 \times 1$ m deep glass-fronted aquarium. Zoom, pan, and tilt adjustments to the two cameras are made to keep track of the 50-mm-long minnows. Images from the two cameras are displayed on a pair of adjacent monitors between the two camera operators before the signals enter the switching circuitry. Each operator can then adjust zoom and camera orientation to roughly match that of the other; exact matching is not necessary with this method. With practice, camera operators can obtain usable material for about half of the recorded time: An automatic servo-operated device with equivalent performance would be expensive to construct.

For very fast movement, normal television scanning is not sufficient, because the tracked object moves during the scan and appears as an elongated blur. Electronically and mechanically gated television cameras are available for overcoming this problem.

This method was developed to deal with a large number of objects to be measured at each time, but if only one or a few positions need to be tracked, an attractively simple alternative is the mixing of one image into the other by using reversed polarity for one of the images. If the reversed polarity image is "faded down" slightly to create a "ghost" alongside the first view, measurement is easy. The calibration grid, or a known reference point generated electronically, must be visible for each of the two viewpoints, however. Electronic "keying" of one image into the other was not as successful as conventional mixing and fading, because the keyed image tended to obscure the direct image. One disadvantage with both of these alternatives to the alternately switched television method is that lighting of the objects must remain closely similar irrespective of position, an arrangement that is difficult to achieve in fish tanks.

The alternately switched television record creates a troublesome flickering image to the naked eye when the record is viewed at normal playback speed. However, most contemporary VCRs allow fast play at twice normal speed, by skipping every other field, and this provides a rapid means of scanning the recorded material. More precise selection of sequences to be measured can be made by replay in slow motion.

REFERENCES

- CULLEN, J. M., SHAW, E., & BALDWIN, H. (1966). Methods for measuring the three-dimensional structure of fish schools. *Animal Behaviour*, 13, 534-543.
- DILL, L. M., DUNBRACK, R. L., & MAJOR, P. F. (1981). A new stereophotographic technique for analysing the three dimensional struc-

- ture of fish schools. *Environmental Biology of Fishes*, **6**, 7-13.
- LANDAU, B., GLAUTMAN, H., & SPELKE, E. (1981). Spatial knowledge and geometric representation in a child blind from birth. *Science*, **213**, 1275-1278.
- LOCK, A., & COLLETT, T. S. (1980). The three-dimensional world of a toad. *Proceedings of the Royal Society (London, Series B)*, **206**, 481-487.
- MAGURRAN, A. E., & PITCHER, T. J. (1984). *Caged mackerel behaviour* (Research Report). Aberdeen, Scotland: Marine Laboratory, Department of Agriculture and Fisheries.
- MAJOR, P. F., & DILL, L. M. (1978). The three dimensional structure of airborne bird flocks. *Behavioral Ecology and Sociobiology*, **4**, 111-122.
- PARTRIDGE, B. L., & CULLEN, J. M. (1977). A low-cost interactive coordinate plotter. *Behavior Research Methods & Instrumentation*, **9**, 473-478.
- PARTRIDGE, B. L., DAWKINS, R., & AMLANER, C. (1978). Voltage-controlled oscillators: Inexpensive alternative analog-to-digital converters. *Behavior Research Methods & Instrumentation*, **10**, 712-714.
- PARTRIDGE, B. L., PITCHER, T. J., CULLEN, J. M., & WILSON, J. P. F. (1980). The three-dimensional structure of fish schools. *Behavioral Ecology and Sociobiology*, **6**, 277-288.
- PITCHER, T. J. (1973). The three-dimensional structure of schools in the minnow. *Animal Behaviour*, **21**, 673-686.
- PITCHER, T. J. (1975). A periscopic method for the three-dimensional positions of fish in schools. *Journal of Fisheries Research Board of Canada*, **32**, 1523-1538.
- PITCHER, T. J., & LEE, D. (1983). *Analysis of caged fish behaviour in relation to sonar target strength* (Research Report). Aberdeen, Scotland: Marine Laboratory, Department of Agriculture and Fisheries.
- PITCHER, T. J., MAGURRAN, A. E., & EDWARDS, J. I. (in press). Schooling mackerel and herring choose neighbours of similar size. *Marine Biology*.
- PITCHER, T. J., & PARTRIDGE, B. L. (1979). Fish school density and volume. *Marine Biology*, **54**, 383-394.
- POTEL, M. J., SAYRE, R. C., & MACKAY, S. A. (1980). Graphic input tools for interactive motion analysis. *Computer Graphics and Image Processing*, **14**, 1-23.
- POTEL, M. J., & WASSERSUG, R. J. (1981). Computer tools for the analysis of schooling. *Environmental Biology of Fishes*, **6**, 15-19.

Appendix A

Three-Dimensional Geometry For Camera System in Air or Water

- A reference grid is set up normal to the camera positions as shown in Figure 2a, with the origin of the grid coordinates ($x, y, z = 0$) at top left. Camera positions are: left: c_{xl}, c_{yl}, c_{zl} ; right: c_{xr}, c_{yr}, c_{zr} . Object O in front of grid, above and inside cameras, has images on left camera at B, coordinates x_l, y_l , and on the right camera at A, coordinates x_r, y_r . The method below finds the x -, y - and z -coordinates of the object. Parallax $P = x_l - x_r$.
- Geometry**

In triangle C2,C,N	: CN	=	$c_{xr} - x_r$; $\alpha = \arctan(c_{zr}/CN)$
in triangle C1,M,D'	: MD'	=	$x_l - c_{xl}$; $\beta = \arctan(c_{zl}/MD')$
in triangle C,F,D	: let angle CFD = γ	=	$180 - (\alpha + \beta)$
*by the sine rule,	CF	=	$p \sin(\beta)/\sin(\gamma)$
similarly,	FD	=	$p \sin(\alpha)/\sin(\gamma)$
and so,	CE	=	$p \sin(\beta) \cos(\alpha)/\sin(\gamma)$
and	DE	=	$p \sin(\alpha) \cos(\beta)/\sin(\gamma)$
and	FE	=	$p \sin(\beta) \sin(\alpha)/\sin(\gamma)$
in triangle H,J,E	: JE	=	$c_{yr} - y_r$; $\gamma = \arctan(c_{zr}/JE)$
in triangle H',J,G'	: J'G'	=	$c_{yl} - y_l$; $\gamma' = \arctan(c_{zl}/J'G')$
and so	JG	=	$FE/\tan(\gamma)$
and	J'G	=	$FE/\tan(\gamma')$
- Estimated z -coordinate of object is FE.
- Estimated x -coordinate of object is the average of two values: $x = ((x_r + CE) + (x_l - DE))/2$.
- Estimated y -coordinate of object is the average of two values: $y = ((y_l + J'G) + (y_r + JG))/2$.
- If the object is *behind the grid*, as detected by $x_l - x_r$ going negative, then some of the signs are reversed. For the x -coordinate, $x = ((x_r - CE) + (x_l + DE))/2$, if z is negative.
- For an object *behind the grid*, the y -coordinate, $y = ((y_l - J'G) + (y_r - JG))/2$, if z is negative.
- For an object *behind the grid*, the z -coordinate, $Z = -FE$.
- If the object is *below the cameras*, as detected by $c_{yr} - y_r$ or $c_{yl} - y_l$ going negative, then the signs in (5) or (7) are reversed, since only y is affected.
- If the object is *outside the camera positions*, then the estimate of the x -coordinate for x greater or equal to 0 is given by the following table:

	outside left	outside right
in front of grid	: $((x_l + DE) + (x_r + CE))/2$	$((x_l - DE) + (x_r - CE))/2$
behind grid	: $((x_l - DE) + (x_r - CE))/2$	$((x_l + DE) + (x_r + CE))/2$
criterion	: $c_{xl} > x_l$	$c_{xr} < x_r$

Figure 3. Diagram showing path of a ray doubly refracted through glass and water. Labeling refers to the iteration algorithm given in Appendix C.

2. refractive index

glass/water = μ_{gw} = 0.883

refractive index air/glass = μ_{ag} = 1.51

$V = v_1 + v_2 + v_3$

$v_1 = w/\tan(k)$; $v_2 = g \tan(\alpha_1)$; $v_3 = a \tan(\alpha_2)$

$\alpha_1 = \arcsin(\mu_{gw} \cos(k))$; $\alpha_2 = \arcsin(\mu_{ag} \sin(\alpha_1))$

3. The main equation is: $V = w/\tan(k) + g \tan \arcsin(\mu_{gw} \cos(k)) + a \tan(\arcsin(\mu_{ag} \sin(\arcsin(\mu_{gw} \cos(k))))$, which is difficult to disentangle, since k appears many times. It may be solved iteratively by taking a starting value for k as: start $k = \arctan((a + g + w)/V)$. The iteration procedure is listed in ALGOL and as a BASIC subroutine below, and has been tested thoroughly under all possible values. Note that total internal reflection limits the possible values for k .

ALGOL

```

REAL PROCEDURE iterk2(v);
VALUE v; REAL v;
COMMENT Procedure iterates for angle subtended at grid position v
measured from camera position in x axis, from ray refracted
through water, glass, and air to camera;

BEGIN REAL rgw,pp,ppp,rag,k,w,g,a,crit,stepp,tir,ldif,dif,rhs;
rgw:=0.883;   rag:=1.51;   COMMENT Refractive indices;
tir:=0.719773; COMMENT Total int. reflecth;
stepp:=0.1;   COMMENT Initial step c.6 deg;
COMMENT Convergence criterion; crit:=v*0.0001;
ldif:=1;      g:=1.9;      COMMENT Glass width in cm;
a:=288;       w:=160;      COMMENT Water & air dists;
k:=arctan((a + g + w)/v);  COMMENT Initial value;
repeat: IF k<tir THEN k:=tir;
IF k>1.57079633 THEN k:=1.5704; COMMENT right angle,almost;
pp:=arcsin(rgw*cos(k));ppp:=rag*sin(pp);ppp:=arcsin(ppp);
rhs:=(w/tan(k)) + (g * tan(pp)) +
(a * tan(ppp));
dif:=rhs - v;
IF abs(dif)<crit THEN GOTO exst;
IF sign(dif)*sign(ldif) THEN stepp:=-stepp/2;
k:=k + stepp; ldif:=dif;
GOTO repeat;
exst: iterk2:=k;   COMMENT NB radians;
END procedure iterk2;
```

BASIC

```

5 REM test of iterk2 procedure, BASIC subroutine version
7 PRINT "Test of Iterk2 Procedure" V cm;TAB 10;"k degrees";TAB 25;
  "iterations"
8 LET rad=57.2957729
10 FOR V=0.02 TO .11 STEP .01
20 GOSUB 1000; PRINT V; TAB 10; k*rad; TAB 25; it
30 NEXT V
40 STOP
1000 REM subroutine for iterk2 ex-ALGOL version tip
1005 LET ldif=1: LET it=0
1010 LET mugw=0.883: LET tir=.72272: LET muag=1.51: LET stepp=0.1:
  LET r90=1.57079633
1015 LET a=200: LET g=1.9: LET w=160
1020 LET crit=v*0.0001
1025 LET k=ATN((a+g+w)/V)
1030 IF k<tir THEN LET k=tir
1040 IF k>90 THEN LET k=1.57076
1070 LET rhs=(w/TAN k)+(g*TAN(ASN(muag*cos(k))))+(a*TAN(ASN(muag*
  SIN(ASN(muag*cos(k))))))
1080 LET dif=rhs-v
1090 IF ABS dif<crit THEN GO TO 1200
1100 IF SGN dif<>SGN ldif THEN LET stepp=-stepp/2
1120 LET it=it+1: GO TO 1050
1200 RETURN
```

(Manuscript received May 24, 1984;
revision accepted for publication November 21, 1984.)