Influence of Range, Angle of View, Image Resolution and Image Compression on Underwater Stereo-Video Measurements: High-Definition and Broadcast-Resolution Video Cameras Compared

Authors
Euan S. Harvey
Jordan Goetzee
Bryce McLaren
Tim Langlois
School of Plant Biology,
University of Western Australia

Mark R. Shortis
School of Mathematical
and Geospatial Sciences,
RMIT University, Australia

Abstract
An investigation of how increasing the distance and angle to objects of interest affected the measurement accuracy and precision achievable with high-definition and medium-resolution PAL stereo-video systems was conducted. A test was also conducted to determine whether varying the compression of the imagery influenced measurement accuracy and precision. Measurements of five different lengths of PVC pipe (ranging from 51.5 to 3,001 mm) that represented the lengths of reef fishes routinely sampled with stereo-video were made at 1 m intervals out to the maximum visibility (9 m range) over three different angles (90°, 80°, and 70°). High-definition stereo-video imagery was compressed at three different bit-rates. The results show that higher definition stereo-video imagery allows objects to be measured more accurately and precisely over greater ranges. When both ends of a target can be clearly seen in high-definition stereo-video imagery, the associated error is approximately 1% of the total length of the object. There was no deleterious effect on accuracy or precision from increasing the angle of view. Lower compression did not result in more accurate and precise length estimates. The configuration of a stereo-video system needs to match the task of a particular survey, as changes in the base separation and angle of convergence will affect the accuracy and precision of measurements. With full high-definition systems, smaller lengths (<50 mm) of PVC could not be accurately measured at distances greater than 5 m whereas longer lengths (500–3,001 mm) could be measured with acceptable accuracy and precision at 9 m.

Keywords: Stereo-video, Accuracy and precision, Camera resolution, Video compression, Fish length

Introduction
Underwater video is commonly used as a non-extractive tool to collect information on the abundance, size and biomass of marine organisms (see reviews in Harvey and Mladenov, 2001; Shortis et al., 2009). Diver operated and baited remote underwater stereo-video systems have been used to sample fish assemblages, and in particular to make accurate and precise measurements of fish length in situ (Harvey et al., 2000a, 2000b; 2002a, 2002b; Harman et al., 2003; Watson et al., 2005, 2007, 2009). Stereo-video can also be used to measure the distance to a fish and the angle relative to the center of the stereo-video system (Harvey and Shortis, 1996), which enables accurate definition of the borders of a sampling unit. This facilitates defensible decisions about whether a fish is inside or outside the sample unit (Harvey et al., 2004) and whether it is beyond a range at which stereo measurements are considered to be accurate and reliable (Harvey et al., 2002b).

Because stereo-video technology has been demonstrated to be capable of providing accurate measurements of fish length regardless of user experience, it has been recognized as a valuable tool for determining the length frequency and biomass of cultured fish in aquaculture (Petrell et al., 1997; Harvey et al., 2003; Costa et al., 2006, 2009; Phillips et al., 2008). Stereo-video has not been widely used to sample wild populations of fish but can be used to collect
reliable length data in uncontrolled environments (Phillips et al., 2008; Watson et al., 2009). Accurate and reliable length information is also an important indicator of the health of wild fish stocks (Jennings and Polunin, 1997; Jennings and Kaiser, 1998). Fishing has been shown to result in a decrease in the mean length, biomass or length frequency of target populations (Roberts, 1995; McClanahan et al., 1999).

In a comparison with SCUBA diver length estimates, Harvey et al. (2000b) demonstrated that measurements from a stereo-video system greatly improved the accuracy and precision of length estimates, which had a flow on effect of improving the statistical power of a sampling program to detect changes in the mean length of a population of fish. This was particularly relevant where there was a requirement to detect small changes (30% or less) in the mean length with a high statistical power (e.g., 90%). Similarly, accurate and precise measurements of length are also important if length is to be used to estimate biomass (e.g., Taylor and Willis, 1998) or fecundity (e.g., Davis and West, 1993). Fecundity in fish is directly related to body weight, with the weight of individuals within a species typically ranging across four orders of magnitude (Werner and Gilliam, 1984). Therefore, even small errors in the length estimate can have a large influence on weight and fecundity estimates (e.g., Harvey et al., 2002a).

The accuracy and precision of stereo-video systems have been rigorously tested in a number of experiments using objects of known sizes (Harvey and Shortis, 1996, 1998; Shortis et al., 2000; Harvey et al., 2002b). Using a pair of Sony VX1E Hi 8 camcorders in a configuration designed for surveying reef fishes, Harvey and Shortis (1996) recorded a mean error length of 12.7 mm (1 SD = 4.1 mm). A change in the calibration strategy resulted in significant improvements in the accuracy and precision of measurements from comparable Hi8 video technology (Sony TRV 516 Hi 8 camcorders, mean error = 7.6 mm, SE = 6.6 mm) but also demonstrated that the newer progressive scan digital video technology (Sony TRV 900E digital handycams) was substantially more accurate (mean error = 1.26 mm) and precise (SE = 3.49 mm; Harvey et al., 2002b). These studies also demonstrated that the accuracy and precision of length estimates degrade as the angle of orientation of a fish relative to the stereo-video system increases (particularly when the head and tail of the fish are aligned at greater than 50° perpendicular to the stereo-video system) and as distance increases. However, these studies used objects of a known length, between 100 and 500 mm. It has not been unequivocally demonstrated whether error is absolute (i.e., constant irrespective of the length of the object) or relative to the length of the object being measured.

There are a number of parameters that may affect the accuracy and precision of measurements made by stereo-video systems that have not been examined. Radial and decentering distortions in the camera lens, along with other optical effects such as vignetting and chromatic aberrations, may increase the uncertainty of measurement when the head or tail of a fish is seen in the outer edges of a video image. The profile of these effects generally produces steep gradients of the distortions near the edge of the field of view of the lens. This may be further exacerbated by increasing the distance of the target from the stereo-video system. Previous experience with underwater observations suggests that measurements of larger fish may have a different accuracy and precision than smaller fish.

Over the last two years, most of the major companies manufacturing consumer handycams have released full high-definition (1,920 × 1,080 pixels, FHD) handycams onto the market, replacing medium-resolution systems based on the 4:3 format broadcast TV standards such as NTSC or PAL. FHD is specific to the 1,920 × 1,080 image resolution of high-definition television (HDTV) and should not be confused with cameras that offer so-called HD image resolutions of 1,280 × 720 pixels. This lower resolution is part of the 16:9 format HDTV standard but is an interim resolution between FHD and the standard definition television (SDTV) resolution of 704 × 480 pixels. In comparison to PAL handycams (720 × 576 pixels), the FHD cameras have −2.6× and −2.2× more horizontal and vertical pixel resolution respectively and −5.9× more pixels in total. In theory, measurements made with the higher resolution FHD handycams should be considerably more accurate and precise than measurements made with stereo-video using older PAL handycams. Many of the existing FHD handycams also have the functionality that enables users to decrease the horizontal pixel ratio to record at 4:3 format 1,440 × 1,080 pixel resolution, often also called high-definition, as well as FHD. Although 1,440 × 1,080 pixel resolution is not part of the HDTV standard and instead is based on the equivalent resolution of the 4:3 format image, for convenience in this paper, the format will be referred to as high-definition (HD) to
Methods and Materials

All calibrations and measurements were recorded in the University of Western Australia’s swimming pool on the 23rd of March 2009 in 1.5 m of water. The horizontal visibility of the pool was estimated at 10 m.

Camera Systems

Two stereo-video systems were used for the tests; one system constructed from two Sony TRV900 handycam camcorders (PAL 720 x 576 pixels) fitted with Hama X 0.7 wide-angle lenses and a second system constructed from Sony CX12 handycam camcorders fitted with Raynox C5050 X 0.5 wide-angle lenses. The camcorders were mounted in purpose built underwater housings mounted on an aluminum frame. The camcorders and housings were separated by 750 mm with both housings inwardly converged at 7°. This design enables stereo overlap at 50 cm from the handycams (Figure 1). To eliminate the systematic error of motion parallax, a Light Emitting Device (LED) mounted to the frame in front of the housings and visible in the fields of view of both handycams was used to synchronize the video images.

Image Capture

The digital video imagery recorded to digital tape by the Sony TRV900 camcorders was captured via the i.LINK® interface (Firewire or IEEE 1394) and converted to an AVI format using Adobe Premier Pro®. MPEG videos recorded by the Sony CX12 camcorders were downloaded directly to the PC from the memory stick storage.

Camera Calibration and Image Measurements

Calibrations were made before and after the recordings. The calibration strategy and procedure followed that described by Harvey and Shortis (1996) and Shortis and Harvey (1998) with the exception that Cal 1.1 software from SeaGIS (www.seagis.com.au) was used. Measurement procedures followed that described by Harvey and Shortis (1996) and Harvey et al. (2002b) using Photomasure 1.51 (www.seagis.com.au) using either AVI or MPEG files. Length data and precisions were output from Photomasure as ASCII text files that could be imported directly into Microsoft Excel™ or Access™.

Length Measurements

Five lengths of polyvinyl chloride (PVC) pipe were prepared to test a range of lengths from 51.5 to
3,001 mm. The ends of each pipe were spray-painted black with a white spot on each end to create a contrast, making the ends more visible on the video. The true lengths of the PVC pipes were determined by making five measurements with a fiberglass tape measure and using the mean measurement as the known value to compare against stereo-video measurements.

The fiberglass tape was laid out in a straight line in the swimming pool and fixed to the center of a stereo-video system at the zero mark. The PVC pipe was held horizontally directly in front of the stereo-video system so that both edges could be clearly seen and recorded at 1 m intervals along the tape measure between 2 and 9 m (at 10 m the pipe could no longer be viewed by the snorkelers). In water, measurements of the PVC pipes were made on the same axis and from the same points for each recording by marking the points to be measured. The PVC pipe was held at each interval for a few seconds to allow five replicate, independent measurements to be made.

Comparison of Different Resolution Handycams

The procedure described above was carried out with the PAL (720 x 576) resolution Sony TRV900 camcorders with a 999 mm length of PVC pipe. The entire procedure was repeated with the Sony CX 12 camcorders at two resolution settings: HD (1,440 x 1,080) and FHD (1,920 x 1,080).

Image Compression

All of the imagery recorded by the Sony CX12 was in a MPEG Transport Stream format (MTS). Elecard Converter Studio (http://www.elecard.com/) was used to convert the recorded imagery from MTS to a high-definition MPEG format. To test how different compression settings affect measurement accuracy and precision, video imagery of the 999 mm length of PVC pipe, recorded directly in front of the stereo-video system at distances of 2 to 9 m, was converted using three different image quality settings. The image quality is directly related to the bit-rate of image recording at a constant frame rate such as 30 fps. High bit-rates correspond to high quality and low image compression factors because more information can be recorded. Conversely, low bit-rates imply that less information can be recorded and higher compression factors are applied. The bit-rates were varied to create high (low bit-rate and low quality, 26,000–39,000 kbps), medium (medium bit-rate and medium quality, 36,000–55,000 kbps) and low (high bit-rate and high quality, 52,000–80,000 kbps) compression settings. The same 9 min and 27 s FHD video resulted in 1.87, 2.76 and 3.43 GB files when compressed with low, medium and high bit-rate settings. As above, the PVC pipe was again held at each distance for a few seconds to allow five replicate, independent measurements to be made.

The imagery recorded to test for differences between HD and FHD resolutions (described above) and the effect of different lengths and angles (described below) was converted using a medium compression setting.

Effect of Different Lengths

In total, five different lengths of PVC pipe (51.5, 252, 500, 999 and 3,001 mm) were individually recorded using the Sony CX12 stereo-video system in FHD resolution mode. Five replicate, independent measurements were made of each length of pipe at each distance.

Effect of Increasing Angle

In addition to recording directly in front of the handycams (i.e., 90° perpendicular to the base bar), the orientation was changed with respect to the base bar to 80° and 70° (Figure 2) so that at least one edge of the PVC pipe would be seen and measured in the outer edge of the image. This was repeated for the five lengths of PVC pipe at distances of 2–9 m. As above, five independent measurements were made at each combination of pipe length, distance and angle.

**FIGURE 2**
The angles at which the five different lengths of PVC pipe were measured from the FHD stereo-video system.

**Statistical Analysis**

**Accuracy and Precision**

The accuracy of a stereo-video measurement was defined as the closeness of a measurement to the actual length of the PVC pipe, whereas the precision of measurements refers to the repeatability of replicate measurements. Precision was calculated as the standard error of the five measurements.
(Andrew and Mapstone, 1987). The known length of a PVC pipe was subtracted from the stereo-video measurement to obtain an estimate of measurement error. To compare the different video resolutions and the effects of compression, errors were converted to absolute values (i.e., negative values were converted to positive values). The mean absolute error for each distance was used as an estimate of accuracy, and the standard error was used as an estimate of precision.

To compare the accuracy and precision of measurements of the five lengths of PVC pipe and the effects of different angles, the errors were expressed as a mean percentage of the total length to present a relative error proportional to the different lengths of pipe. For example, a 1 mm error on a 100 mm length of PVC pipe would equal a 1% error. The percentage errors are also presented as actual mean errors to determine whether measurement errors are constant or relative to the length of the object.

Comparisons of Different Resolutions and Compression Settings

Data on the accuracy of the different handycam resolutions and the compression settings set were analyzed with two separate two-factor ((1) Resolution, three levels, fixed: PAL, HD, FHD; Distance, 6 levels, fixed: 2–7 m; (2) Compression, three levels, fixed: high, medium, low; Distance, 8 levels, fixed: 2–9 m)) Permutational Analysis of Variance (PERMANOVA; Anderson et al., 2008) tests. The tests were both based on a Euclidean distance resemblance matrix of correlation factors. Where significant interactions occurred, pair-wise comparisons were made. All analyses were undertaken on absolute errors.

Length, Angle and Distance

To test for differences in the measurement error of the five lengths of PVC pipe and for the three different angles, a three factor (Angle, three levels, fixed: 90°, 80° and 70°; Distance, nine levels, fixed: 1–9 m; Length, five levels, fixed: 51.5, 252, 500, 999 and 3,001 mm) PERMANOVA based on a Euclidean distance resemblance matrix was used. Where appropriate, pair-wise comparisons were undertaken. To investigate whether the errors were relative to the length of the object being measured, the statistical analysis was conducted on the errors expressed as a percentage of the total length of the object.

Results Differences in Handycam Resolutions

From video imagery recorded by the Sony TRV900 handycams, it was not possible to make measurements of the PVC pipe at distances greater than 7 m because the ends of the pipe could not be distinguished reliably. The ends could still be seen reliably on the Sony CX12 imagery at distances up to 9 m. Because of these differences, a direct statistical comparison of accuracy and precision has been limited to measurements made between 2 and 7 m (Table 1).

Pair-wise comparisons of a significant Distance and Resolution interaction show that at some distances there are differences in the accuracy between the three different resolutions of imagery. There is no consistency in the pattern of the errors (Figure 3), but with the exception of measurements taken at 7 m from the PAL imagery, all absolute mean errors are less than 1% of the total length of the PVC pipe. The accuracy of all three resolutions of image capture compares favorably with previous investigations, with errors tending to degrade with distance, which accords with both theory and previous investigations. Generally, the higher the resolution of the imagery, the more precise the resulting measurements (Figure 3), which also accords with the theoretical expectation that greater image resolution will result in a finer resolution of measurement of the 3-D position of the object.

At distances of 8 and 9 m, both accuracy and precision begin to degrade rapidly for the HD and FHD video imagery (13 mm for both HD and FHD at 8 m, and 42 and 58 mm, respectively, at 9 m). Measurements using the HD imagery are 50–70 % less precise than those from the FHD

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>df</th>
<th>MS</th>
<th>Pseudo-F</th>
<th>P (perm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distance</td>
<td>5</td>
<td>115.36</td>
<td>10.765</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Resolution</td>
<td>2</td>
<td>15.654</td>
<td>1.460</td>
<td>0.261</td>
</tr>
<tr>
<td>Distance × Resolution</td>
<td>10</td>
<td>47.748</td>
<td>4.455</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Error</td>
<td>72</td>
<td>10.717</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>89</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Based on 4,999 permutations of the raw data using a Type III sum of squares.

January/February 2010 Volume 44 Number 1
FIGURE 3
The absolute mean error for PAL, HD and FHD stereo-video measurements. Error bars = ±1 SE. SD, standard-definition; HD, high-definition; FHD, full high-definition.

imagery, which is commensurate with a 75% horizontal resolution reduction (i.e., from 1,920 to 1,440 pixels).

Differences in Image Compression Settings
The PERMANOVA analysis indicates that there is an interaction between Compression and Distance (Table 2).

In general, as distance increases, the absolute measurement error also increases (Figure 4). There appear to be anomalies in the errors at ranges of 2 and 4 m for the measurements taken from imagery subject to "low" compression because the errors are much larger than would be expected from the overall trend (Figure 4). It is expected that measurements made from the higher quality imagery should be more accurate and precise than the lower quality imagery.

Effect of Distance and Size on the Accuracy of Measurements
The three-way PERMANOVA analysis indicates a significant interaction between distance to the PVC pipe and the length of the PVC pipe (Table 3). Because no significant differences occurred in the measurements recorded at different angles to the base bar (e.g., for the 500 mm pipe at 70°, 80° and 90° mean error = -0.44 mm ± 1 SE 0.56 mm, -3.31 mm ± 1 SE 0.11 mm, 4.21 mm ± 1 SE 0.14 mm, respectively), the data presented in the graphs have been pooled across the three angles.

Pair-wise comparisons reveal that this interaction is largely driven by the length of the object being measured. For example, for the 51.5 mm PVC pipe, the percent error increases rapidly up to a distance of 5 m, at which point the edges of the pipe could not be distinguished from the background (Figure 5). The same pattern occurs with the 252 mm PVC pipe where reliable measurements could not be made at distances greater than 6 m. The longer PVC pipes (500, 999 and 3,001 mm) could be seen at 8 or 9 m, but again percent error increases as distance increases. Regardless of the length of the PVC pipe, the precision of measurements degrades with increasing distance (Figure 5). The longer PVC pipes could not be measured at smaller distances because the stereo overlap did not allow the full length to be visible in both images. The mean percent error for all lengths of PVC pipe was 1.5% of the total length (SE ± 0.07%).

For the 51.5 mm PVC pipe, the actual mean errors are less than 1 mm and reduce to 0.5 mm (~1% error) when only measurements between 1 and 3 m are included (Figure 6). The mean percentage errors on the longer lengths of PVC pipe are relatively small in comparison. Hence, the percentage length error for the 51.5 mm pipe does not agree with the error trend for the longer lengths of pipe.

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>df</th>
<th>MS</th>
<th>Pseudo-F</th>
<th>P (perm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compression</td>
<td>2</td>
<td>28.782</td>
<td>2.935</td>
<td>0.056</td>
</tr>
<tr>
<td>Distance</td>
<td>7</td>
<td>43.968</td>
<td>4.484</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Compression × Distance</td>
<td>14</td>
<td>17.907</td>
<td>1.8262</td>
<td>0.041</td>
</tr>
<tr>
<td>Error</td>
<td>96</td>
<td>9.805</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>119</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Based on 4,999 permutations of the raw data using a Type III sum of squares.
and the percentage errors are larger than those indicated by previous investigations. When the real mean errors are plotted (Figure 6), it becomes apparent that the error for the 3,001 mm pipe exhibits much larger errors than the other lengths of pipe. These apparent anomalies require further investigation; however, as a generalization, it would appear that when the edges of an object are clearly visible, it is possible to measure, on average, within 1% of the true length of the object.

**Discussion**

From a FHD stereo-video system, the mean error for a PVC pipe at distances less than 8 m is 0.7%. This is much improved on older PAL systems tested by Harvey and Shortis (1996), who found the total mean percentage error of measurements to a distance of 7 m of a 458 mm sized fish silhouette to be 4.6%. An NTSC system used by Petrel et al. (1997) measured the fork lengths of anesthetized Chinook Salmon to within 3% at a range of approximately 2 m. The Sony CX12 FHD handycams tested here provide increases in the accuracy and precision of measurements over a range of angles and at much greater distances in comparison to stereo-video systems constructed from PAL or NTSC resolution handycams.

This research also demonstrated that when the edges of the object of interest can be clearly seen, the error associated with length measurements is accurate to approximately 1% of the total length of the object being measured. However, objects can only be measured when they can be clearly seen. This limits the distance at which smaller objects can be usefully measured and has implications for multispecies studies attempting to count and measure fish assemblages, as bigger fish will be able to be identified and measured across a much larger range of distances than smaller fish.

There are a number of advantages to using HD and FHD stereo-video systems in comparison to the older digital video technology, the main one being the increase in the distance at which it is possible to clearly see and measure objects of interest. At close range (1–4 m), the real differences in accuracy between the PAL, HD and FHD stereo-video systems tested were minimal, but because of the ability to more clearly define the edges of objects of interest at distances

**TABLE 3**

Results of three-factor PERMANOVA on the accuracy and precision of measurements of fish silhouettes.

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>df</th>
<th>MS</th>
<th>Pseudo-F</th>
<th>P (perm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Angle to target (An)</td>
<td>2</td>
<td>4.095</td>
<td>1.326</td>
<td>0.261</td>
</tr>
<tr>
<td>Distance to target (Di)</td>
<td>8</td>
<td>48.072</td>
<td>15.572</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Length of target (Le)</td>
<td>4</td>
<td>141.780</td>
<td>45.925</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>An × Di</td>
<td>15</td>
<td>5.243</td>
<td>1.698</td>
<td>0.052</td>
</tr>
<tr>
<td>An × Le</td>
<td>7</td>
<td>2.217</td>
<td>0.718</td>
<td>0.657</td>
</tr>
<tr>
<td>Di × Le</td>
<td>20</td>
<td>27.337</td>
<td>8.855</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>An × Di × Le</td>
<td>24</td>
<td>3.353</td>
<td>1.086</td>
<td>0.362</td>
</tr>
<tr>
<td>Res</td>
<td>324</td>
<td>3.087</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>404</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Significant results are in bold, using a significance level of 0.05.
of 4 m or greater, the accuracy of measurements from the HD and FHD systems is better than PAL technology. Although there does not appear to be any great differences between the accuracy of HD and FHD systems, a great advantage of FHD is its improved repeatability (precision). During analysis of field data, often only one measurement of a fish is made, for example, when taking length frequency data (e.g., Watson et al., 2009). In these circumstances, FHD systems will offer far greater efficiency.

In PAL imagery recorded from objects at close range (2–3 m), the edges of the lengths of the PVC pipe were easier to see and hence measurements were more accurate than those using the HD and FHD systems. This was not expected and could be attributed to three factors. The first is that the PAL camera recorded in a progressive scan format. The PAL progressive scan captures the entire image simultaneously whereas the interline capture format used by the HD and FHD cameras captures the odd and even lines of the image separately (Shortis and Snow, 1995). The cameras and the object were stationary during the image capture, so image motion displacement between the odd and even lines is not a significant factor. Nevertheless, the constant motion of the water medium will have a small influence on the quality of the image when interline image sensors are employed. It is recommended that stereo-video systems be developed using progressive scan HD or FHD handycams. It was also possible that this result might be an artifact of the calibration procedure or compression effects associated with the conversion and compression of the CX12 MTS files into MPEG files in comparison to the AVI files captured for the PAL handycams. To check this factor, camera calibrations, recordings, compression of the imagery and measurements were repeated. There were no changes in the interpretation of the results. A third possibility is that the Sony TRV 900 PAL cameras contained three CCD sensors whereas the Sony CX12 used to record the HD and FHD imagery contains one CMOS sensor. It is not clear whether this unexpected anomaly was caused by the effects of the numbers and types of sensors contained within a camera, or the effects of progressive scan versus interlaced cameras, or a combination of both. This is an area that would benefit from further investigation.

The other unexpected result was the measurement error associated with the 3,001 mm length of PVC pipe. This was much greater than was predicted and differs from other results (Harvey and Shortis, 1996; Harvey et al., 2002b). Although the ends of this piece of PVC pipe would have been at the extreme left and right edges of the field of view, it is unlikely that lens distortions played a large role because this effect would have been noticeable in some of the shorter lengths of PVC at closer ranges as well. Radial and decentering distortions are accurately modeled in the calibration procedure, and no significant
optical aberrations were apparent in either stereo-video system. The higher than expected errors are mostly likely associated with the separation of the underwater housings on the base bar and their angle of convergence. These systems were designed to measure reef fish from 5 to 200 cm in length at distances between 1.5 and 7 m. For optimal performance, the configuration of a stereo-video system may need to change depending on the size of the targets to be measured and their distance from the stereo-video system. The base separation of the cameras should be one third to one fifth of the camera to object range for acceptable precision (Harvey and Shortis, 1996). As an example, a system measuring smaller targets, e.g., the 10 mm diameter inhalant siphons of a shellfish (Shortis et al., 2000), was constructed with a much smaller 25 cm base separation and an increased angle camera convergence. A system specifically designed to measure longer targets at larger ranges (e.g., sharks or whales) would have a base separation of the order of 1.5 to 2 m. A much larger system would have to employ a different calibration approach as a consequence.

It is apparent from Figure 6 that the stereo-video systems consistently overestimated the lengths of the PVC pipes, irrespective of their length or distance. This outcome may be caused by the blurring of the edges of objects from attenuation through the water medium, the optical transfer through the lens (also known as the point spread function) and the quantization by the image sensor (conversion of the continuous color scene into an image composed of discrete pixels, each with 256 intensity levels). Research in medical imaging suggests length measurement errors as large as 1 to 1.5 pixels (Goldstein, 2000) are possible, which would correspond to apparent increases in length of the pipes from approximately 1 mm at 2 m range to 5 mm at 7 m range for the stereo-video systems. Additional research is needed to quantify this effect in the underwater environment.

Although FHD provides more precise measurement compared to HD or PAL, its disadvantage is the increase in the memory required to store imagery and the computer processing power required to analyze the imagery. The difference in file sizes between the HD and FHD footage is approximately 65%, with a 10 min HD video being 1.75 GB in size and a FHD video of the same length being 2.68 GB in size.

Although high compression may decrease the sizes of the files, the visibly prominent disadvantage of high compression is a loss in the “crispness” of the imagery. This effect, combined with loss of contrast from attenuation, makes it more difficult to select the edges of the objects of interest and was expected to result in a degradation in the accuracy and precision of measurements, particularly at distance. It was therefore surprising to find no significant improvements in measurement precision and accuracy associated with lower compression. Previous research on the quality of digital images used for photogrammetric applications suggests that moderate levels of compression have little impact on the accuracy and precision of the measurements, and the degradation is essentially linear (Kiefner and Hahn, 2000). It is only when very high levels of compression, corresponding to still images compressed at ratios of 1:100, that there is a significant degradation of accuracy and precision caused by the disturbance of image locations by compression artifacts (Zhilin et al., 2002). Results for these tests show that reduced file sizes can be effectively exploited to overcome storage and processing issues associated with the FHD format. This comparison must be qualified in that the MPEG compression algorithm for video images and, for example, the JPEG compression algorithm for still images share some common features such as the discrete cosine transform (DCT), but there are also some fundamental differences in the MPEG compression technique used for video. Irrespective of the similarities and differences between the compression algorithms, the common outcome is that moderate levels of compression have little impact on the accuracy and precision of measurements. One important advantage of using low compression and the FHD video cameras is that the increased image quality makes the taxonomic identification of fish easier.

The increase in the resolution of handycams has some significant advantages when they are incorporated into stereo-video systems. One of the major hindrances to the uptake of this technology is the manual processing of the imagery to generate data for statistical analysis and interpretation. Even with a sophisticated software interface to enable operators to count and measure fish, the manual processing of the imagery represents a significant cost. The higher resolution imagery increases the likelihood of developing a solution that will facilitate implementation of research strategies aimed at semi-automated processes for recognizing, counting, and measuring fish.

Acknowledgments

This study was conducted with financial assistance from the University of Western Australia (UWA), with
funding from the Marine Science Honours program. The authors would like to thank Peter Barnes, Russell Cole and three anonymous reviewers for useful comments on the drafts leading to this manuscript. Staff from the university’s swimming pool are gratefully acknowledged for allowing access to the pool and their support.

Lead Author:
Euan Harvey
School of Plant Biology
The University of Western Australia
35 Stirling Highway, Crawley,
Western Australia 6009
Email: euanh@cyllene.uwa.edu.au

Appendix of Acronyms and Abbreviations
AVCHD: Advanced Video Codec
High Definition
AVI: Audio Video Interleaved (a video file format)
CCD: Charge Coupled Device (a type of image sensor)
CMOS: Complementary Metal Oxide Semiconductor (a type of image sensor)
FHD: Full High Definition (1920 x 1080 pixels)
HD: High Definition (1440 x 1080 pixels)
HDTV: High Definition Television
LED: Light Emitting Device
MPEG: Moving Picture Experts Group. MPEG is a group of compression standards for video and audio.
MTS: MPEG Transport Stream (a video format)
NTSC: National Television System Committee. NTSC is the analog television system used in most of the Americas, Japan, South Korea, Taiwan, Burma, and some Pacific island nations. In the context of handycams NTSC video is often 720 x 480 pixels.
PAL: Phase Alternating Line (720 x 576 pixels). PAL is an analogue television encoding system used in broadcast television systems in large parts of the world.
PVC: Polyvinyl Chloride
SCUBA: Self Contained Underwater Breathing Apparatus
SDTV: Standard Definition Television (704 x 480 pixels)

References
Davis, T.L.O., West, G.J. 1993. Maturation, reproductive seasonality, fecundity, and spawning frequency in Lutjanus vittatus (Quoy and Gaimard) from the North West Shelf of Australia. Fish Bull. 91:224-236.
Harvey, E.S., Cappo, M., Shortis, M., Robson, S., Buchanan, J., Speare, P. 2003. The accuracy and precision of underwater measurements of length and maximum body depth of southern bluefin tuna (Thunnus maccoyii) with a stereo-video camera system. Fish Res. 63:315-326.


