INTRODUCTION

The direct linear transformation (DLT) (Abdel-Aziz and Karara, 1971; Shapiro, 1978; Walton, 1981) has been the most commonly used three-dimensional (3-D) cinematography/videography method in biomechanics research. The DLT uses two or more two-dimensional views of a well-surveyed control object to solve for camera parameters relating the image space to the object space. Both the reference frame and the units of measurement are defined by the (known) 3-D coordinates of points on the control object. Proposed as a low-cost alternative to the DLT, the non-linear transformation (NLT) method (Dapena et al., 1982; Dapena, 1985a) also uses a control object, but the precise 3-D coordinates of points on the control object need not be known. The reconstruction algorithm uses these points and approximate information about the camera arrangement to solve for one camera's orientation relative to the other camera using an iterative approach. A scaling factor and reference frame transformation external to the cameras must be defined in the calibration.

To be most accurate, one should not extrapolate the DLT beyond the limits set by the control object (Wood and Marshall, 1986; Chen et al., 1994). In reality, many researchers are faced with the task of studying activities that exceed the size of their DLT control objects. In many instances it may be impractical to build and carefully survey a large enough DLT object so that extrapolation is unnecessary. One alternative is to use the NLT. Because one does not need to survey the control object for the NLT, it is much less costly to implement than the DLT especially for large control volumes. A single PVC pole with two marks on it to define one known length can be carried from place to place within the activity volume to 'build' any size control object desired. Additional marks on the pole and/or additional pole locations can be used to increase the number of control points. Accuracy of the NLT method, however, still needs to be established. The purpose of this study was to compare the accuracy of the NLT to that of the DLT with and without extrapolation.

METHODS

A cuboid DLT control object was constructed using 27 strings hung vertically from the ceiling of a large room with spherical markers, 3 cm in diameter, located at surveyed positions along the strings. The horizontal positions of the strings were determined by triangulation based on the horizontal distances measured from two points (A and B) on the floor (see Fig. 1). The vertical positions of the spheres were measured from the floor using a sliding caliper from an anthropometry kit. The floor was not assumed to be level. The relative height of the floor below each string was determined precisely using a transit theodolite and used to correct the vertical measurements made with the sliding caliper. The dimensions of the cuboid were approximately 185 cm x 490 cm x 213 cm (X, Y and Z, respectively) with 108 points (four points per string) surveyed to be accurate within ± 0.05 cm (see Fig. 1).

Two control volumes were defined for the DLT which used varying numbers of control points for determining camera parameters. Control Volume I (CV-I) encompassed the entire surveyed volume of strings (185 cm x 490 cm x 213 cm). Four different combinations of control points (16, 24, 40, and 60; see Table 1) were used to predict the location of 48 'non-control points' (i.e., surveyed points which were not used in the computation of camera parameters) located throughout the control volume (four points each on strings 4–7, 11–14, 18, 19, 23, and 24; see Fig. 1).

The smaller control volume used for extrapolation (CV-II) was defined within one end of CV-I (185 cm x 121 cm x 213 cm, see Fig. 1). Camera parameters based on these 24 control points were used to predict the location of the 16 non-control points inside CV-II and 68 non-control points outside CV-II. These non-control points were referred to as being at a level of extrapolation defined by the difference between the y-coordinates of the extrapolated points of interest and the right side edge of CV-II expressed as a percentage of CV-II's y-dimension. This
created six levels of extrapolation, (approximately) 50, 100, 150, 200, 250 and 300% (see Fig. 1).

The NLT control volume was designed to approximate the dimensions of the large DLT control volume. A single, 223-cm-long pole made from 2.5-cm-diameter white PVC tubing was used to create the NLT control object. Six points were marked on the pole with black electrical tape. The distances between points were 50, 50, 23, 50 and 50 cm, respectively. The pole was moved to 11 different vertical and four horizontal positions encompassing the activity volume which, along with five reference frame points (see below), created a 95-point NLT control object. The two 100 cm distances between points 1 and 3 and between points 4 and 6 on the pole provided 30 known lengths (2 lengths x 15 pole locations) throughout the calibrated volume which were used in the NLT calculations. A 1 m x 1 m ‘cross’ placed at a known distance from each camera was used to determine the NLT’s normalized scaling factors (d and l; see Dapena et al., 1982).

Five additional points (two on a vertical string and three on the laboratory floor) were used to define a laboratory reference frame. The two string points, approximately 2 m apart, were used to define the Z-axis (vertically up). A provisional X-axis was defined by two points placed approximately 3 m apart in the Y-direction on the floor. (Since the floor was not assumed to be horizontal, Y’ was not necessarily orthogonal to Z.) The X-axis was defined by the cross product of Y’ and Z. The true Y-axis was then defined by the cross product of Z and X. Finally, a third point on the floor was used to translate all points to an origin common with the DLT control object. This reference frame transformation is an integral part of the NLT but is optional with the DLT. A transformation should be used with the DLT whenever it is difficult to orient the DLT control object precisely with the desired laboratory reference frame. To determine the effect of this transformation on accuracy, the standard DLT was compared with and without a reference frame transformation. The extrapolated DLT, however, was not tested with a reference frame transformation since three of the five reference frame points (including the origin) fell outside of CV-II and hence were themselves extrapolated points (see later).

Two gen-locked Panasonic WD-5100 video cameras plus AG-7400 S-VHS portable VCRs were used to record the 2-D images of all points used in the experiments. A Peak Performance Technologies (Englewood, CO) motion analysis system was used to digitize the videotapes. Control and non-control point locations were digitized three times and averaged together prior to computing 3-D coordinates. The error of prediction, $e$, was calculated as the resultants of the difference between the predicted and surveyed locations of selected points from the DLT control object in the X-, Y-, and Z-directions ($d_x, d_y, d_z$, respectively).

$$e = \sqrt{d_x^2 + d_y^2 + d_z^2}.$$ 

The resultant errors of prediction were normalized to the longest 3-D diagonal of the appropriate volume of points being predicted. For example, the largest diagonal was for CV-I (when repredicting control points). The smallest diagonals were those for the (essentially) flat areas occupied by the non-control points at each extrapolation level.
Table 2. Comparison of NLT to standard DLT using different numbers of DLT control points and with and without a reference frame transformation

<table>
<thead>
<tr>
<th>Method/control volume showing which points were predicted</th>
<th>Diagonal length (cm)</th>
<th>Absolute error without transformation in cm. Mean (S.D.)</th>
<th>Normalized error without transformation (% diagonal)</th>
<th>Absolute error with transformation in cm. Mean (S.D.)</th>
<th>Normalized error with transformation (% diagonal)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DLT CV-IA-16 cp repredicting same 16 cp</td>
<td>561.3</td>
<td>0.53 (0.30)</td>
<td>0.09</td>
<td>1.48 (0.43)</td>
<td>0.26</td>
</tr>
<tr>
<td>DLT CV-IB-24 cp repredicting same 24 cp</td>
<td>561.3</td>
<td>0.90 (0.35)</td>
<td>0.16</td>
<td>1.43 (0.59)</td>
<td>0.25</td>
</tr>
<tr>
<td>DLT CV-IC-40 cp repredicting same 40 cp</td>
<td>561.3</td>
<td>0.98 (0.30)</td>
<td>0.17</td>
<td>1.26 (0.55)</td>
<td>0.22</td>
</tr>
<tr>
<td>DLT CV-ID-60 cp repredicting same 60 cp</td>
<td>561.3</td>
<td>1.03 (0.43)</td>
<td>0.18</td>
<td>1.20 (0.54)</td>
<td>0.21</td>
</tr>
<tr>
<td>DLT CV-IA-16 cp predicting 48 ncp</td>
<td>436.4</td>
<td>1.34 (0.52)*</td>
<td>0.31</td>
<td>1.72 (0.60)**</td>
<td>0.39</td>
</tr>
<tr>
<td>DLT CV-IB-24 cp predicting 48 ncp</td>
<td>436.4</td>
<td>1.21 (0.56)+</td>
<td>0.28</td>
<td>1.38 (0.54)+†</td>
<td>0.32</td>
</tr>
<tr>
<td>DLT CV-IC-40 cp predicting 48 ncp</td>
<td>436.4</td>
<td>1.05 (0.46)‡</td>
<td>0.24</td>
<td>1.14 (0.56)†‡</td>
<td>0.26</td>
</tr>
<tr>
<td>DLT CV-ID-60 cp predicting 48 ncp</td>
<td>436.4</td>
<td>0.85 (0.37)§</td>
<td>0.19</td>
<td>0.93 (0.42)§</td>
<td>0.21</td>
</tr>
<tr>
<td>NLT-95 cp predicting 48 ncp</td>
<td>436.4</td>
<td>N.A.</td>
<td>N.A.</td>
<td>1.55 (0.51)</td>
<td>0.36</td>
</tr>
</tbody>
</table>

*Not significantly different from NLT ($t = 1.98$, df = 94, $p = 0.05$).
**Not significantly different from NLT ($t = 1.48$, df = 94, $p = 0.14$).
†Significantly more accurate than NLT ($t = 3.08$, df = 94, $p = 0.0027$).
‡Not significantly different from NLT ($t = 1.57$, df = 94, $p = 0.12$).
§Significantly more accurate than NLT ($t = 4.99$, df = 94, $p = 2.8 	imes 10^{-6}$).
**Significantly more accurate than NLT ($t = 3.71$, df = 94, $p = 3.5 	imes 10^{-6}$).
$Significantly more accurate than NLT ($t = 7.62$, df = 94, $p = 2.0 	imes 10^{-11}$).
$Significantly more accurate than NLT ($t = 6.43$, df = 94, $p = 5.2 	imes 10^{-16}$).
Note: cp = control points, ncp = non-control points.

RESULTS

The smallest resultant errors were recorded with the standard DLT repredicting control points without a reference frame transformation (see Table 2). The errors in repredicting control point locations grew, however, as the number of control points increased from 16 to 60 (see Fig. 2). The errors in predicting non-control points, on the other hand, decreased as the number of control points increased (see Fig. 2). Subjecting the DLT to a reference frame transformation increased the errors for both control points and non-control points, especially for the 16 control point condition (Table 2). There were no significant differences in predicting non-control points between NLT and DLT for the 16 control point condition without a reference frame transformation and for the 16 and 24 control point conditions when the DLT was subjected to a reference frame transformation. When a greater number of control points were used, the DLT became significantly more accurate than the NLT.

As expected, the accuracy of the DLT was reduced substantially as the level of extrapolation increased (see Table 3). At 300% extrapolation, the absolute errors had grown nearly tenfold compared to the 0% condition (6.01 vs 0.61 cm). Based on errors of predicting identical non-control points, the extrapolated DLT was significantly more accurate than the NLT only if the level of extrapolation was limited to 50% or less (Table 3).

![Fig. 2. Plot of mean absolute error for control points and non-control points for the standard DLT using various numbers of control points in the calibration. Shown is the interaction between control point errors and non-control point errors.](image)

The crossover point (beyond which the NLT became significantly more accurate than the extrapolated DLT) was at about 100% extrapolation.

DISCUSSION

When the need arises to study 3-D motion occurring in large activity volumes, the researcher often is faced with a dilemma. If
the activity volume exceeds the size of the available DLT control object, should a person (1) construct and survey a custom large-scale (often single use) DLT control object for that particular project, (2) extrapolate beyond a smaller (but re-useable) DLT control object, or (3) choose an alternative method like the NLT? The results of this study provide some valuable insights to help the researcher choose an appropriate course of action.

When predicting non-control points (the 'true test'), the standard DLT using 60 control points was found to be most accurate. When re-predicting control points, however, the standard DLT using the fewest (16) control points produced the greatest apparent accuracy. André et al. (1990) refer to errors in re-predicting control points as 'DLT error' and show that the actual errors can be much higher especially if the control points themselves contain errors in their surveyed positions. Baseline accuracy on the re-prediction of a small number of control points is misleading because of the potential of 'overfitting' the equations to those few control points. The fact that the reference frame transformation had the greatest effect on the 16 control point condition and that this 16 point calibration was least accurate at predicting non-control points is evidence that 'overfitting' occurred.

The addition of a reference frame transformation tended to increase the errors for the highly accurate standard DLT. With the reference frame depending on only five points (rather than the entire set of control points) random errors in the five point set statistically are more likely to produce an inaccurate reference frame than random errors in the 16-60 point sets. However, the effects of such a transformation on the extrapolated DLT may be different. Here the errors are already fairly large. It is possible that such a transformation could actually improve the accuracy for certain extrapolated points, namely those which fall in the vicinity of the points used to define the new reference frame (especially the origin). To investigate this question in full, however, would require more space than we are allowed in this technical note. Thus a comparison of the NLT (which uses a reference frame transformation implicitly) and the extrapolated DLT also with a reference frame transformation will have to wait for some future study.

The results of the present study are similar in some ways but different in others to the findings of Chen et al. (1994). Using a similar shape but smaller DLT control object than in the present study, Chen et al. had a total of 32 surveyed points available to use as control points or non-control points. Using different numbers of control points (n = 8, 12, 16, 20, and 24) they predicted all remaining (32 - n) non-control points for each test. Chen et al. reported increased accuracy of the DLT as the number of control points increased from 8 to 16, but found no additional reductions in error as the number of control points increased to 24. A severe limitation of the Chen et al. study, however, is that they did not make these comparisons using the same set of non-control points each time as was done in the present study. Hence, their results cannot be compared directly with ours. We found continued improvement in the DLT when the number of control points increased beyond 24 all the way to our maximum of 60. Additional improvement may be possible beyond 60 since the errors had not yet reached an apparent minimum.

Dapena (1985b) has presented the only previous comparison between the DLT and NLT methods of 3-D reconstruction. Dapena reported the accuracy of the NLT to be about the same as the DLT. However, Dapena later altered his definition of percent error. When applied to his original data, the new criterion led him to conclude that the standard DLT is more accurate than the NLT (Dapena, personal communication, 18 January 1993). The present research generally supports this conclusion. The DLT inherently uses more information (i.e., the known 3-D coordinates of the control points) when solving for camera parameters. The NLT solves for camera parameters by iteration without this information.

The present results show that the NLT is not far behind the DLT in accuracy, however. If one uses only 16-20 DLT control points, as Chen et al. (1994) recommend, then either method could be used with essentially the same accuracy. Comparing the NLT to the extrapolated DLT (see Table 3), the main advantage of the NLT is revealed. The DLT appears to be remarkably accurate when the extrapolation is limited to 50% or less. However, if the project requires extrapolation of 100% or more, the researcher should consider using the NLT. The NLT is inexpensive and only slightly less accurate than the standard DLT. It lends itself well for field work where transporting a large DLT control object is not feasible. A single pole with two marks on it to define one known length can be carried from place to place within the activity volume to 'build' any size control object desired. Additional marks on the pole and/or additional pole locations can be used to increase the number of control points. Most importantly, no surveying of control points is required other than to measure the one known length on the pole.

Table 3. The effect of extrapolation on the accuracy of the DLT (CV-II using 24 control points). The NLT is included for comparison by predicting the same non-control points as the DLT at each level

<table>
<thead>
<tr>
<th>Level of extrapolation relative to DLT</th>
<th>Diagonal length (cm)</th>
<th>DLT absolute error in cm. Mean (S.D.)</th>
<th>DLT normalized error (% diagonal)</th>
<th>NLT absolute error in cm. Mean (S.D.)</th>
<th>NLT normalized error (% diagonal)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0% extrapolation—16 ncp</td>
<td>275.2</td>
<td>0.61 (0.14)*</td>
<td>0.22</td>
<td>1.93 (0.52)</td>
<td>0.70</td>
</tr>
<tr>
<td>50% extrapolation—16 ncp</td>
<td>273.7</td>
<td>0.81 (0.33)*</td>
<td>0.30</td>
<td>1.32 (0.31)</td>
<td>0.48</td>
</tr>
<tr>
<td>100% extrapolation—12 ncp</td>
<td>275.0</td>
<td>2.11 (0.38)*</td>
<td>0.77</td>
<td>1.81 (0.35)</td>
<td>0.66</td>
</tr>
<tr>
<td>150% extrapolation—8 ncp</td>
<td>212.2</td>
<td>2.21 (0.46)*</td>
<td>1.04</td>
<td>1.30 (0.46)</td>
<td>0.61</td>
</tr>
<tr>
<td>200% extrapolation—12 ncp</td>
<td>274.5</td>
<td>3.63 (0.23)*</td>
<td>1.32</td>
<td>1.84 (0.71)</td>
<td>0.67</td>
</tr>
<tr>
<td>250% extrapolation—8 ncp</td>
<td>210.8</td>
<td>3.73 (0.77)†</td>
<td>1.77</td>
<td>1.39 (0.37)</td>
<td>0.66</td>
</tr>
<tr>
<td>300% extrapolation—12 ncp</td>
<td>274.5</td>
<td>6.01 (1.42)**</td>
<td>2.19</td>
<td>1.88 (0.82)</td>
<td>0.68</td>
</tr>
</tbody>
</table>

*Significantly more accurate than NLT (t = 9.49, df = 30, p = 1.5 x 10^-10).
†Significantly more accurate than NLT (t = 3.25, df = 30, p = 0.0028).
‡Significantly less accurate than NLT (t = 1.39, df = 22, p = 0.067).
§Significantly less accurate than NLT (t = 5.07, df = 22, p = 4.4 x 10^-7).
||Significantly less accurate than NLT (t = 7.25, df = 14, p = 4.2 x 10^-6).
**Significantly less accurate than NLT (t = 8.35, df = 14, p = 2.9 x 10^-8).

Note: These tests compare the NLT with a reference frame transformation to the extrapolated DLT without a reference frame transformation.
Acknowledgements—The authors would like to acknowledge Phil Cheetham, John DeWitt, Gary Heise, Al Hreljac, Tony Marsh, Barry Munkasy, Rachael Seidler, Chris Sherwood, and Peter Vint for their assistance in this project.

REFERENCES


