An Image-Based Three-Dimensional Digitizer for Pre-Decorating Thermoformed Parts

J.P. Mellor
Rose-Hulman Institute of Technology*
jpmellor@rose-hulman.edu

Abstract

Thermoformed plastic parts are pervasive in modern products. Decorating these thermoformed parts is a critical step in their manufacture. For parts with significant relief, decorating the part after it is formed can be prohibitively expensive. Prior to thermoforming, the plastic which is in flat sheets or rolls can easily be fed through a printer. However, during thermoforming the plastic flows and stretches to conform to the shape of the mold distorting the original decoration. This paper describes a device called the *Image-Based Three-Dimensional Digitizer* (IB3D) which measures the plastic deformation using vision techniques. Knowing the deformation allows us to pre-distort the graphic printed on the flat sheets such that as the plastic flows into the thermoforming mold the decoration undistorts.

1   Introduction

Plastic has become a vital raw material. Most products have at least some plastic in them. A large portion of these plastic parts are thermoformed. Thermoforming produces quality, close tolerance parts quickly and economically (see [3]). The aesthetic value, and to a large extent the overall value, of the thermoformed part is determined by the quality with which it is decorated.

Prior to thermoforming the plastic is in flat sheets or rolls and can easily be fed through a printer. During thermoforming the plastic flows and stretches to conform to the shape of the mold distorting the original decoration. For parts with complex shapes or significant relief, this distortion is intolerable. One obvious solution is to decorate the part after it is formed. The decoration will no longer deform, but because of its 3-D shape, it is no longer possible to feed the plastic through a printer. In general, complicated and expensive techniques are needed to decorate formed parts and in many cases decoration becomes prohibitively expensive.

Another approach is to pre-decorate the part, but also *pre-distort* the decoration so that as the plastic flows and stretches the decoration aligns properly and deforms to its undistorted form. The challenge with this approach is figuring out how to pre-distort the decoration. Until now, the cost of determining the necessary pre-distortion has been prohibitive in all but the most extreme cases. This paper describes a novel method for recovering the pre-distortion necessary for pre-decorating thermoformed parts with complex shapes

and/or significant relief. We also describe a device, the *Image-Based Three-Dimensional Digitizer* (IB3D), which is capable of efficiently and cost effectively recovering this pre-distortion.

2   Approach

The IB3D uses a camera platform, which supports a trinocular stereo rig, to recover the two-dimensional grid coordinates as well as three-dimensional coordinates. A special striped grid is printed onto a piece of plastic prior to thermoforming. The IB3D recovers the location of each corner in the grid. The two-dimensional grid coordinate is equivalent to the corner's location before forming and its three-dimensional coordinates are its location after forming. The IB3D processes the grid in small sections which are then "stitched together" or mosaiced. An in depth discussion of our approach can be found in [7].

2.1   Image Acquisition

Three high quality digital cameras (Sony DFW-SX900 [9]) are used to acquire images. Poor lighting conditions may result in degraded images (e.g. low contrast, either under exposed or over exposed, and specularities). In some cases, it may be necessary to control the lighting, however, we have obtained good results from ambient room lighting.
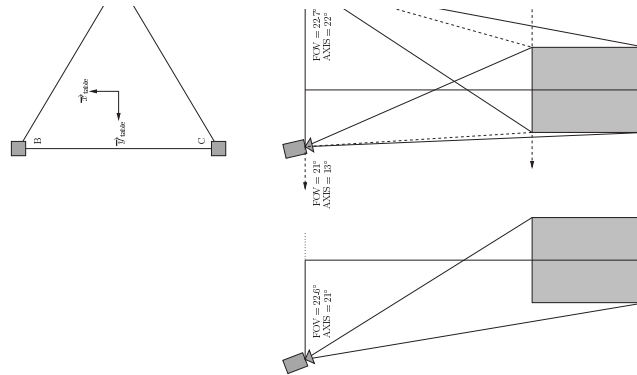
Figure 1: Camera Geometry.

The cameras are mounted in a triangular pattern about $40''$ apart and image the target from directly overhead. The desired imaging geometry is shown in Figure 1. At the top is shown an overhead view (parallel to the *xy* plane) of the imaging configuration. The center bottom shows a view parallel to the *yz* plane. The left and right bottom show views parallel to the *xz* plane. The light gray region is the working volume which must be simultaneously

visible from all 3 cameras – this volume ensures that $1' \times 1'$ tiles with $6''$ of overlap on each side can be acquired anywhere in the working volume. The minimum required field of view (FOV) is $22.7° \times 21°$. Excessive foreshortening may limit the usefulness of images. For our configuration the limiting case occurs at a slope of $35°$. If the slope exceeds $35°$ one or more cameras may not be able to properly image the steep area. In our configuration, the maximum depth (distance along the optical axis) achievable is $80.2''$. At this depth, a pixel images a $0.034''$ square on a plane perpendicular to the optical axis. Real cameras can only keep a limited range of depths in focus. For a brief introduction to lenses see [1]. The range of depths which are acceptably focused is known as the *depth of field*. In our configuration an F-stop of 4.5 or larger must be used to ensure that the entire working volume is in focus. Images with good contrast can easily be obtained with this aperture setting in a well lit room.

2.2   Camera Calibration

The pin-hole camera model used throughout this paper. Tsai's algorithm [10] is used to find the internal $(fs_x, fs_y, x_0, y_0, k_1)$, and external ($\mathscr{R}$ and $\overrightarrow{T}$) parameters which map world points to image points. Since fs_x and fs_y cannot both be separated, we set $s_y = 1.0$.

A checkerboard pattern mounted on a planar backing, is used to build a set of corresponding image and world points. Image points where 2 black and 2 white squares meet can be precisely located using a corner detector (e.g. [4]). The coordinates of the world points are determined by the physical dimensions of the checkerboard pattern and its placement in the world coordinate system. Typically one of the corners is arbitrarily chosen as the origin. The only remaining issue is matching each image point with the correct world point. The regularity of the checkerboard pattern can be exploited to automatically establish the correspondence and reduce the number and size of neighborhoods which must be explored for corners.

Calibration not only gives us information about individual cameras, but also the relationship between cameras. For example if P_1 is a point expressed in the coordinate system of camera 1, then P_2 the same point expressed in the coordinate system of camera 2 is

$$P_2 = \mathscr{R}_2\mathscr{R}_1^{-1}(P_1 - \overrightarrow{T}_1) + \overrightarrow{T}_2. \qquad (1)$$

This relationship will be used later to reconstruct the world coordinates of imaged points.

2.3   Grid Detection

A grid constructed from stripes is used to measure the deformation of a plastic sheet as a result of thermoforming. Like the checkerboard pattern used for camera calibration, there are lots of high contrast corners which can be precisely located. The stipes are printed onto the plastic prior to thermoforming and produce a distorted grid on the formed part. The distorted grid, like the checkerboard pattern above, can be automatically recovered. The recovered grid serves two important functions: 1) it identifies the original two-dimensional grid coordinates of each corner, and 2) it facilitates correspondence between cameras and between tiles. Together these two functions enable the distortion generated by thermoforming to be measured.

Unlike the checkerboard pattern, the distorted grid can have significant three-dimensional shape and even occlusions which make recovering the grid a challenge. An adaptive

threshold is used to compensate for brightness variations caused by the distorted grid's 3-D shape. The adaptive threshold uses the knowledge that the grid contains only black and white pixels and tries to find a threshold which will locally separate the two. We are interested in quadrilaterals whose corners lie on the closed contours of white squares. Quadrilateral contours are defined as contours for which the convex hull of the associated corners are quadrilaterals. Quadrilaterals which are too small or too skinny are rejected. When the plastic stretches during thermoforming the stripe orientation does not change drastically. In our experience the change in direction is less than $45°$. Therefore, using the original stripe directions can establish relative grid coordinates – if the recovered grid coordinates are $[u, v]$, then the actual coordinates are $[u + u_0, v + v_0]$.

2.4   Reconstruction

At least two images are needed to recover 3-D coordinates via triangulation [6, 2, 5]. If p_l and p_r are both images of P, then calculating the 3-D coordinates of P is simply a matter of geometry. Thus, for calibrated cameras, recovering 3-D coordinates reduces to establishing correspondences (i.e. finding the p_r that matches p_l).

Determining the grid offsets, $[u_0, v_0]$, for each image not only yields the actual grid coordinates, but also establishes the correspondences needed to recover 3-D coordinates. The grid offsets are specified interactively by selecting a corresponding grid point in each image and giving its actual grid coordinates.

Ideally, P lies along the line of sight $\overrightarrow{C_i p_i}$. Because of noise, however, this is rarely the case – particularly when P is visible in all three cameras. Assuming that the noise can be modeled as zero mean additive Gaussian noise then the best estimate of P is the one which minimizes the sum of squared errors. We use the distance (in 3-D space) between P and the line of sight through p_i as our error measure which is defined as

$$\varepsilon_i(P) = \overrightarrow{C_i p_i} \times \left((P - C_i) \times \overrightarrow{C_i p_i} \right) \cdot (P - C_i) \tag{2}$$

where $\overrightarrow{C_i p_i} = [d_x, d_y, d_z]$ is the unit vector in the direction from C_i to p_i and $p_i = [x' - x_0, y' - y_0, f]$. We use

$$\underset{P}{\operatorname{argmin}} \sum_i \varepsilon_i(P) \tag{3}$$

to find the best P. $\sum_i \varepsilon_i(P)$ has a single minimum which occurs when the gradient equals $[0, 0, 0]$ or

$$AP = B \tag{4}$$

where

$$A = \sum_i DD^{\mathrm{T}},$$

$$B = \sum_i DD^{\mathrm{T}}C_i$$

and

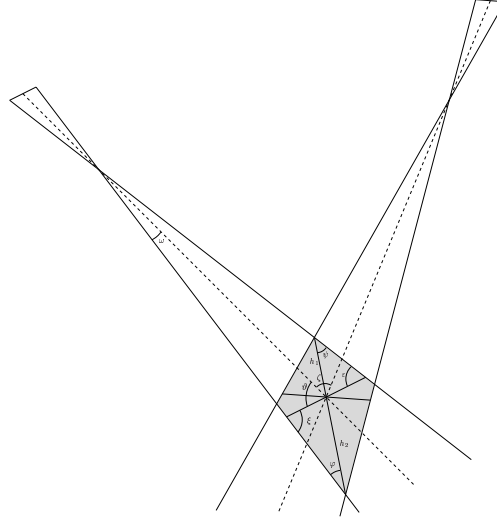$$D = \begin{bmatrix} 0 & -d_z & d_y \\ d_z & 0 & -d_x \\ -d_y & d_x & 0 \end{bmatrix}.$$

Figure 2: Intersection of Two Pixels Projected into 3-D Space.

Equation 4 is an exactly determined system of linear equations which can be easily solved. The number of cameras n which view P (in our case 2 or 3) and $\sum_i \varepsilon_i(P)/n$ provide a measure of confidence for the reconstructed points.

Figure 2 shows the intersection of two pixels projected into 3-D space. $h_1 + h_2$ is an upper bound on the uncertainty along the z axis for reconstruction using two corresponding points. For our configuration the maximum $h_1 + h_2$ is $0.15''$. Errors in the camera calibration typically increase this uncertainty by about 50% (see [10] for details). Increasing the baseline of the cameras (i.e. putting them farther apart) would reduce the vertical uncertainty, but would also make correspondence more challenging. Similarly reducing the height of the cameras would reduce the vertical uncertainty, but would also reduce the depth of field.

2.5   Mosaicing

To enable objects larger than $1' \times 1'$ to be reconstructed a gantry is used to position the cameras above the object. Individual tiles ($18'' \times 18''$ or $1' \times 1'$ with $6''$ of overlap on each side) are reconstructed and then "stitched together" or mosaiced. After mosaicing all of the reconstructed points from all of the tiles are expressed in a single global coordinate system. If the position, $\overrightarrow{T}$, and orientation, $\mathscr{R}$, of the gantry relative to the global coordinate system are perfectly known for each tile reconstructed, then points in the tile, P_{TILE}, can be converted to global coordinates using

$$P_{\text{GLOBAL}} = \mathscr{R}P_{\text{TILE}} + \overrightarrow{T}. \tag{5}$$

In practice, we only have estimates for the position and orientation of the gantry. The grid coordinates recovered above are used to correspond 3-D points recovered from different tiles. The error for grid coordinate $[i, j]$, $\varepsilon_{i,j}$, is

$$\varepsilon_{i,j} = \sum_k \kappa_{i,j,k} \left(\mathscr{R}_k P_{i,j,k} + \overrightarrow{T}_k - \overline{P}_{i,j} \right)^2 \tag{6}$$

where $\mathscr{R}_k$ and $\overrightarrow{\mathrm{T}}_k$ are the gantry rotation and translation for the kth tile, $\mathrm{P}_{i,j,k}$ is the reconstructed 3-D coordinates of grid point $[i,j]$ in the kth tile,

$$\overline{\mathrm{P}}_{i,j} = \frac{\sum_k \kappa_{i,j,k} \mathrm{P}_{i,j,k}}{\sum_k \kappa_{i,j,k}}, \tag{7}$$

and $\kappa_{i,j,k}$ is 1 if tile k contains grid point $[i,j]$ and 0 otherwise. Equation 6 leads to the straightforward minimization

$$\underset{\{\mathscr{R}\},\{\overrightarrow{\mathrm{T}}\}}{\arg\min} \sum_{i,j} \varepsilon_{i,j} \tag{8}$$

where $\{\mathscr{R}\}$ is the set of all $\mathscr{R}_k$'s except $\mathscr{R}_0$ and $\{\overrightarrow{\mathrm{T}}\}$ is the set of all $\overrightarrow{\mathrm{T}}_k$'s except $\overrightarrow{\mathrm{T}}_0$. $\mathscr{R}_0$ is fixed as the identity matrix and $\overrightarrow{\mathrm{T}}_0$ as $[0,0,0]$. The optimized $\mathscr{R}$'s and $\overrightarrow{\mathrm{T}}$'s enable the reconstructed tile points to be placed into a single global coordinate frame. Equation 8 is a good candidate for optimization techniques [8] such as Powell's method.

3   Results

A $32.5'' \times 42.5''$ grid pattern with stripes $0.5''$ wide printed on paper was to test the IB3D. The paper grid was draped over several boxes to produce a 3-D surface. Figure 3 shows the complete set of images acquired by the IB3D to reconstruct the grid points (both 2-D grid coordinates and 3-D positions). Each row is from a single camera. The reconstructed surface is shown in Figure 4. The printed grid contains a total of 5376 (84×64) grid points. Not all of the rows are visible in 2 or more cameras. For example, the bottom 8 rows (4 black stripes) are not completely visible in second and third cameras (see Figure 3). Over 99% of the visible grid points, a total of 4688, were recovered. The reconstruction certainly looks good, but because we don't have ground truth values for the 3-D position of the grid points, evaluating it is challenging. The paper grid we used, unlike thermoformed plastic, did not stretch while forming a 3-D surface, therefore adjacent grid points will be $0.5''$ apart along the surface. If the surface curvature is not too large, the Euclidean distance between adjacent grid points will be very close to $0.5''$. The absolute value of the difference of the distance between adjacent grid points and $0.5''$ is used as an error measure to evaluate the quality of the reconstruction. Figure 5 shows the fraction of points below a given error measure. 98.9% of the points are below the design goal of $0.050''$ accuracy.

The IB3D has also been used to produce a large pre-decorated 3-D thermoformed part for outdoor advertising. A $9' \times 14'$ milkshake to be placed on a roadside billboard was designed for *Steak 'n Shake*, Inc. A grid with $0.5''$ stripes was printed on a large piece of plastic and thermoformed into the final shape as shown in Figure 6. The major contours are highlighted with red dots. These dots were not present during the digitization. Notice the significant distortion of the grid and the specular reflections. The milkshake contained a total of 71,318 grid points (211×338)of which over 85%, or 60,693, were reconstructed with high confidence. The reconstructed milkshake is shown in Figures 7. The holes indicate places where the grid points were not successfully reconstructed. The two major causes were specular reflections and excessively foreshortening or occlusion both of interfere with recovering a connected grid structure. As with the previous testing
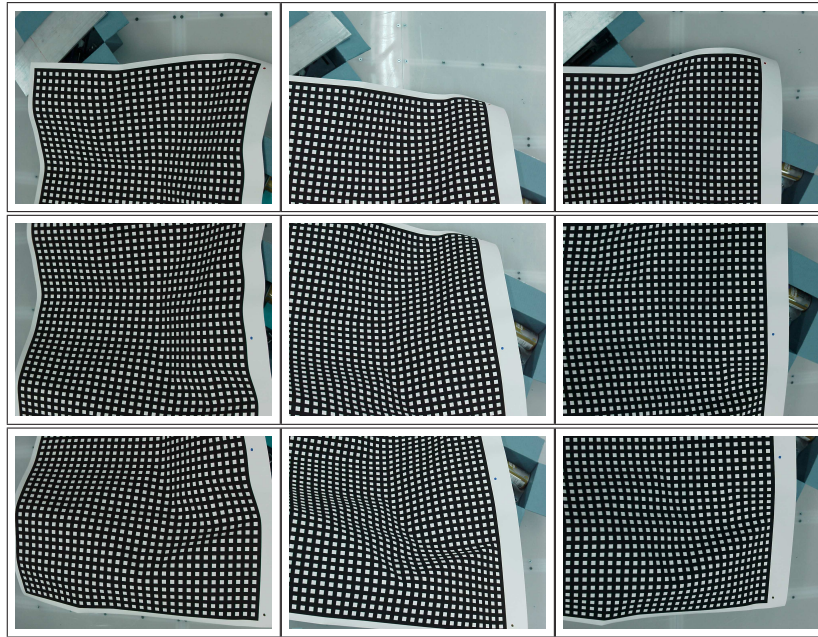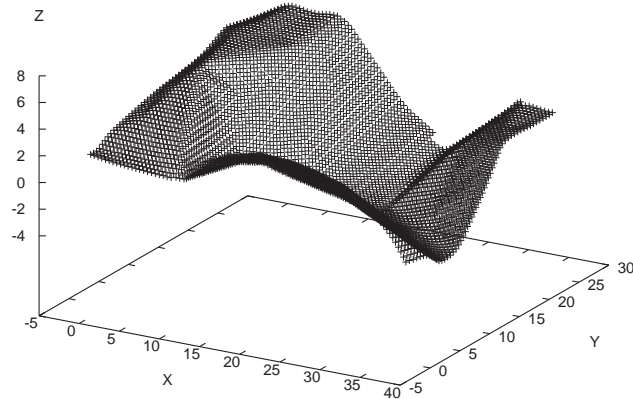
Figure 3: Images of the Test Grid.

Figure 4: Reconstructed Surface of Test Grid.

we do not have ground truth values for the 3-D position of the grid points. However, after fixing the holes by hand, the reconstruction produced by the IB3D was sufficiently accurate to produce the final decorated version of the milkshake. The finished billboard is shown in Figure 8.

4   Conclusions

This paper describes a novel method for recovering the pre-distortion necessary for pre-decorating thermoformed parts which have complex shapes and/or significant relief. We
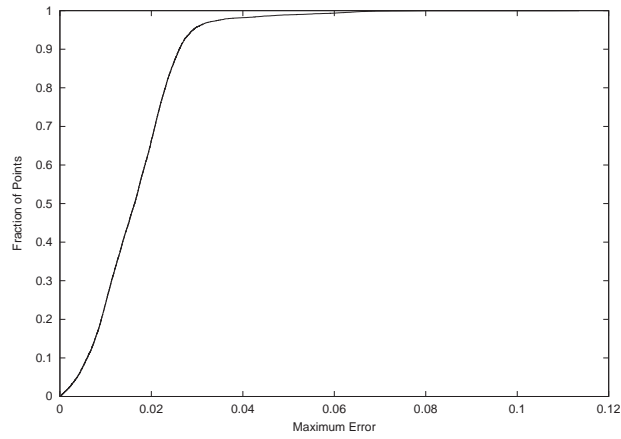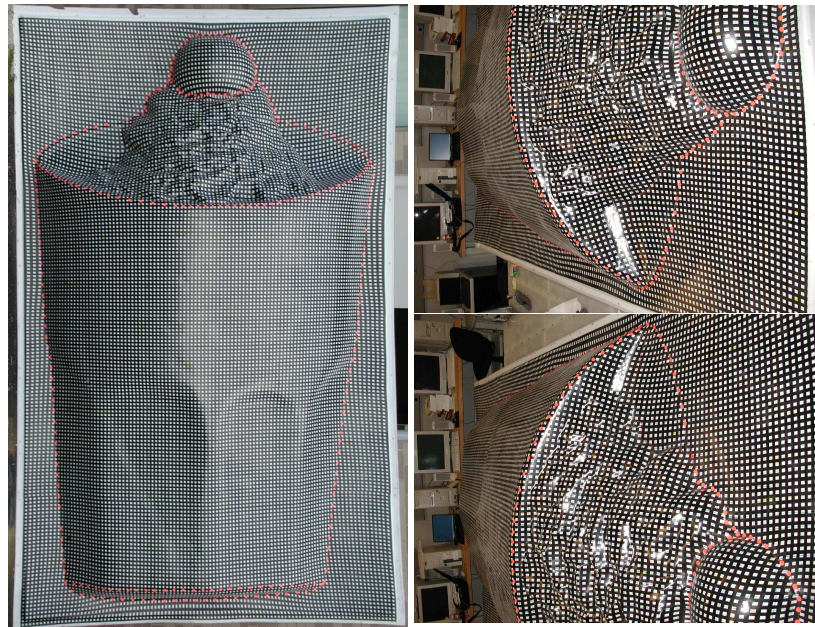
Figure 5: Test Grid Reconstruction Error.

Figure 6: Milkshake With Grid.

also describe a device, the *Image-Based Three-Dimensional Digitizer* (IB3D), which is capable of efficiently and cost effectively recovering this pre-distortion. The IB3D uses a test part which is pre-decorated with a striped grid. During thermoforming the plastic stretches distorting the grid. The IB3D recovers both the 2-D grid coordinates and the 3-D positions of each grid point. This data leads directly to the pre-distortion which enables thermoformed parts to be pre-decorated.
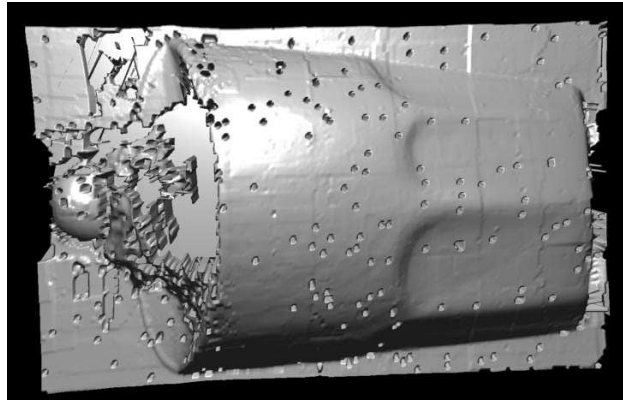
Figure 7: Reconstructed Milkshake Surface.

Figure 8: Finished Billboard.

4.1 Future Work

While the IB3D has been used successfully to produce large 3-D thermoformed objects for outdoor advertising, it is essentially a proof of concept. The basic approach is sound and produces good results, however there is room for improvement in all phases of operation.

4.1.1 Image Acquisition

Specularities and foreshortening are the two major challenges for image acquisition. Dulling spray has been successful in reducing specularities, but may not be sufficient

in all cases. Dulling spray must be applied to the part after it is formed. Other methods of eliminating specularities include controlled lighting and polarizing filters. Controlled lighting, by providing more uniform illumination, may also improve grid detection.

The positioning system gantry allows only limited control over the imaging geometry and as a result some surfaces may be excessively foreshortened. Better control of the imaging geometry would improve both grid detection and the reconstruction accuracy. One approach is to allow the gantry and/or the part to tilt. Another possibility is to mount the camera platform on the end of a digitizing arm, allowing extensive control of imaging geometry. The digitizing arm could also be used to manually digitize points which cannot be properly imaged (e.g. deep channels).

4.1.2   Grid Detection

The grid detector identifies quads and then links them together into a grid. If a quad is not detected it is simply omitted. If adjacent quads are detected, their positions can be used to estimate the locations of missing quads. Occasionally the grid detector connects grid points across occlusion boundaries. Omitting the occluded grid points causes the grid point numbering to be incorrect. Dense stereo techniques can be used to detect occlusion boundaries and prevent incorrect grid links.

4.1.3   Reconstruction

Reconstruction requires that one correspondence per tile be established interactively. It is possible to establish correspondences automatically. One possibility is to start at a corner of the grid and track grid points as the tiles are collected. Another possibility is to use a laser pointer to identify corresponding points. Typically, a 3-D CAD model is used to fabricate a thermoform mold. The 3-D data in this model, even though it describes the mold and not the thermoformed part, can be used as an additional confidence measure.

References

[1]   M. Bass, E. w. Van Stryland, D. R. Williams, and W. L. Wolfe, editors. *Handbook of Optics*. McGraw-Hill Professiona, 1994.

[2]   Olivier Faugeras. *Three-Dimensional Computer Vision*. MIT Press, Cambridge, MA, 1993.

[3]   Charles A. Harper, editor. *Modern Plastics Handbook*. McGraw-Hill Professional, 2000.

[4]   C. J. Harris and M. Stephens. A combined corner and edge detector. In *Proceedings of the 4th Alvey Vision Conference.*, pages 147–151, 1988. Manchester.

[5]   R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2000.

[6]   Berthold Klaus Paul Horn. *Robot Vision*. MIT Press, Cambridge, MA, 1986.

[7]   J.P. Mellor, M. Oder, J. Starr, and J. Meador. An image-based three-dimensional digitizer for pre-decorating thermoformed parts. Technical Report RHV TR-1, Rose-Hulman Ventures, April 2002.

[8]   William H. Press, Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery. *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge University Press, 1992.

[9]   Sony. DFW-SX900 technical manual, 2001. http://theimagingsource.com/prod/cam/dfwsx900/dfwsx900_ug.pdf.

[10] Roger Y. Tsai. A versatile camera calibration technique for high-accuracy three dimensional machine vision metrology using off-the-shelf TV cameras and lenses. *IEEE Journal of Robotics and Automation*, RA-3(4):323–344, August 1987.