(139)

Real-Time Close-Range 3-D
Non-Contact Volume and Motion Measurements

R. Wilson and F. Leberl

Vexcel Corporation
2477 55th Street
Boulder, Colorado, USA 80301

Abstract

Medical and industrial needs exist to track the position, volume and geometry of moving objects whose shapes change. Video imaging offers frame rates of sufficient frequency to support real-time sensing of fairly rapid changes in the real world. It is the processing of the video frames which represents the bottleneck in a "real-time" implementation.

However, under certain limiting constraints, namely near-symmetry around a curved axis, real-time measurements are feasible with standard PC-based software to obtain 3-D positions, shape and volume of a moving and changing 3-D object.

The paper describes a novel stereo-photogrammetric analysis system based on object silhouettes. Implementation involves the end-to-end system from video cameras to user interface and reporting of results. A medical application's scenario puts high demands on robustness and simplicity of the user interface, while operating in an entirely automated fashion.

"Real-time" represents the need to process an image pair each 2 seconds. Accuracies of ±1% in dimension and 3% in volume are required and permit one to accept certain shortcuts in the photogrammetric approach.

Introduction

Vexcel has developed a computer-video system to make non-contact spatial measurements on objects which change size and/or shape. Objects are modeled as having a nearly circular cross section about a straight or curved central axis, with a diameter free to vary along the length of the central axis. The objects are thus assumed to be delimited by bent surfaces of revolution. While many familiar objects (e.g., growing zucchini squash, inflating balloons, etc.) may reasonably be approximated by this model, the motivation for developing the system was to provide medical instrumentation to be used by a urologist in the diagnosis of human erectile dysfunction (impotence).

The system, called **Priapus** after the Greco-Roman god of procreation, was designed at Vexcel in cooperation with the Urological Clinic at the Medical University, Hannover, and F.M. Wiest GmbH and Company in Munich, FR Germany. Priapus employs two digital CCD cameras to observe the object under investigation. The observations provide non-contact, real-time measurements of object length, diameter, volume, curvature and angle. Real-time performance is achieved by use of single-image morphological and photogrammetric processing and a specialized approach to 3-D reconstruction of the object.

System Requirements

A variety of computer-video systems have been developed to digitize images of discrete targets (LEDs or retro-reflective markers), and track them as they move through 3-D space.[1] [2] However, the physicians for whom the system was designed ruled out placing artificial landmarks on the subject or marking the object in any way.

Although structured lighting (e.g., a projected grid or pattern of random dots) might be medically acceptable, lighting compatible with normal office lighting was required. However, the doctors did not object to the patient's wearing a special cloth drape during the measurement process, thereby providing a simple background against which the object could be imaged. Because an observation session might require processing many frames of data, the measurement process had to be fully automatic. This ruled out relying on user interaction to determine corresponding points in the paired images. Consequently, we chose to extract all necessary image information from simple silhouettes of the object.

Another goal of the development was to make calibration of the two cameras as simple and as flexible as possible. Camera calibration entails determining for each camera the perspective transformation which maps locations in 3-D object space into the camera's 2-D image plane. Assuming that the lens acts as a simple central projector (i.e., ignoring distortion of the lens), the camera's perspective depends upon its exterior orientation (location and 3-D orientation) and interior orientation (location of principal point and principal distance).[3] But standard video cameras are not *metric* cameras, and their interior orientation is not generally known to the videographer. Moreover, a calibration procedure which requires direct measurement of camera parameters is quite cumbersome, unless the cameras are firmly fixed in space and their lenses rendered unfocusable. Especially for a prototype system, it should be possible to place the cameras almost anywhere, yet not be required to precisely survey their locations, orientations, etc. If the locations of discrete targets in 3-D space have been carefully surveyed, and their projected locations onto the image planes of the cameras are also known, one should be able to solve for the perspective transformations of the cameras. The calibration software must provide the user adequate feedback to ascertain whether accurate calibration has in fact been achieved. Targets must be localized in the image with sub-pixel accuracy in order to determine their object space locations to within ±0.5 mm.

The system had to yield measurements of *length, diameter, volume, curvature* and *angle*, but the accuracy requirements of the system were not clear. The doctors felt that accuracies of ±1% length and diameter and ±3% in volume regardless of the attitude of the object would be more than adequate for their application. Thus we could afford to make certain simplifying assumptions in order to accelerate system performance.

Medical studies indicated that a frame rate of 1 stereo frame (a stereo pair of images) every 15 seconds over a period of time lasting up to one hour (240 frames) was adequate to describe the morphological changes under investigation.[4] This generous period of time, 15 seconds to capture and process a stereo pair of images, facilitates performing all computation steps in *real-time*. Our goal was to eventually pare the required time down to a few seconds, which is more appropriate for possible industrial applications.

Since the system was to be used by people without extensive experience with computers, a window-based, mouse-driven user interface was required.

The system hardware was required to consist largely of off-the-shelf items: standard monochrome video cameras, a frame-grabber/image processing board and moderate-performance graphics workstation.

System Hardware

The system is implemented on a Sun 386i workstation with a Matrox MVP-ATi video processor. It also uses two Pulnix TM-840 high-resolution video cameras supported by a camera stand with fluorescent lamps, a Video Junction Box (containing the camera power supply and connections enabling the two cameras to be genlocked), a video monitor and a custom-made calibration rig. The calibration rig is cubical in shape, with an edge length of 30 cm. It consists of two 30 by 30 cm square aluminum plates connected by three aluminum rods. Wires are strung between the plates, and support 20 white plastic beads 12 mm in diameter. These are the calibration targets.

System Software

In its current state of development, the system software consists of the following seven Unix shell-level commands:

Name	Description
calibrate	interactively calibrate cameras
measure	calculate measurements
measuretool	interactive front-end for measure
outline	repetitively capture outlines
patient	interactive patient database manager
show	graphically display outlines and measurements
threshold	interactively set video threshold

The software is written in the C programming language. The interactive, mouse-driven, iconic user interface is implemented in SunView. All Priapus data files consist of ASCII text.

Data Capture and Image Processing

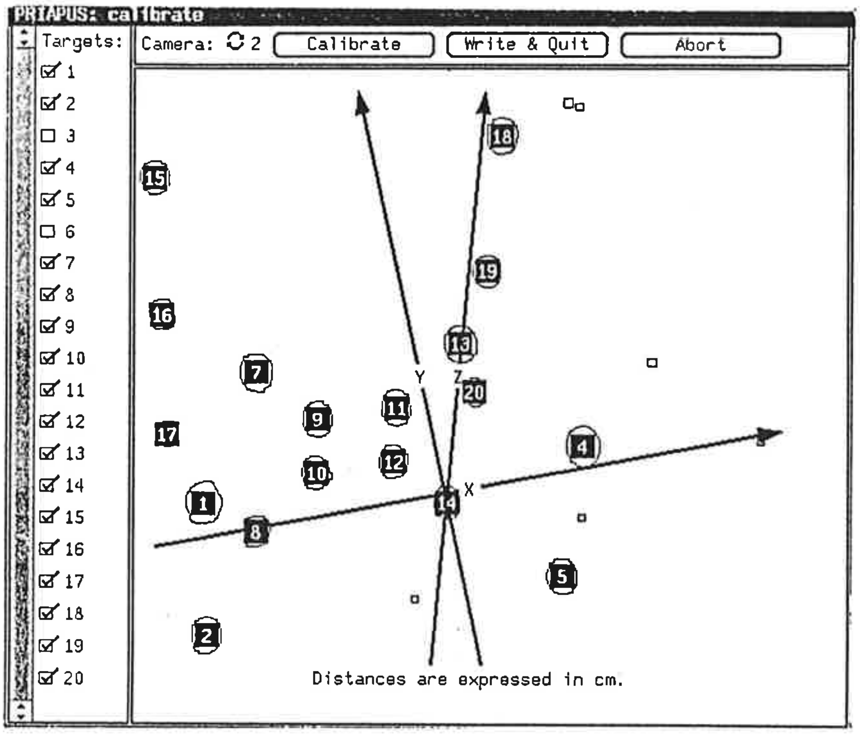
Both for calibration of the cameras and performing measurements, paired images are captured and processed to yield outlines. The video cameras are synchronized by connecting the sync output of one camera (the master) to the sync input of the other (the slave) through the Video Junction Box. The image of one camera is digitized by the Matrox video processor to yield a 512 by 480 pixel 8-bit deep raster image. In the subsequent 1/30th of a second, the other camera's image is similarly digitized. The video processor thresholds each image to produce binary images. A separate video threshold is supported for each camera. The resulting binary "blobs" are eroded by one pixel to eliminate background speckle, and then dilated by one pixel to restore their original size. The images are segmented and outlines are vectorized by a routine which crawls along the perimeters of the blobs. Blobs with open outlines are excluded from subsequent analysis. Centroids (average horizontal and vertical image coordinates) are computed for those with closed outlines.

Objects for which we have captured data are generally light in color and imaged against black backgrounds. Accurate outlines are gathered only if portions lying along the silhouette (the limbs) of the object are adequately illuminated and the video thresholds are carefully adjusted. The *threshold* command enables the user generate snapshots of objects on the video monitor while interactively adjusting the thresholds by manipulating graphical sliders using the mouse. This allows the user to confirm that the cameras are properly positioned and adjusted and to select thresholds which yield crisp binary images.

Data capture and generation of outlines may be performed by the *outline* command. It writes an ASCII file of object outlines, which may be used as input to the *calibrate* and *measure* commands. Calibration only requires one image for each camera. For performing measurements, images are repeatedly captured and processed using a user-specified interval between frames (e.g., 15 seconds) and a user-specified number of frames (e.g., 120) or duration (e.g., 30 minutes). For real-time measurement, no outline file is produced and the *outline* command is not invoked. The *measure* command executes the routines to capture and process images, and the measurement routines receive outline data structures as input.

Calibration

Calibration of the cameras and subsequent stereo intersection is based on the the *Direct Linear Transform*, or DLT.[5] Important advantages of the method are that it does not require metric cameras and does not require that initial approximations be supplied for the camera parameters. If lens distortions are neglected, the DLT requires calculation of eleven coefficients which embody the perspective of each camera. A minimum of six non-coplanar targets (or control points) must be imaged by each camera and their locations in 3-D object space must be specified. Given their known positions in 3-D space and the positions of their images (central projections) in 2-D space, a least squares solution for each camera's calibration coefficients is obtained without iteration. This enables more accurate estimation of the calibration coefficients when additional (redundant) control points are used.



Targets:
☑ 1
☑ 2
☐ 3
☑ 4
☑ 5
☐ 6
☑ 7
☑ 8
☑ 9
☑ 10
☑ 11
☑ 12
☑ 13
☑ 14
☑ 15
☑ 16
☑ 17
☑ 18
☑ 19
☑ 20

Distances are expressed in cm.

Figure 1. Window displaying user-identified control points superimposed on an image of the calibration rig. After the camera is calibrated, the origin and axes of the object coordinate system are depicted.

The calibration rig establishes the object coordinate system for calibration. The centers of the 20 targets were surveyed by means of a numerical control machine to within ±0.1 mm. These coordinates are stored in an ASCII control point file which is read by the *calibrate* command. Calibration also requires outlines of targets viewed by each camera. They are captured using the *outline* command, which produces an outline file for each camera.

calibrate allows the user to interactively identify targets in each camera's image. This is illustrated in Figure 1. Each control point is numbered. The user identifies a point by using the mouse to click on its number and then clicking on its outline. The target is then "checked" to show that it has been identified. Alternatively, the user can first click on an outline and then the number. A mis-identification is removed by double-clicking on it. When enough points have been identified, the user can *calibrate* the camera. After solving for the calibration coefficients, the camera parameters are displayed and statistics are computed to provide feedback concerning accuracy. Using the newly computed coefficients, the known locations of the control points in object space are projected into the image. Standard deviations of the horizontal and vertical differences between projected and observed locations in the images are calculated. Disparities which exceed 0.5 pixels indicate a possible mis-identification of one or more control points.

After cameras have been calibrated, further statistics are generated. Using coefficients from both cameras, the observed image coordinates are intersected in object space. Standard deviations of the x, y and z differences between intersected and known control point locations are calculated. Disparities which exceed 0.5 mm indicate problems with the calibration. 0.5 mm represents 1 part in 600 within the 30 by 30 by 30 cm calibrated region. With cameras located roughly 2 meters from the origin in object space, this is 1:4000 of the camera-to-object distance.

Measurement

The *measure* command requires calibration data and outlines for each camera. It is normally invoked by *measuretool*, which provides an interactive user interface for specifying input files and command options. *measuretool* also enables the user to specify the location of the *base* of the object in each camera's view. The software distinguishes the base of the object from its tip by tracking the base frame-by-frame.

Once the cameras have been calibrated, the geometric requirements for locating a point in 3-D space are fairly straightforward. Given images from two different cameras, and a point in each image which is known to correspond to the same target, we know (assuming the camera to be a central projector) that each 2-D point denotes a ray in 3-D object space. The target is located where the rays (almost) intersect. However, all Priapus measurements are based on silhouettes in the views of two cameras with dissimilar vantage points. It is possible that no point along one silhouette corresponds (i.e., is homologous) to a point on the other silhouette. This is the primary challenge which our measurement software must overcome.

Our technique begins by computing a 2-D axis or centerline for each silhouette. The locations of the base and tip are found by representing the outline in polar coordinates relative to its centroid. The series of polar radii are iteratively low-pass filtered until two clear polar maxima are found. If such maxima are not discovered, it is possible that the object is viewed "head on" or that an image of the object under investigation has not been properly captured. (The software must treat these as special cases.) Given locations for the base and tip, a quadratic is fit to the base, tip and centroid. This function is evaluated at 40 points. At each of these points a line normal to the function is constructed and a point is chosen midway between the two points where the normal slices through the outline. A quadratic is then fit to the new set of axial points. The process is repeated until the amount of adjustment is less than a predetermined threshold. It is assumed that corresponding points lie along the 2-D centerlines in the two images. The problem is how to select the points that match.

Each resulting 2-D centerline extends from one end of the silhouette to the other. If it is to represent an idealized 3-D centerline in the object, its ends must be trimmed in order to take perspective into account (unless, of course, we have a perfect side view of the object). We have experimented with an algorithm which requires generation of an initial 3-D centerline by performing stereo intersection for points equispaced along the two 2-D centerlines. This is an initial approximation. The points are not genuine homologs because of the poorly defined base and tip locations and due to foreshortening. Nevertheless, the angles of the proximal and distal ends of the object relative to the gaze of the camera are calculated and used to trim the ends by amounts proportional to the cosines of the angles. The trimmed 2-D centerlines are again subjected to stereo intersection (again using proportional matching) to yield a final 3-D centerline.

Once a final 3-D centerline has been computed, the measurements are calculated. The *length* is computed by summing the lengths of all segments along the 3-D centerline. *Diameter* is calculated by performing a simple geometric construction for each camera. For each point along the 2-D centerline which was used for stereo intersection, a plane is constructed which contains the corresponding point on the 3-D centerline and is orthogonal to the ray extending from the perspective center to the object point. We assume that the local portion of the silhouette may be approximated as lying within this plane. The pair of points in the image outline closest to the 2-D centerline is determined. The corresponding silhouette points are located in object space by intersecting rays extending from the image outline points and through the perspective center with the plane. This yields two estimates of object radius at a point along the objects axis for each camera. *Volume* is calculated summing the volumes of cylindrical slices (of known length and diameter) along the length of the object. The sum of the angles of contiguous segments relative to their neighbors is calculated and divided by the length to yield a measure of *curvature* (expressed in degrees/mm). The attitude of the object is expressed in spherical coordinates in order to quantify *angle*. These quantities are written frame-by-frame to an ASCII measurement file. The data may be graphically displayed using the *show* command.

System Performance

We have tested *measure* both by making measurements on physical models and on simulated objects generated in software. As of this writing, we still have not attained our originally stated goals for accurate mensuration (± 1% for length and diameter and ±3% for volume). In a recent test we positioned the video cameras to look down on a model object from vantage points which afforded roughly orthogonal views. The cameras were calibrated. Then successive outlines of a model object were captured, with the object being rotated in the horizontal plane. On the average, Priapus underestimated length by 4.5%, diameter by 2.5% and volume by 7%. The coefficients of variation were 2.8% for length, 2.9% for diameter and 3.8% for volume. The variations in the computed measurements were (largely) deterministic functions of angle. Less favorable viewing conditions (e.g., an almost "head on" view in one of the images) causes yet further degradation in accuracy. The measurement process currently requires 13 seconds per stereo pair.

Current System Development

Based on our simulations, we attribute the current inaccuracies of *measure* to errors in correspondence of points matched along the 2-D centerlines and used in stereo intersection. As of this writing, we are currently developing software to more precisely determine the base and tip of the object in 2-D image space by fitting ellipses of constrained major axis and orientation to the end regions of the silhouette. The trimmed 2-D centerlines will then be subjected to proportional matching to yield an initial 3-D centerline. Points equispaced along the initial 3D centerline will be chosen and projected back into the image planes of the cameras. Points will be selected along the 2-D centerlines which are nearest to the projected points. These points will be treated as homologs and used to generate a final 3-D centerline via stereo intersection, thereby matching image points in a manner consistent with foreshortening due to perspective.

References Cited

1. Maurice, K., F. Leberl, S. Curry, and W. Kober, "Real-time close-range 3-D motion measurements for dental medicine," in *Close-Range Photogrammetry Meets Machine Vision*, Proceedings of SPIE, Vol. 1374. 1990..

2. R. Wilson, "Software for automatic tracking of moving objects in three dimensions," in *High Speed Photography, Videography, and Photonics IV*, B.G. Ponseggi (Ed.), Proceedings of SPIE, Vol 693, pp. 269-276. 1986.

3. Slama, C., C. Theurer, and S. Henrikson (Eds.), *Manual of Photogrammetry*, Fourth Edition, American Society of Photogrammetry, Falls Church, VA. 1980.

4. Bruins, J.L., A.E.J.L. Kramer, and U. Jonas, "RigiScan penile rigidity and tumescence monitoring in impotent patients as a diagnostic tool and as an objective assessment of the effect of intracorporeal injection of papaverine hydrochloride," in *Investigative Urology 2*, G.H. Jacobi *et. al.* (Eds.), Springer-Verlag, Berlin, pp 153-162. 1987.
5. McGlone, J.C., E.M. Mikhail, and F.C. Paderes, "Analytic data-reduction schemes in non-topographic photogrammetry," in *Non-Topographic Photogrammetry*, Second Edition, H.M. Karara (Ed.), American Society for Remote Sensing, Falls Church, VA, pp. 37-57. 1989.

(138)

ISSN 0256-1840

ISPRS

ISPRS Commission V Symposium

Close-Range Photogrammetry Meets Machine Vision

A. Gruen, E. Baltsavias
Editors

3–7 September 1990
Zurich, Switzerland

Organised by

ISPRS International Society for Photogrammetry and Remote Sensing (ISPRS)
Commission V "Close-Range Photogrammetry and Machine Vision"
ETH Institute of Geodesy and Photogrammetry, Swiss Federal Institute of Technology (ETH) Zurich

Sponsored by

SGPBF	Swiss Society for Photogrammetry, Image Analysis and Remote Sensing

Cooperating Organisations

SPIE	The International Society for Optical Engineering
IEEE	The Computer Society, TC on Pattern Analysis and Machine Intelligence
FIG	Fédération Internationale des Géomètres
ITG	Information Technology Society of the SEV
SGBT	Swiss Association of Biomedical Engineering

Published by
SPIE—The International Society for Optical Engineering
P.O. Box 10, Bellingham, Washington 98227-0010 USA
Telephone 206/676-3290 (Pacific Time) • Telex 46-7053

SPIE PROCEEDINGS SERIES

SPIE Volume 1395
Part One of Two Parts