The Recovery of 3-D Facial Shape for Analysis in the Pre- and Post-Operative Phases of Reconstructive Surgery by Hand-Actuated Light-Striping

P. Butler[1], E. O'Broin[2], and D. Vernon[1]
[1]Department of Computer Science, Maynooth College, Co. Kildare, Ireland
[2]Temple Street Children's Hospital, Temple Street, Dublin, Ireland

Telephone: 708 3847
Fax. 708 3848

ABSTRACT

Laser-based light-striping techniques are routinely employed to compute the surface profile or shape of objects, including the human face. Such systems exploit either motorised rotational or translations fixturing in order to sample the surface in a uniform manner, extracting the translation or the angle of rotation from encoders. In particular, face or head profiling systems depend on the rotation of the scanner about an vertical axis centred in the subject (or, alternatively, having the subject itself rotate about a similar axis). These devices are bulky and do not lend themselves to use in remote clinical trials where portability is mandatory. This paper describes an alternative configuration which only requires manual actuation by a clinician and obviates the need for any bulky motorised fixtures. The rotation angle is extracted instead directly from the image and the resultant set of non-uniformly sampled surface profiles, which are collected in a cylindrical co-ordinate reference frame are subsequently transformed to a 3-D Cartesian frame of reference and the final surface is computed using Lagrangian interpolation. Work in progress to effect face registration and face characterisation as a replacement of conventional anthropometric techniques is described.

1. Introduction

Medical analysis in general involves the collation and subsequent assessment of information accumulated from a wide range of sources. However, the structural information required in reconstructive surgery contrasts sharply with the information required in traditional medicine in that it is primarily concerned with the geometric structure of the body rather than its physiology.

The goal of the work described in this paper is to provide plastic surgeons with good reliable data to assess and prepare for surgery. Young patients' faces alter with growth and thus old and new data must be available to take into consideration such changes when on-going treatment is being carried out. However prior to such analysis a faithful reproduction of the human face must be developed. Ideally such a system should be portable to aid collection of large amounts of face data in remote clinical trials and to promote cost effectiveness. The use intended for such a portable system is for the 3-D reconstruction of face data of young children suffering such common malformations as cleft pallets thus placing major emphasis on the lower region of the face. The implications of this is that the system must not be intrusive. Also, acquisition time must be minimal and the reconstruction technique must be as robust as possible. The face data acquisition system described is hand-actuated and thus requires manipulation of non-uniformly sampled data. A technique developed to accommodate this will be described.

The organisation of this paper is as follows. Section 2 deals with the components of the system along with a brief description of its construction. Section 3 describes the calibration of the system, image acquisition considerations and processing issues. Section 4 covers current research with respect to the design of the system and sheds some light on how the data may be utilised as a means of uniquely representing a face. Finally section 5 presents some conclusions about the system and its feasibility.

2 Scanner Configuration

The imaging system is set-up to incorporate a CCD camera and a laser light stripe source[1][2]. These basic components are fixed and adjacent to one another on a rotational construction such that they are at a constant distance from one another and are free to rotate about a controlled axis, see Figure 1. The light source and camera are configured in a manner that allows the plane of light emanating from the laser to intersect the principle axis of the camera at this pivot. The system is free to rotate about the axis without affecting the orientation or relative distances of the 2 components. The camera must be a sufficient distance from the

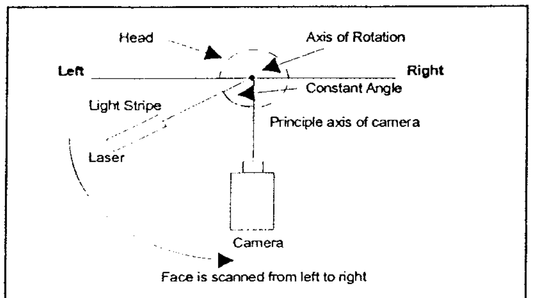

Figure 1: Schematic of scanner configuration.

$\Rightarrow pos\_scale = \dfrac{c}{C}$ pixels per millimetre.

3.4 *Rotation angle*

Since the face is non-uniformly scanned in a cylindrical fashion[7], it is necessary to determine the angular position at which each contour image is acquired. It is possible to calculate the angle of rotation of the system using the image data alone. Two horizontally collinear LEDs are placed above the object at either side of the axis of rotation of the system in the field of view of the camera as in figure 4. In the image frame of reference the apparent distance between these LEDs changes as the system is rotated about them. There is a relationship between this maximum distance, the changing distances due to rotation, and the angle of rotation of the system. The maximum distance between the LEDs is considered to be the maximum distance in a sequence of images. The rotation angle is said to be negative when the successive LED distances are increasing. It is said to be zero degrees when this apparent distance reaches a maximum and positive when LED distances are decreasing. A brief outline of the angle calculation follows.

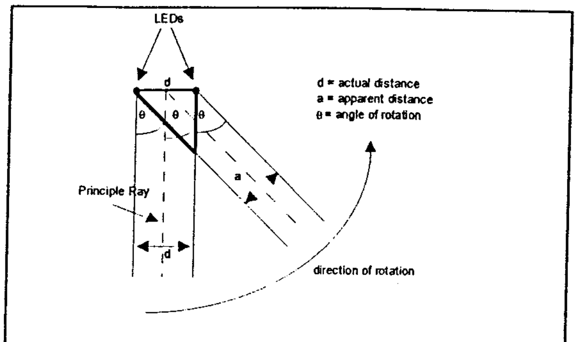

Figure 3: Calculation of angle θ.

Acquire images at different θ_i, where θ_i is the rotation angle w.r.t. image *i*.
Compute θ_i from LED distance.

Let apparent distance between LEDs due to rotation $= a_i$ pixels
Let actual distance between LEDs $= d$ pixels

$\Rightarrow \theta_i = \cos^{-1}\left(\dfrac{a_i}{d}\right)$ see figure 3.

Note that θ is the angular position of the stripe with respect to the camera. We are concerned with the angular position of the stripe with respect to the light source. Thus we must calculate the angle ϕ between the camera and the light source. Knowing the distance between the camera and the light source (A), the distance between the camera and the axis of rotation (B), along with the distance between the light source and the axis of rotation, (C) we can deduce ϕ using the cosine rule.
Thus the angular position of the stripe in image *i* is thus $\theta = (\theta_i - \phi)$.

3.5 *Subsequent Processing*

For each subsequent thresholded contour image *i*, the skeleton of the stripe is computed. This is achieved by scanning through the image row by row. When an object pixel is encountered its horizontal image location is recorded. The centroid of the object pixels along a row is then computed. This value is recorded and the object pixels in that row are replaced with one object pixel corresponding to the horizontal centroid rounded to the nearest pixel. This is repeated for all rows of the image. Let the skeletonised image be denoted $S(p,q)$.

For each point (pixel) on $S(p,q)$, the corresponding cylindrical height is calculated. Let p be the horizontal position of the pixel.

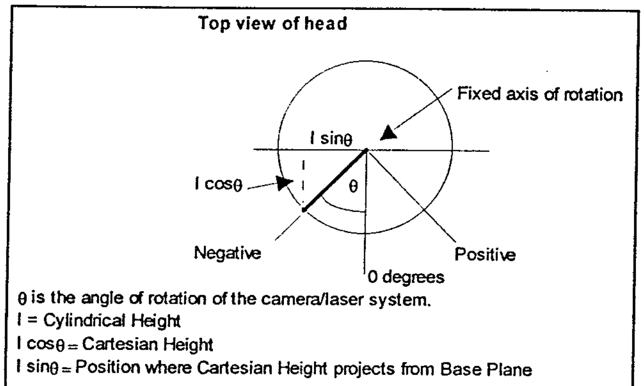

Figure 4: Cartesian height calculation

Thus,

$(p - \omega_b) \times heightscale = l$ the cylindrical height of point (p,q) in millimetres.

proposed object (face) in order to accommodate a large depth of field and to enforce the certainty that the object is entirely within the field of view of the camera. The object must be placed in a position such that the axis of rotation of the system is at a distance which allows the deepest depression in the object to project beyond it. This allows a single scan to encompass a sufficient amount of the face surface. The system is set-up to rotate about the head which immediately eliminates the requirement for an extra component such as a motorised rotating chair[3][4] and thus contributes to the portability of the system. The rotational motion of the system is smooth, allowing minimal interference from other spurious movements.

3 Face Reconstruction

3.1 *Image Acquisition*

The most important point to note in an application of this kind is the quality of the image data acquired. In this case we are using a laser which when striking the surface of an object produces a speckle effect. Speckle refers to the granular appearance of reflected light from a rough surface. A rough surface in this context implies one with random microscopic surface variations whose scale is fractionally greater than or less than the wavelength of the incident laser light. The deflection of this light off such a surface gives rise to random constructive and destructive interference thus resulting in such an effect[5] . This can give rise to an image where the stripe isn't solid. This is a disadvantage in that it interferes with segmentation and subsequent thinning of the stripe. Although this speckle effect can be modelled, it is countered here by opening the aperture of the camera slightly thus giving rise to slight over-exposure resulting in an image of a solid stripe.

3.2 *Local Object Height*

The local height measurement technique requires a simple calibration step using an object consisting of two parallel planes. The orthogonal empirically measured distance between the 2 planes is known. The calibration object is placed in the path of the laser beam in a position such that the plane of light strikes both planes orthogonally as shown in Figure 2. An image of the stripe is acquired under minimal illumination conditions (to promote segmentation of the stripe). Using this calibration technique we can determine the image distance between these two planes (stripe offset) from the camera's point of view.

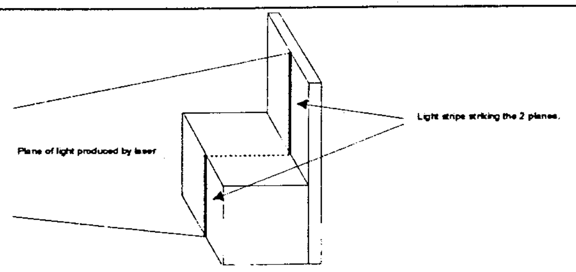

Figure 2: Light stripe strikes calibration object orthogonally to produce an image consisting of 2 disjoint vertically offset lines.

The image histogram is bi-modal due to the illumination constraints and the threshold lies between these two modes[6]. The image is segmented on the basis of this threshold value.

Let the base line of the calibration image $= \omega_b$

Let the top line of the calibration image $= \omega_t$

Let the offset between the 2 planes of the calibration object $= h$

Thus, stripe offset $\delta = (\omega_b - \omega_t)$

$\Rightarrow heightscale = \frac{h}{\delta}$ units: millimetres per pixel.

3.3 *Calibration for Horizontal placement of Processed Stripe*

This step is needed to determine the correct position of each processed stripe in an accumulator image. The first thing to note is that this horizontal scaling matches the vertical scaling in the scene. The same applies to the image since a pixel is defined to be square. We can deduce a simple ratio to implement this scaling. The length of one of the stripe segments in the calibration image corresponds to the vertical height of the corresponding plane of the calibration object. From these two lengths a ratio is formed giving units in pixels per millimetre. Thus, when calculating the local height at each vertical position along a single contour, the corresponding appropriate position in the image may be expressed appropriately in pixels.

Let the vertical length of one plane of the calibration object $=$ C millimetres

Let the length of the corresponding stripe segment in the calibration image $=$ c pixels

The position of a contour calculated to be at the horizontal position C millimetres along the base plane is equivalent to a horizontal position c pixels in the accumulator image.

The cartesian height of the same point may be deduced through analysis of the diagram in Figure 4. The cylindrical height is in bold. The dotted line denotes the corresponding cartesian height. $l\sin\theta$ is the horizontal position of the point in the cartesian co-ordinate reference frame and is expressed in millimetres. In order to place the processed stripe in an accumulator image at its appropriate position, it is necessary to scale it to the same measurement but in pixels.

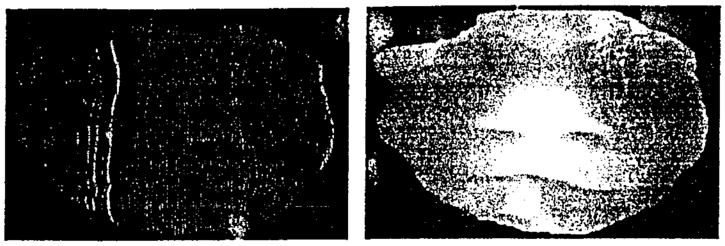

Figure 5: Accumulated stripes and interpolated surface.

Having accumulated all the contours into a single image, it is necessary to fit a continuous surface to the data in order to reconstruct the face. It is also necessary to collect between 150 and 200 contours in order to reconstruct the facial surface as faithfully as possible leaving perhaps between 1 and 3 pixels to be calculated horizontally between each contour, thus allowing the recovered surface to comprise mainly of actual data and as little predicted data as possible. The technique adopted here is Lagrangian interpolation.

4 Current research

Currently, research in the area of face analysis and registration is in progress. Some conventional techniques and newer more robust techniques have been taken into account.

Traditionally such matters have been dealt with anthropometrically through interaction between technical and medical parties. For registration of facial data in pre- and post- operative phases of plastic surgery, several approaches for comparing biological forms (in this case faces) are available. Such approaches involve primarily, the detection of special points of interest on the face. These influential points of interest are characterised as those that contribute most to the difference between forms. Adopting such techniques involves the presence of a qualified biologist whose primary input involves the localisation of these landmarks. Such data would subsequently be compared with other data through the application of such mathematical techniques as euclidian distance matrix analysis (EDMA)[8]. This entails the construction of a matrix whose elements are normalised distances between every possible pair of landmark points. The result being a scale invariant matrix of data pertaining to a face that is unique and may be compared with other similarly constructed matrices through inversion and multiplication.

In recent years, with the increasing interest in fully automated face data acquisition and description, attention has been turned to computer vision scale-space techniques[9]. In this technique individual vertical contours may be described in terms of the curvature of the curve segments comprising the contour. These segments may be isolated through gaussian scale space filtering which involves the successive convolution of gaussian filters of varying size to the contour in order to introduce a parameter of scale. The qualitative description of the contour is deduced by finding all the inflection points (zero crossings) in the second derivative.

More research is to be carried out before conforming to a particular registration methodology. In parallel with the research described above is the identification of the sources of error which play a crucial role in the overall accuracy of the system. It is intended that on identification of all such error sources, they may be taken in turn and modelled as accurately as possible in order to reduce if not eliminate them.

5 Conclusions

This paper has introduced the possibility of a cost effective non intrusive face scanning system. Employing computer vision range sensing techniques, it is possible to collect and describe the 3-D surface points of the face in terms of local height variations with respect to some predefined, pre-calibrated base plane. The technique addressed here is a variation on this concept and has the advantage of being simpler and more intuitive. This local height measurement technique requires a simple calibration step using a very basic calibration object consisting of 2 parallel planes. The camera system is set-up to incorporate a harmless light source from which a stripe of light emanates. The raw data may be acquired in the form of multiple images of the stripe as it strikes the surface of the face at different positions dictated by the constant axis of rotation. A simple yet novel method has been used to remove the need for sampling the stripes at pre-determined angles, thus the system is more robust and easy to use. Using a cylindrical co-ordinate reference frame, each set of 3-D co-ordinates associated with the points constituting the stripe in each of the 2-D images may be mapped to their appropriate positions in the base co-ordinate reference frame. Thus, from a sequence of 2-D stripe images a complete 3-D surface representation of the face is obtained. Future work includes modelling and eliminating sources of error plus conformation to a particular automatic face registration technique.

References

[1] L. C. Sanz, 'Advances in Machine Vision', Springer Verlag (NY) Inc., 1989, pp139-155.

[2] R. Nevatia, 'Machine Perception', Prentice-Hall Inc., Englewood Cliffs, (NJ), 1982, pp 167-168.