Estimation of a 3-D Surface Profile of a Human Face using a Hand-Actuated White Light Profilometer

P. Butler
Department of Computer Science
National University of Ireland, Maynooth
Co. Kildare, Ireland
pabutler@cs.may.ie

D. Vernon
Department of Computer Science
National University of Ireland, Maynooth
Co. Kildare, Ireland
dvernon@cs.may.ie

E. O'Broin
Temple Street Children's Hospital
Temple Street
Dublin, Ireland

Abstract

In this paper, we present a white-light striping configuration which addresses all three of the usual difficulties encountered with conventional laser configurations: size, accuracy and monochromaticity. This system is portable, yields highly accurate and repeatable stripe co-ordinates and produces a full colour 3-D rendering of the original surface. The portability is achieved by dispensing with servo-mechanisms and by relying on a user to rotate the scanner about a fixed fulcrum by hand. Consequently, since positional encoders are not used, the scan position has to be computed from the image itself. We use a fixed pulse-shaped registration target with vertically aligned edges. The apparent width of the pulse varies with viewing angle and, hence, the angle of rotation is computed from the distance between the edges. Concerning accuracy, the system uses a white-light intensity discontinuity rather than a laser. The chief advantage of this configuration is that the reflected stimulus - the discontinuity - to be measured is an inherent point like entity. We present and evaluate an algorithm for estimating the position of this intensity discontinuity to sub-pixel accuracy. Results are presented which demonstrate that the technique offers stripe localisation with between 0.1 and 0.01 pixel accuracy. Finally, the approach has the additional benefit that, because it uses white light illumination, it is also possible to sense the local colour and intensity of the surface being profiled. This colour information can be combined with the computed elevation of height data to yield a 3-D rendered model of the surface complete with the colour of the original object. This incorporation of the original colour variation is of significance in applications where the aesthetic appearance of the 3-D model is of importance, such as in the application - assessment of facial structure for pre- and post-operative surgical procedures - for which the system described in this paper was developed.

1. Introduction

Successful surgical procedure relies greatly on careful pre-operative planning aided by accurate diagnostic information. In the case of reconstructive surgery this information is mainly of a structural nature. The automatic recovery of such information has been a goal for more than 2 decades and many successful systems have been designed and implemented. The techniques used are many and varied from laser profilometry[1] to more expensive systems utilising computerised axial tomography (CAT)[2][3] , with the latter technique giving the option of also acquiring information regarding the underlying bone structure. There can be a risk involved in the use of such scanning techniques as they involve exposure of the patient to small doses of radiation which can be dangerous. Also, most of the existing systems to-day incorporate bulky motorised fixtures such as a rotating chair[1][4] or the use multiple imaging components housed in a special room[5]. Although such systems have proven to be

very accurate they do not lend themselves to use in remote clinical trials where portability is essential. The issues of accuracy and patient safety are accommodated for in the work described in this paper. However, the main objective of this paper is to address the implications of hand-actuation on the recovery of positional information in a rotational scan of the face and to ensure that the profile data acquired may be accurately characterised in terms of 3-D real-world height information and texture.

The paper is organised as follows. Section 2 provides a brief overview of the system configuration. Sections 3, 4 and 5 deal with the computation of the fundamental information from the raw image data. Section 6 provides some brief information regarding the processing of this information in order to create a 3-D elevation model of the face. Results are covered in section 7 and finally section 8 provides some conclusions.

2. System Overview

The current configuration consists of a CCD camera and a collimated white light source which is offset to the right of the camera by angle ϕ. An earlier version of this system is described in [6]. The light beam when striking the surface of the face, produces a vertical bar which deforms in accordance with the profile defining the structure of the face at that position (Figure 2(a)). The imaging system is mounted on an arm which is free to rotate about a controlled axis so that multiple profiles may be acquired within a cylindrical frame. The camera is positioned on its side so that maximum resolution in the vertical direction may be achieved. The reflectance function of the imaged light approximates a step-edge discontinuity in the horizontal direction. The light source is positioned so that the approximate location of the discontinuity coincides with the principle axis of the camera at the fulcrum of the system. This ensures that the local height variation along each profile is defined with respect to the same base position, i.e. the rotation axis. The rotation of the scanner is actuated by hand. The recovery of the facial structure in 3-D is accomplished on the basis of describing each point of the profile as a cylindrical co-ordinate[6][7][8]. This information may be computed from image data alone. The initial task to perform is the localisation of each profile within a sequence of imaged contours. The image processing software has been developed on a Sun Sparc 20 using the Khoros system.

3. Localising the Profile

Since each point of the profile is defined as the horizontal position of the discontinuity along its associated row in the image, it is necessary to localise this edge transition to as high a degree of accuracy as possible given that the information is represented in a discrete domain. For this purpose a simple but effective algorithm has been developed, which yields repeatability accuracy of between 0.1 and 0.01 of a pixel for any single edge.

3.1 Localisation of Intensity Discontinuity

The steps involved in localising a dark to light intensity discontinuity to sub-pixel accuracy are as follows:

- Identify the approximate position of the discontinuity.
- Find the bounds of the transition.
- Compute the intensity associated the edge transition.
- Compute the horizontal position associated with this intensity-level to sub-pixel accuracy.

The approximate position of the discontinuity may be found by scanning from left to right along a horizontal cross-section of the image, computing at each position the difference between the current intensity value and the next i.e. $f(x+1)-f(x)$, until the difference is greater than a pre-specified threshold. The position at which this occurs is considered to be the approximate position of the edge transition. The effectiveness of using a constant threshold value is dependent on the sharpness of the edge transition.

Next, it is necessary to find the maximum and minimum grey level values associated with the transition. This involves searching to the left and right of the approximate edge position and finding

the bounds of the transition. The bounds of the edge transition are defined as the position on either side, where the inter-grey level difference is a fraction of the inter-pixel threshold level specified earlier, e.g. 0.1 of the threshold. This allows for robust computation of the boundary positions in cases where surfaces exhibit varying reflectance characteristics (e.g. skin). This may cause the transition to take place over more or fewer pixels, making the use of a constant search window size invalid.
The mid-step grey-level is defined as the average of the maximum and minimum grey-levels within the bounds of the identified search window. This grey level defines the intensity associated with the edge transition. It is however, unlikely that this value will be present among the existing image information, so its associated position will not be defined as a discrete location.
In order to compute the position associated with the mid-step grey-level, it is necessary to find in the increasing part of the transition the 2 grey levels that occur on either side of it. These are defined as first, the grey-level closest to in value but less than the mid-step (position p) and second, the grey-level closest to but greater than the mid-step (position q). These are found by comparing the mid-step value (M) to each grey-level about the approximate edge position within the window specified earlier. Using $(p, f(p))$, $(q, f(q))$ and M it is possible to deduce the edge position E which is the position of M between p and q (Figure 1).

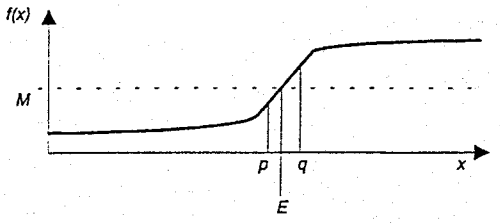


Figure 1 The localisation of an edge to sub-pixel accuracy

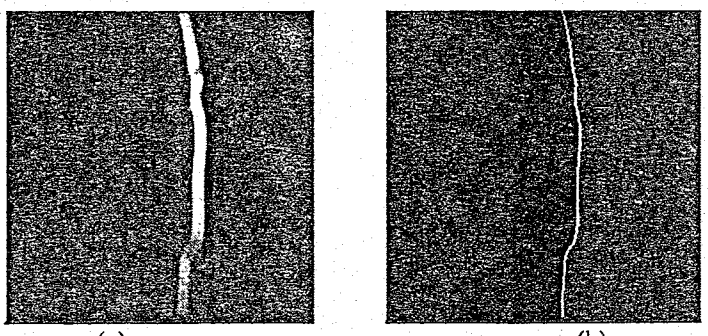
The assumption made here is that the transition is linear between the p and q, and from similar triangles:

$$\frac{E-p}{q-p}=\frac{M-f(p)}{f(p)-f(q)} \qquad (1)$$

Since q and p are adjacent pixels their difference is 1 pixel:

$$E=p+\frac{M-f(p)}{f(p)-f(q)} \qquad (2)$$

E is the position of the dark to light intensity discontinuity. The position of a bright to dark intensity discontinuity may be computed in a similar way. On computation of the edge position, this information is represented in the output image as a floating point intensity value, and stored in the nearest pixel corresponding to it for visualisation purposes (Figure 2(b)). All subsequent operations are performed on the real edge position.



(a) (b)

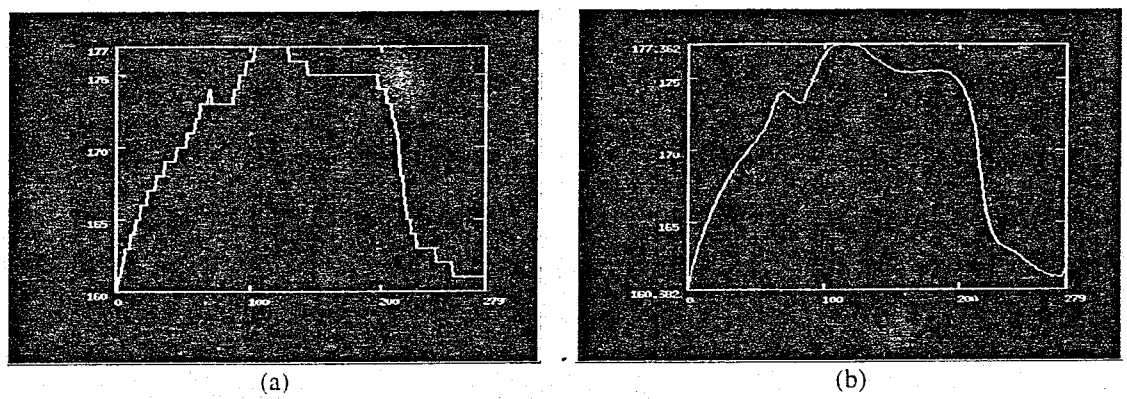
Figure 2 (a) an imaged contour, (b) the associated localised edge profile.

3.1.1 Error due to edge localisation

This edge localisation technique yields repeatability accuracy of between 0.1 and 0.01 of a pixel. Although very accurate, this still contributes to an error in edge localisation, approximately equal to the standard deviation of the computed edge set (σ has been computed to be 0.018 pixels over 100 images). This error is particularly important in the case of rotation angle computation. Referring to section 5, the computation of the inter-pulse edge distances are subject to 2 edge localisation errors, which are propagated through division and subsequently through an inverse trigonometric function to give the angle.

3.1.2 Quantisation Noise

Since each of the edges comprising a contour may be localised to sub-pixel accuracy, the inherent continuity of the profile may be preserved giving a very smooth transition from one edge to the next (Figure 3). When the height associated with each edge point is deduced (section 4), this smoothness propagates to produce a vertically smooth cylindrical profile.



(a) (b)

Figure 3 (a) plot of a profile localised to pixel accuracy, (b) the same profile to sub-pixel accuracy.

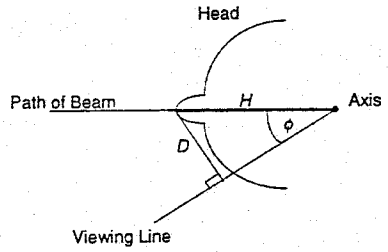
3.1.3 Texture Localisation

The significant advantage of using white light not only pertains to safety, but also to texture recovery. The sequence of profile images are acquired in RGB format. Each image of the sequence is converted to grey-scale (convert to HSV space[9] and use V) and passed through the edge localisation algorithm above. Using the edge information associated with a single image, the corresponding RGB texture information may be deduced from the original RGB image. The edge is however, likely to occur somewhere between the dark and bright region of the step discontinuity, so its intensity may be too low to represent, the texture associated with the edge point. It is because of this that the approximate texture corresponding to the point may be taken 2 to 3 pixels to the right of the edge position. This allows the creation of a set of RGB colour profiles matching in structural appearance, the set of edge profiles. This provides the raw material for producing a texture image of the face or object being scanned.

4. Computing the Cylindrical Height of a Profile Point

Having localised the profile edges it is now necessary to compute the real height associated with each point along it. The first thing to note is that the height of a point determines the extremity of the contour deflection thus implying that there is a relationship between the real height and the deflection value. The deflection is defined with respect to the image position of the axis of rotation which represents the centre of the cylinder. This relationship is may be computed from Figure 4, where D is the deflection of the beam in the real-world expressed in *mm*, and H is the associated height also expressed in *mm*. However, the imaging process causes this deflection information to be transformed to a 2-D discrete domain which is defined in different units, i.e. pixels. This projection process also incurs perspective distortion on the imaged information which implies that the mapping between the real height and the imaged deflection cannot be represented linearly. Thus, a calibration step is

introduced which allows an approximation of the perspective function to be deduced on the basis of deflection and real world height.



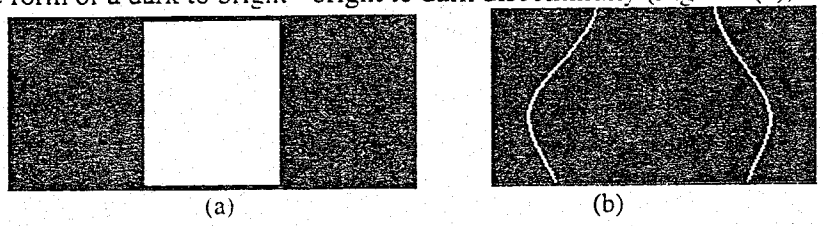
$$H = \frac{D}{\sin\phi} \qquad (3)$$

Figure 4 Computation of local height based on profile deflection.

In order to model the non linear mapping between the real-world height and the image deflection, its is necessary to base the calibration on a number of heights and their associated deflection values. This is achieved by using a calibration object with a varying set of heights that may be matched to deflections in the calibration image. The mapping between the 2 data sets represents a curve which in this case is approximated by a second order polynomial. The common set of 3 coefficients associated with an over determined system of equations (10+) is deduced via a least squares error solution. The x co-ordinates represent the deflection values calculated from the image and the y values represent their associated real-world heights. The real world height associated with a deflection may thus be computed by evaluating *f(d)* where *d* is the imaged deflection of the contour point. Thus the radial height may be computed for all points on a profile.

5. Computing the Angular Position associated with a Profile

In the case of hand actuation, the contour data comprising the object is sampled in non-uniform intervals as the scanner rotates about the face. It is thus necessary to compute on the basis of the image information the positions at which each contour has been acquired. This is achieved using pulse shaped registration target with vertically aligned edges which exhibit a strong contrast horizontally on either side in the form of a dark to bright - bright to dark discontinuity (Figure 5(a)).



(a) (b)

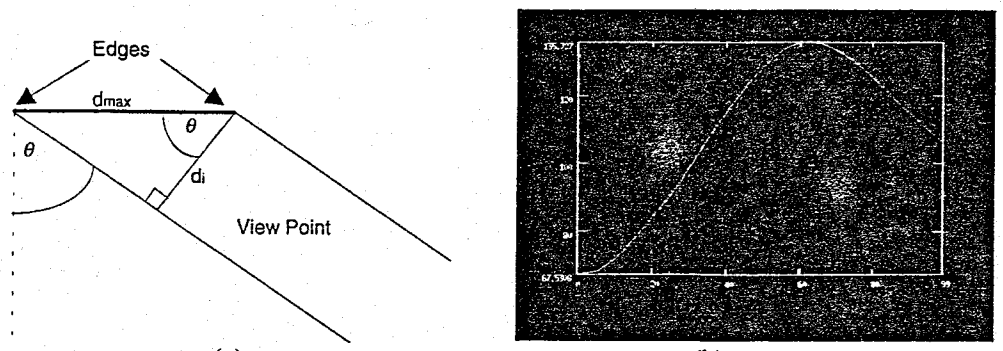
Figure 5 (a) pulse shaped registration target, (b) time image of the change in inter-edge distance.

The registration target is placed above the head in the field of view of the camera and centred at the axis of rotation. As the camera rotates about this the white region appears narrower in horizontal cross-section. By performing edge localisation on each varying pulse at a constant vertical offset into each image the position of both discontinuities may be deduced. The edge localisation algorithm employed allows sub-pixel accuracy (section 3.1) to be achieved thus allowing the width of the pulse to be calculated to similar accuracy. Figure 5(b) shows how the inter edge distance varies with rotation.

This change in inter-edge distance is a function of the rotation angle associated with each image (Figure 6(a)). The angle is computed thus:

$$\theta_i = \cos^{-1}\left(\frac{d_i}{d_{\max}}\right) \qquad (4)$$

where d_i is the apparent distance between the pulse edges in image i and d_{max} is the actual distance between them when they appear a maximum distance apart. θ_i is the rotation angle. This evaluates to zero degrees when both distances are equal, defined to be negative as d_i approaches d_{max}, and positive as d_i decreases.



(a) (b)

Figure 6 (a) the relationship between inter-edge distance and θ, (b) typical plot of set of inter-edge distances.

As can be seen from equation (4), it is necessary to have prior knowledge of the maximum inter-edge distance in order to compute the angle. It is not entirely accurate to assume that this will appear among 100 - 150 images in a scan sequence, so this value is deduced by fitting a high order polynomial to the set of distances and computing the maximum. This action is justified by the fact that a scan is performed with a smooth motion giving rise to inter-edge registration data that resembles a polynomial (Figure 6(b)). Having acquired both d_i and d_{max}, the rotation angle is deduced through the application of the formula above.

The uncertainty in estimating the inter edge distance of the pulse function is the sum of the uncertainty in estimating the first edge and the second edge. However, since the estimation of both edges is subject to the same error, the composite error may be evaluated as twice the error in localising a single edge (section 3.1.1).

The angle θ represents the position of the profile with respect to the position of the camera. The profile position should be defined with respect to the light source. This correction is implemented by computing the angle between the camera and the light source. This is then subtracted from each θ to transform the angular position so that it is defined with respect to the light source. The inter-edge distance associated with ϕ is deduced as part of the calibration step, by aligning the calibration object with the registration target so that the beam of light strikes it orthogonally. The angle associated with this distance is computed in the same way as θ.

6. Subsequent Processing

Having extracted the height and positional information from the raw image data, it is now necessary to organise the profiles using this positional information and generate an elevation image of the face.

6.1 Placing the Profiles

Once all of the cylindrical heights have been computed from a sequence of images, its is then necessary to find their position horizontally in relation to one another so that they may be placed together in an accumulator. In this case the position is calculated on the basis of a cylinder whose radius is the maximum height among the set of profiles (in the case of a face this will probably correspond to the highest point along the mid-line profile). The position is determined as the length of the arc to the left or right of the position associated with a rotation angle of zero degrees. If theta is negative then the placement will be to the left, otherwise it will be to the right. The arc length associated with an angle position within the global cylinder may be computed from equation (5).

$$x = H_{\max} \times \theta_i \quad (5)$$

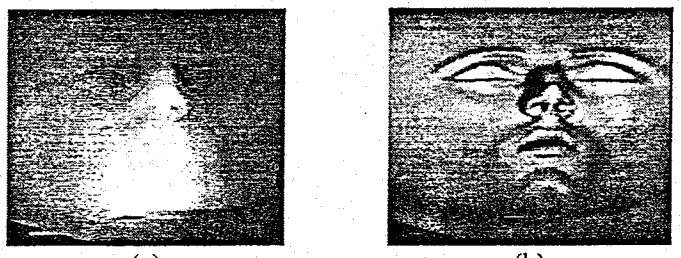
Since the image domain doesn't represent negative positions, all positions are scaled up by half the width of the accumulator image, so that they may be all expressed as positive positions. The Khoros Polymorphic Data Model (PDM) allows storage of explicit location data associated with each pixel thus the accuracy to which the angle and height information is computed is propagated to the next

stage of processing. The texture profiles are compiled separately in a similar manner so that there is a matching colour profile associated by position with each height encoded profile.

6.2 Computing the Inter-Profile information

The surface comprising the profiles is estimated on the basis of Lagrange's interpolation method[10][11] . The algorithm devised allows the inter profile points to be deduced on the basis of 1 or more existing points on either side of each. For the purposes of this application the use of 2 points on either side has proved sufficient to represent the transitions along each cross-section of the profile image.

For each point to be deduced this requires that 2 existing points be recorded on either side. The position of each point along the horizontal represent the x values for the function and the associated heights (grey-levels) represent the y values, and are taken from the accurate location data associated with the image computed when the profiles were accumulated. Each term of the function is evaluated as a function of the position of the required height point, and the x and y values associated with each of the specified surrounding points. The application of this methodology to each of the horizontal cross-sections comprising the profile image will allow a surface to be approximated (Figure 7(a)).



(a) (b)

Figure 7 (a) the interpolated elevation image based on the profile information and (b) its associated texture.

The inter-profile texture information is approximated in a similar but less rigorous manner through interpolation between each pair of points. The interpolation is performed on each band separately to model the gradual change in band intensity from point to point. The result is a colour surface matching in structure the associated grey-scale cylindrical height map (Figure 7(b)).

6.3 Transformation to Cartesian Frame

A cylindrical height map has now been generated on the basis of the original profile data. The cylindrical data points are represented as 3-D x, y, z locations. Each profile position is computed as the size of the arc separating it from the zero degrees position . So, from equation (5):

$$x = \left(H_{\max} \times \theta\right) \Rightarrow \theta = \frac{x}{H_{\max}} \qquad (6)$$

To 're-wrap' the cylindrical data the position of a point of height z on this profile must be expressed in cartesian form[8]. The transformation per co-ordinate is shown in equations (7), (8) and (9).

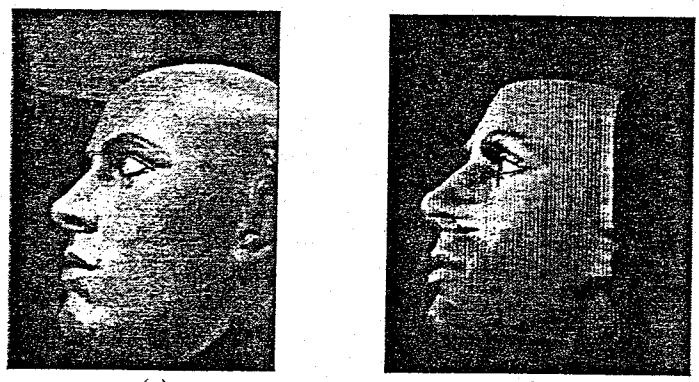
$$x\_cart = z \times \sin\theta \quad (7), \qquad y\_cart = y \quad (8), \qquad z\_cart = z \times \cos\theta \quad (9)$$

A quadrilateral mesh geometry is applied to these transformed points and allows the 3-D surface to be rendered for visualisation. Some visual results are provided in the next section.

7. Results

A manequin's head was scanned and the profile data processed to produce a rendered result. Figure 8 allows visual comparison between the result and an image of the head. Although the dimensional information appears to be approximately equal between both images, some anthropometric analysis techniques must be used to validate the comparison. Future work includes the design of some

interactive software to help achieve this quantitative analysis accurately. In view of this, automatic registration software is currently under development.



(a) (b)

Figure 8 (a) image of a manequin's head and (b) a similar view of the rendered surface with over-laid texture.

Referring again to the result (Figure 8(b)), it may be noted that the rendered surface exhibits some distortion about the nose area (particularly about the nostril) and this is caused by occlusion. The contour data about one side of the nose becomes invisible to the camera because the nose blocks its field of view. The interpolation software tries to interpolate across this gap thus degrading the result. Some considerations regarding minimisation of this problem are currently being assessed including possible introduction of a second camera or light source to allow data associated with occluded regions to be recorded.

8. Conclusions

A rotational white light based profilometry system has been described. The use of white light significantly reduces the risk of damage to the eyes and skin. The system achieves automatic profile edge and texture localisation with the aid of a simple algorithm. The data is processed to produce a colour rendering of the face surface. The portability of the system is promoted by hand-actuation of the rotating arm. The position of a profile is computed from registration information in the image. Future work includes modification of the imaging system to eliminate, or at least reduce, the affects of occlusion. Also, the design of some anthropometric analysis software is being considered.

References

[1] Moss J. P., Linney A. D., Grindrod S. R. & Mosse C. A., *A laser scanning system for the measurement of facial surface morphology*. Optics and Lasers in Engineering vol. 10. 1989, pp 179-190.

[2] Linney A. D., Moss J. P., Richards R., Mosse C. A., Grindrod S. R. & Coombes A. M. *Use of 3-D visualisation systems in the planning and evaluation of facial surgery*. SPIE vol. 1380 Biosteriometric Tech and Appl. 1990, pp 190-199.

[3] Vannier M. W., Marsh J. L., Warren, J. O., *Three dimensional computer graphics for cranio-facial surgical planning and evaluation*. Computer Graphics vol. 17, no. 3, July 1983, pp 263-273.

[4] Grindrod J. R., *Non-contact 3-D surface digitisation of the human head*, NCGA Conf. Proc. 1989, vol. 1, pp 132-141.

[5] Vannier M. W., Commean P., Pilgrim T. K., Bhatia G., Brundsen B., *Facial Surface Scanner*, IEEE Computer Graphics Appl. 1991, vol. 11, no. 6, pp 72-80.

[6] Butler P., O'Broin E., & Vernon D., *The recovery of 3-D facial shape for analysis in the pre- and post-operative phases of reconstructive surgery by hand-actuated light striping*, OESI Lasers and Optical Eng. '96.

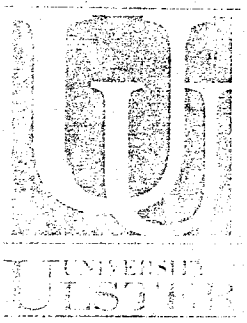
[7] Paul R. P., *Robot Manipulators*. MIT Press Cambridge Massachusetts, 1981,

[8] Thomas J. B. & Finney R. L., Calculus an analytical geometry, Addison Wesley, 1988, pp 852-854, 1004.

[9] Foley J. Van Dam A., et al., *Introduction to computer graphics*, Addison Wesley, 1994, pp 410-419.

[10] Kronsjo L., *Algorithms: their complexity and efficiency*, 2nd ed. J. Wiley & Sons, 1987, pp 10-15.

[11] Press W. H., *Numerical recipes in C*, Cambridge University Press, 1992, pp 85-110.



School of Information and Software Engineering
Faculty of Informatics

Irish Machine Vision and Image Processing Conference (IMVIP-97)

and

Eight Ireland Conference on Artificial Intelligence (AI-97)

Volume 1