

A PHOTOGRAMMETRIC METHOD FOR THE ASSESSMENT OF FACIAL MORPHOLOGY IN FETAL ALCOHOL SYNDROME SCREENING

Tania S. Douglas^{a,*}, Fernando Martinez^a, Ernesta M. Meintjes^a, Christopher L. Vaughan^a, Denis L. Viljoen^b

^a MRC/UCT Medical Imaging Research Unit, Dept of Human Biology, University of Cape Town, Observatory 7925, South Africa - (tdouglas, fernando, ernesta, kvaughan)@cormack.uct.ac.za

^b Dept. of Human Genetics, National Health Laboratory Service, University of the Witwatersrand, PO Box 1038, Johannesburg 2000, South Africa – denisv@mail.saimr.wits.ac.za

KEY WORDS: biometrics, image processing, feature extraction, medicine

ABSTRACT:

Photogrammetry has been introduced as an alternative to direct measurements to obtain facial distances for a variety of anthropometric applications. We apply photogrammetry to assist in the screening of children for Fetal Alcohol Syndrome (FAS). We have developed a stereo-photogrammetric image acquisition platform to obtain digital photographs of 6 to 7 year-old children, and image processing algorithms to extract the eye features required for FAS screening from these photographs. Images are calibrated using the Direct Linear Transform. Eye features are extracted by examining peak and valley maps to locate the eyes and extract the iris contour, and applying a genetic algorithm to match cubic splines to eye contours.

1. INTRODUCTION

Photogrammetry has been introduced as an alternative to direct measurements to obtain distances between facial landmarks using both two-dimensional and three-dimensional methods. Obtaining measurements from photographs is less intrusive to the patient and less costly, provides a permanent record of the face that can be accessed at a later time, and offers consistency in longitudinal studies in which different observers with different direct measuring techniques might participate.

In this paper we report on the use of a stereo-photogrammetric tool and image processing algorithms (written in MATLAB) to obtain and to extract eye features from digital photographs of 6-7 year-old children for Fetal Alcohol Syndrome (FAS) screening. Among the facial features characteristic of FAS, those present around the eyes can be evaluated by distance measurements; the required points are the eye corners and the pupil centres. The eye distance measurements used in FAS screening are shown in Figure 1.

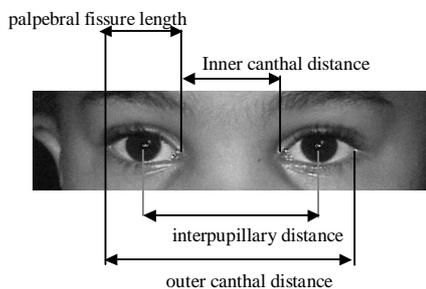


Figure 1. Eye measurements used in Fetal Alcohol Syndrome Screening

2. METHODOLOGY AND RESULTS

2.1 Image Acquisition

The image acquisition platform allows the face of a child to be photographed in a calibration frame simultaneously on a pair of high-resolution digital cameras (Sony DKC-FP3 Digital Still cameras) mounted at the same distance (104 cm) from the child. The cameras are triggered simultaneously by remote control. The equipment is shown in Figure 2.

The control frame contains a chin- and head-rest. Eleven retro-reflective control markers with known three-dimensional coordinates are mounted on the calibration frame so that they are visible on both the left and the right images.

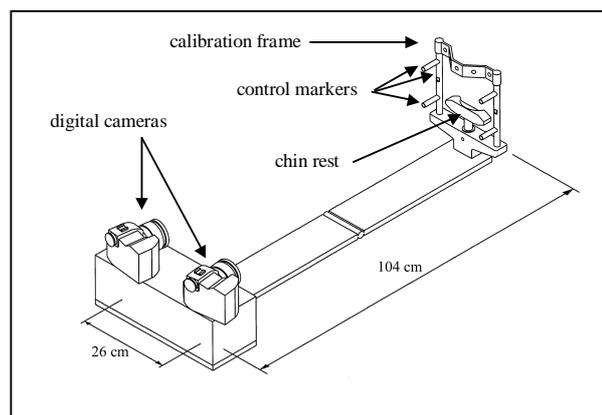


Figure 2. Image acquisition tool

Some of the control markers are raised above the surface of the frame in order to achieve a better depth-of-field while the face

* Corresponding author.

of the child is located within the three-dimensional volume defined by the eleven control points. The control frame was measured on a 3-axis Bridgeport digital milling machine with a spatial resolution of 0.01 mm.

2.2 Image Calibration

An algorithm locates automatically the two-dimensional coordinates of the control frame markers in each pair of images, and obtains the mapping from two-dimensional image coordinates (x and y) into three-dimensional object-space coordinates (X , Y and Z), based on the known three-dimensional locations of the markers. The Direct Linear Transformation (DLT) (Abdel-Aziz & Karara, 1971) is used to transform two-dimensional image coordinates (x_k and y_k) into three-dimensional object-space coordinates (X , Y , and Z). The DLT equations are:

$$x_k + x_k K_k r_k^2 = \frac{B_{k1}X + B_{k2}Y + B_{k3}Z + B_{k4}}{B_{k9}X + B_{k10}Y + B_{k11}Z + 1} \quad (1)$$

$$y_k + y_k K_k r_k^2 = \frac{B_{k5}X + B_{k6}Y + B_{k7}Z + B_{k8}}{B_{k9}X + B_{k10}Y + B_{k11}Z + 1} \quad (2)$$

where $k = L$ or R denotes values pertaining to the left or right images, respectively, and $r_k^2 = (x_k^2 + y_k^2)$, $B_{k1} \dots B_{k11}$ are the eleven parameters that describe the relationship between the two-dimensional image space and the three-dimensional object space, and K_k is a term that corrects for lens distortion (Karara & Abdel-Aziz, 1974).

Each of the eleven control markers, with known object-space coordinates (X , Y , Z) and two-dimensional image coordinates, (x_L , y_L) and (x_R , y_R), gives rise to four equations. For each pair of photographs, the system of 44 equations is solved for the 22 B parameters, K_L , and K_R using a least squares adjustment. Calibration accuracies better than 0.5 mm are achieved in each direction. Once the B and K parameters are known, the three-dimensional object-space coordinates (X , Y , Z) of any point that appears on both the left and right images can be calculated.

2.3 Estimating Eye Locations

Peak and valley maps (Esme et al., 1996) are calculated as shown in equations (3) and (4).

$$I_v(x, y) = \frac{u(I_w(x, y) - I(x, y)) \times |I_w(x, y) - I(x, y)|}{I_w(x, y)} \cdot 255 \quad (3)$$

$$I_p(x, y) = \frac{u(I(x, y) - I_w(x, y)) \times |I_w(x, y) - I(x, y)|}{I_w(x, y)} \cdot 255 \quad (4)$$

The valley map (I_v) selects dark regions in the image, such as the iris. It compares pixels in an image window (I) with the window average (I_w), assigns the value '0' to those pixels greater than the window average, and assigns to those pixels less than the window average a value proportional to the difference between the pixel and the average. The new pixel values are normalized to the range 0 to 255. I denotes the original image, and u denotes the unit step function.

The peak map (I_p) selects bright regions in the image, such as the sclera, in a similar manner: it assigns the value '0' to those pixels less than the window average, and assigns to those pixels greater than the window average a value proportional to the difference between the pixel and the average.

Since the face is always photographed in a fixed position using the equipment shown in Figure 2, an image window in which the eyes are expected to appear can be specified in order to reduce the image area being examined. The peak and valley maps for window of a facial image are shown in Figure 3.



Figure 3. (a) Peak and (b) valley maps

Integral projection is a useful technique for determining the position of facial features (Brunelli & Poggio, 1993; Sobottka & Pitas, 1998). The vertical $V(x)$ and horizontal $H(y)$ integral projection of an image $I(x, y)$ in the window $[x1, x2] \times [y1, y2]$ are defined as:

$$V(x) = \sum_{y=y1}^{y2} I(x, y) \quad (5)$$

$$H(y) = \sum_{x=x1}^{x2} I(x, y) \quad (6)$$

Integral projections of the combined peak and valley map are used to obtain estimates of eye centre locations. The eye centre estimates are defined as the points where the peaks in $V(x)$ and $H(y)$ coincide. Once a location estimate has been obtained for each eye, further image processing concentrates on regions around the location estimate, and each eye is extracted separately.

2.4 Extracting the Iris Contour

The peak map is converted to binary and a small window in which the sclera region of the eye could be expected to occur is defined around the estimate of the eye centre. The small size of this window ensures that areas other than the sclera around the eyes that may appear bright in the image due to non-uniform lighting are excluded. The sclera window is divided further into two sub-windows to the left and right of the eye centre estimate. The largest 8-connected component in each sclera sub-window is regarded as belonging to the sclera. The corresponding 8-connected region in the original peak map is extracted to represent an estimate of the sclera region. In order to obtain the contour of the iris, a circle is fitted to the inner boundary points of the sclera region for each eye. The extracted sclera region and iris contour are shown in Figure 4.

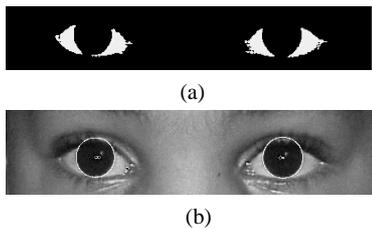


Figure 4. Extracted (a) sclera region and (b) iris contour.

2.5 Extracting the Eye Contour

Yuille et al. (1992) proposed, and Xie et al. (1994) and Esme et al. (1996) modified a method of eye feature extraction based on deformable templates in which the eyelids were modeled by 2 parabolic sections. Our method models the eye by a set of cubic splines (one representing the upper and one the lower eyelid).

The eye contour is found by using a genetic algorithm (GA) to match these cubic splines to bright contours in the valley map. Two four-point cubic splines are used, one for the upper and one for the lower eye contour. Each cubic spline is described by 4 control points, and the end points are common.

2.5.1 GA Parameters: The parameters being optimised by the GA are the locations (x- and y-co-ordinates) of the six cubic spline control points that specify the eye template.

2.5.2 GA Initial Population: An initial random population of chromosomes (each consisting of 6 sets of co-ordinates representing 6 cubic spline control points) is generated.

2.5.3 GA Fitness function: The area corresponding to the extracted iris is removed from the valley map. The sum of the pixels in this modified valley map I_{mod} , that correspond in position to the pixels of the eye template (x_i, y_i) , normalised over the length n of the eye template, as shown in Equation 5, is the fitness function.

$$f = \frac{1}{n} \sum_{i=1}^n I_{\text{mod}}(x_i, y_i) \quad (7)$$

2.5.4 GA Reproduction: An elitist strategy is employed: one third of the chromosomes in each new generation are those with the highest fitness values from the previous generation. Another third of the new generation is obtained by performing simple crossover between randomly selected pairs of strings from the previous generation with no weighting based on string fitness, and the final third consists of new randomly generated chromosomes. The population size remains constant.

2.5.5 Automatic selection of points: The endpoints of the extracted eye contour splines and the centroids of the extracted iris contours represent the eye corners and the pupil centres respectively. The distances shown in Figure 1 can be calculated using the three-dimensional object-space co-ordinates of the eye corner and pupil centre points. Figure 5 (a) shows 3 extracted eye contours, judged by visual inspection to be representative of the actual features.

Figure 5 (b) shows eye contours that underestimate the inner eye corner in images in which sclera and skin cannot be clearly distinguished.

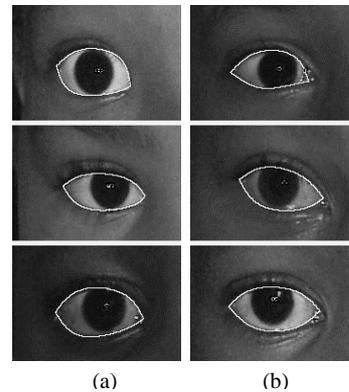


Figure 5. Extracted eye contours: (a) a good fit; (b) underestimation of inner corners.

3. CONCLUSION

The algorithm described here may be applied as a step in the screening of children for FAS, in order to reduce the manual load required in such an endeavour. It may also be used in other applications of facial anthropometry that require eye distance measurements. The extent to which the observed inaccuracies (cf. Figure 5 (b)) would affect the accuracy of distance measurements and therefore screening accuracy must be determined.

ACKNOWLEDGEMENT

This work was conducted with the aid of grants from the Medical Research Council and the National Research Foundation in South Africa.

REFERENCES

- Abdel-Aziz YI, Karara HM., 1971. Direct Linear Transformation from Comparator Coordinates into Object Space Coordinates. *American Society of Photogrammetry Symposium on Close-Range Photogrammetry*, Falls Church, Virginia, U.S.A., pp. 1-18.
- Brunelli R, Poggio T., 1993. Face Recognition: Features versus Templates, *IEEE Trans. PAMI*, 15, pp. 1042-1052.
- Karara HM and Abdel-Aziz YI., 1974. Accuracy Aspects of non-metric imageries. *Photogrammetric Engineering*, 40, pp. 1107-1117.
- Sobotta K, Pitas I., 1998. A novel method for automatic face segmentation, facial feature extraction and tracking, *Signal Processing: Image Communication*, 12, pp. 263-281.
- Xie X, Sudharkar R, Zhuang H., 1994. On improving eye feature extraction using deformable templates, *Patt. Recog.*, 27, pp. 791-799.
- Yuille AL, Hallinan PW, Cohen DS., 1992. Feature extraction from faces using deformable templates, *Int. J. Computer Vision*, 8, pp. 99-111.