# 3D CURVATURE ANALYSIS WITH A NOVEL ONE-SHOT TECHNIQUE.

*J. Matías Di Martino, Alicia Fernández, José A. Ferrari.*

Universidad de la República, Facultad de Ingeniería
J. Herrera y Reissig 565, 11300 Montevideo, Uruguay

## ABSTRACT

In this work we will show that it is possible to perform three-dimensional (3D) analysis and extract features without performing a 3D surface reconstruction. We propose a novel approach that consists in acquiring the scene under the projection of structured light (plus the natural ambient illumination), which allows us to obtain the depth gradient map using a simple processing step with very low hardware requirements.

***Index Terms*—** Face curvatures, 3D feature extraction, HK classification, Nose location.

## 1. INTRODUCTION

Landmark detection is one of the first steps on face recognition or gesture detection systems. In the past five years, most of the approaches began to include 3D information for the sake of face detection[1], face recognition[2, 3, 4] and landmarks localization[4, 5, 6, 7]. Due to the increasing number of three-dimensional scanners (e.g. the Kinect device of Microsoft) some 3D databases are becoming popular (e.g. BJUT-3D and FRGC 2.0) and a large number of publications have recently addressed this topic. Even though a great amount of progress has been made on 2D color or intensity face analysis, it has been demonstrated that the expression, pose and illumination variations have significant impact on 2D systems performance.

In this work, we propose acquiring the scene under the projection of structured light (plus the natural ambient illumination) and measure surface curvature from the deformation of the projected pattern. The main difference with respect to traditional one-shot 3D shape profiling methods [8] is that we do not need to compute the scene depth, instead, depth gradient is directly calculated from the acquired image and so surface curvature. This is extremely useful for 3D curvature analysis and other feature based techniques that use the depth partial derivatives. In addition, the pattern we are projecting is binary and static, we do not need the projection of several patterns or the use of color coding techniques (this means that no gamma or color calibration [9, 10, 11], nor synchronization are needed). Because of that, the structured illumination can be provided by a simple flash (including a square binary amplitude grating and a small lens) and it could eventually be embedded on a smart phone in the near future.

In Section 2, we present an overview of the proposed technique explaining the key features and the different processing steps. Section 3 presents the main results showing the potentiality and usefulness of the method. Finally, Section 4 concludes the work and enumerates some future lines of research and applications.

## 2. METHOD DESCRIPTION

When a scene is being captured by two cameras and assuming that just a translation exists between them, the apparent image shift (disparity) gives information about relative depth of the scene as illustrated on Fig. 1. Assuming $H \gg max\{h, f\}$, the disparity $D(x, y)$ between both images (i.e., the shift of the images on the detector arrays of the cameras) and the depth $h(x, y)$ of the test surface relative to the reference plane will be related by,

$$h \approx \frac{\overline{Y'Y}}{B/H}\ , \quad \overline{Y'Y} \approx D\frac{H}{f}. \tag{1}$$

Hence,

$$h \approx D\frac{H^2}{B\,f} \tag{2}$$

where $f$ is the focal length, $H$ is the distance of the reference plane and $B$ is the distance between the cameras' centers [12].

The previous expression was obtained by considering two cameras, but it also holds when we have a camera and a projector, as shown in literature.[13, 8, 14, 15] For illustrative purpose in Fig. 1 the relative translation between camera and projector is along the $y$ coordinate. In the following, the shift between camera and projector will be along a direction at $45$ degrees with respect to the $x$ and $y$ coordinates.

The procedure followed for 3D-shape retrieval was proposed in [16] and consists in measuring the partial derivatives of the disparity ($D$). Let us assume that we are projecting a rectangular fringe pattern of period $p$ in the $x$ and $y$ direction over a test surface, where $I(x, y)$ is the pattern acquired by the camera (see Fig.2(a)). The 2D-spatial spectrum of this image

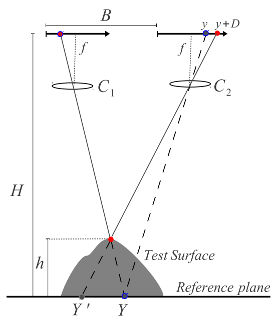



**Fig. 1**. Principle of stereoscopic vision. The figure shows two cameras $C_{1,2}$ with lenses of focal distance $f$ separated a distance $B$, placed at a distance $H$ from a reference plane.

is shown at the center of Fig. 2, in which the red points correspond to the vicinity of the spatial carrier's frequency (i.e. $2\pi/p$). By performing a simple Fourier filtering (red regions in Fig. 2), one can obtain a pattern ($I_v(x, y)$) with deformed vertical fringes and another one ($I_h(x, y)$) with deformed horizontal fringes, as shown in Figs.2(b) and 2(c), respectively. [In Fig. 2 we are assuming that the $x$ and $y$ direction are horizontal and vertical, respectively.]

Therefore, it does not matter if the projected fringes are binary or sinusoidal. Without loss of generality, by filtering the spatial spectrum of $I(x, y)$, one obtains

$$I_h(x, y) = I_0(x, y) \cos\left((2\pi/p)(y + D(x, y))\right) \quad (3)$$

and

$$I_v(x, y) = I_0(x, y) \cos\left((2\pi/p)(x + D(x, y))\right) \quad (4)$$

where $I_0(x, y)$ is a function of the reflectance of the test surface. As usual, we are assuming that $I_0(x, y)$ and $D(x, y)$ are low-frequency functions in comparison with the frequency of the spatial carrier, i.e. $D_i \ll 1$ and $I_{0i} \ll 2\pi/p$, where the subscript denotes partial derivative with respect to the variable $i(= x, y)$.

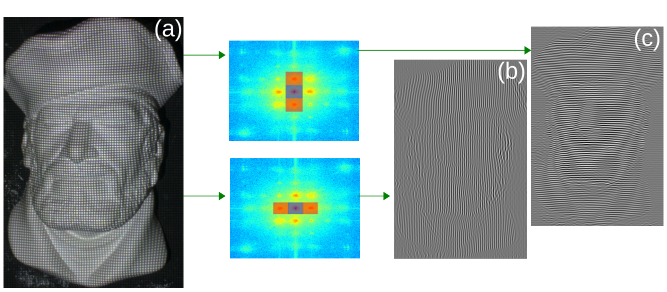



**Fig. 2**. (a) Acquired image. The colored figures show the 2D-spatial spectrum of the image. By filtering the horizontal and vertical Fourier components (red in the figures), one obtains the images shown in (b) and (c), respectively.

Then, by taking partial derivatives with respect to the $x$ and $y$ coordinate, from Eq. (3) and Eq. (4) one obtains

$$\begin{aligned} I_{hi}(x, y) \approx & -(2\pi/p) I_0(x, y) \\ & \sin\left[(2\pi/p)\left(y + D(x, y)\right)\right]\left(y_i + D_i(x, y)\right) \end{aligned} \quad (5)$$

$$\begin{aligned} I_{vi}(x, y) \approx & -(2\pi/p) I_0(x, y) \\ & \sin\left[(2\pi/p)\left(x + D(x, y)\right)\right]\left(x_i + D_i(x, y)\right) \end{aligned} \quad (6)$$

where $x_i = 1$ and $y_i = 0$ for $i = x$, and $x_i = 0$, $y_i = 1$ for $i = y$.

Hence, it is easy to demonstrate that

$$D_x(x, y) \approx \frac{I_{hx}(x, y)}{I_{hy}(x, y)} \quad (7)$$

and

$$D_y(x, y) \approx \frac{I_{vy}(x, y)}{I_{vx}(x, y)}. \quad (8)$$

We conclude that the gradient of the disparity ($D$) can be calculated in a simple manner as the ratio of the derivatives of the (horizontal and vertical) Fourier components of the image $I(x, y)$ acquired by the camera.

In order to analyze the face curvature, let $S$ be the surface defined by a twice differentiable real valued function $D : U \to \mathcal{R}$, defined on an open set $U \subseteq \mathcal{R}$:

$$S = \{(x, y, z) \,|\, (x, y) \in U;\ z \in \mathcal{R}; D(x, y) = z\}. \quad (9)$$

For every point $(x, y, D(x, y)) \in S$ we consider two different curvature measures, the Mean ($H$) and the Gaussian ($K$) curvature defined as[1, 17]:

$$H(x, y) = \frac{\left(1 + D_y^2\right) D_{xx} - 2 D_x D_y D_{xy} + \left(1 + D_x^2\right) D_{yy}}{2\left(1 + D_x^2 + D_y^2\right)^{3/2}}, \quad (10)$$

$$K(x, y) = \frac{D_{xx} D_{yy} - D_{xy}^2}{\left(1 + D_x^2 + D_y^2\right)^2}. \quad (11)$$

As the presented technique allows us to directly compute the first derivatives of the scene depth, it is only necessary to compute the second order derivatives ($D_{xx}$, $D_{yy}$ and $D_{xy}$) e.g. using finite differences (i.e. $D_{xx}[i, j] = (D_x[i, j + 1] + D_x[i, j - 1])/2)$) or any equivalent numeric method for differentiation.

Once the Mean and Gaussian curvatures are calculated, it is possible to classify the different areas of the face according to its shape [18]. Depending on the $H$ and $K$ values, points on the surface are classified following the rule shown in Table 1.

In this work, $z$ coordinate (named $h$ in Fig. 1) is measured from a reference plane, it is clear that for this coordinate system the area of the nose can be identified as a well defined elliptical concave region while the corner of the eyes

| | $K < 0$ | $K = 0$ | $K > 0$ |
|---|---|---|---|
| $H < 0$ | Hyperbolic concave | Cylindrical concave | Elliptical concave |
| $H = 0$ | Hyperbolic symmetric | Planar | Impossible |
| $H > 0$ | Hyperbolic convex | Cylindrical convex | Elliptical convex |

**Table 1**. $HK$ classification.

must present an elliptical convex shape. To remove smooth regions from the areas of interest a threshold approach was followed[1, 19], then those points with high absolute value of $K$ and $H$ were isolated. For those points with $K > 0$ and $H > 0$ the eye candidate label was assigned while those with $K > 0$ and $H < 0$ were selected as nose candidates.

### 2.1. Implementation details.

The main ideas were presented in the previous section. In this subsection, some details concerning the implementation are addressed. Figure 3 shows the main processing steps and illustrative images of the outputs obtained.

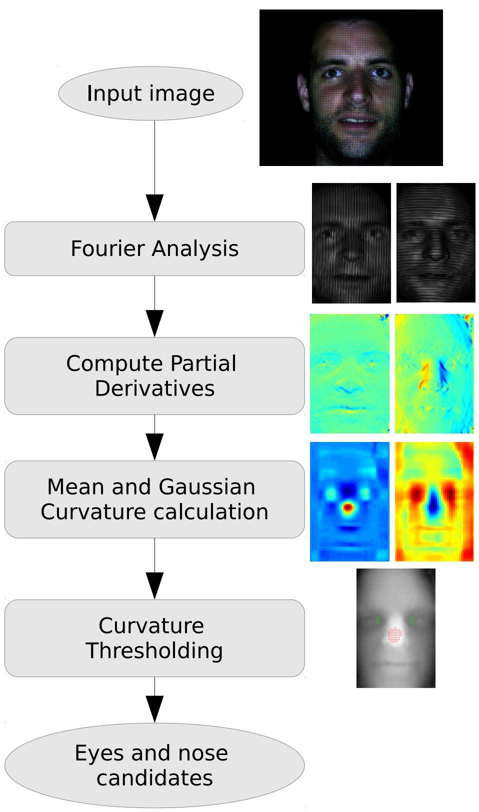



**Fig. 3**. Block diagram.

In this work, a black background was used, and the input image was filtered using *Hann* windows over the Fourier domain (we also tried other kinds of windows such as *rectangular* or *hamming* ones, they all yielded very similar results). Partial derivatives are computed following Eqs. (7) and (8). Due to shadows or surface discontinuities some singularities appear in $D_x$ and $D_y$ (areas where $I_{hy} \approx 0$ or $I_{vx} \approx 0$), these noisy points were removed by applying a median filter. Before proceeding with the curvature analysis, a downsampling step may be required to set the correct scale of the shapes we are looking for (also a multi-scale approach can be followed if the scale of nose and eyes is unknown on input image). In this work the resolution was $600 \times 400$ pixels (after removing the black edges) and we used one tenth of resolution for the curvature analysis. Finally, concerning the thresholding step, values of $H_{th}$ and $K_{th}$ are selected proportional to the mean values (to make the selection independent of the image range): $H_{th} = 3\overline{|H(x,y)|}$, $K_{th} = 5\overline{|K(x,y)|}$, these values were empirically set looking for a compromise between false positive and false negative rates. The values of thresholds obtained are similar to those reported in [1] and [20].

## 3. RESULTS

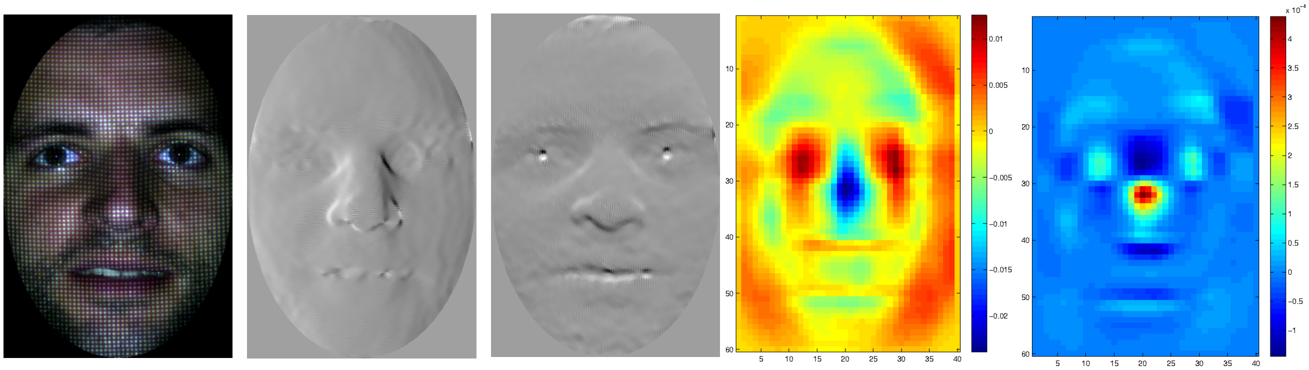
We projected a rectangular pattern over the face as shown in Fig.4(a). As described above, by filtering in the Fourier domain we obtained two different images with just horizontal and vertical fringes from which we computed depth map partial derivatives (Figs. 4(b) and 4(c)). Finally Mean and Gaussian curvatures were calculated (Figs. 4(d) and 4(e)). The last step was the isolation of those areas with large curvature values ( i.e. the thresholding step) and then by the HK-classifications candidates to nose and corners of the eyes were found.

Figure 5 shows the three-dimensional face surface, eye candidate points are displayed in green and the nose ones in red. The 3D reconstruction of the face was achieved by the integration of $D_x$ and $D_y$, the integration was performed only for illustrative purposes, it is not used in any of the processing steps.

The last set of experiments illustrated in Fig. 6 consist in testing the proposed framework in different conditions such as: different poses, expressions and faces partially covered by a scarf or a bonnet. In all the tested cases, nose and eye positions were detected (in conjunction with other false positive points). However, it is possible to remove the false positive detections by considering all the possible triplets (left eye, right eye and nose) and removing those with abnormal proportions, distances and orientations as was suggested in [1].

## 4. CONCLUSION

A novel framework for 3D curvature analysis was proposed showing that it is possible to perform three-dimensional feature extraction without the need of 3D reconstruction. Instead we propose an approach that is capable of providing depth gradient maps using very low hardware resources. Besides we used a commercial projector in the experiments presented (because it was what we have available), it is clear that just a modified flash can be used (as the pattern is binary and static).



(a) Input image (b) $D_x$ (c) $D_y$ (d) Mean curvature ($H$) (e) Gaussian curvature ($K$)

**Fig. 4**.

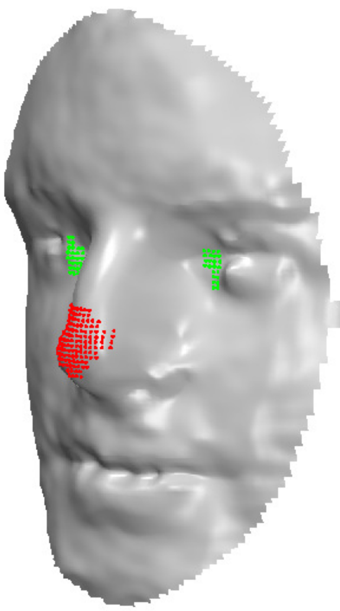


**Fig. 5**. Reconstructed 3D surface and landmarks detected.

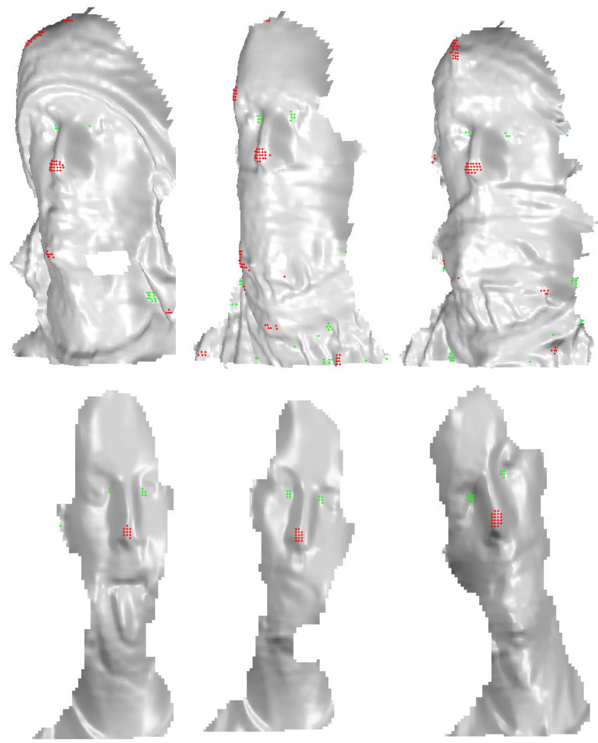


**Fig. 6**. Results under different poses, expressions and different kinds of face occlusion (using bonnet and/or scarf).

Also this kind of illumination pattern can be generated using a led source like the ones already available in most smartphones. We presented the main steps to compute the gradient depth map from the input image as well as some implementation details that must be taken into account. Results are very promising and we think the proposed method is suitable for many commercial and academical applications due to its simplicity and robustness.


# Acknowledgements

The authors thank Alvaro Gómez for fruitful discussions, PEDECIBA (Uruguay), the Comisión Sectorial de Investigación Científica de la Universidad de la República (CSIC, UdelaR, Uruguay) and Agencia Nacional de Investigación Científica (ANII, Uruguay) for their financial support.


## 5. REFERENCES


[1] Alessandro Colombo, Claudio Cusano, and Raimondo Schettini, "3D face detection using curvature analysis," *Pattern Recognition*, vol. 39, no. 3, pp. 444–455, Mar. 2006.

[2] Hengliang Tang, Baocai Yin, Yanfeng Sun, and Yongli Hu, "3D face recognition using local binary patterns," *Signal Processing*, vol. 93, no. 8, pp. 2190–2198, Aug. 2013.

[3] Yinjie Lei, Mohammed Bennamoun, and Amar a. El-Sallam, "An efficient 3D face recognition approach based on the fusion of novel local low-level features," *Pattern Recognition*, vol. 46, no. 1, pp. 24–37, Jan. 2013.

[4] Dirk Smeets, Johannes Keustermans, Dirk Vandermeulen, and Paul Suetens, "meshSIFT: Local surface features for 3D face recognition under expression variations and partial data," *Computer Vision and Image Understanding*, vol. 117, no. 2, pp. 158–169, Nov. 2012.

[5] Mauricio P. Segundo, Chaua Queirolo, Olga R.P. Bellon, and Luciano Silva, "Automatic 3D facial segmentation and landmark detection," *14th International Conference on Image Analysis and Processing (ICIAP 2007)*, , no. Iciap, pp. 431–436, Sept. 2007.

[6] Hong Kong, Wei Jen Chew, Kah Phooi Seng, and Li-minn Ang, "Nose Tip Detection on a Three-Dimensional Face Range Image Invariant to Head Pose," vol. I, pp. 0–4, 2009.

[7] Xiaoming Peng, Mohammed Bennamoun, and Ajmal S. Mian, "A training-free nose tip detection method from face range images," *Pattern Recognition*, vol. 44, no. 3, pp. 544–558, Mar. 2011.

[8] Jason Geng, "Structured-light 3d surface imaging: a tutorial," *Adv. Opt. Photon.*, vol. 3, no. 2, pp. 128–160, Jun 2011.

[9] Hongwei Guo, Haitao He, and Mingyi Chen, "Gamma correction for digital fringe projection profilometry," *Appl. Opt.*, vol. 43, no. 14, pp. 2906–2914, May 2004.

[10] Gastón A Ayubi, Jaime A Ayubi, J Matías Di Martino, and José A Ferrari, "Pulse-width modulation in defocused three-dimensional fringe projection," *Optics letters*, vol. 35, no. 21, pp. 3682–3684, 2010.

[11] Gastón A. Ayubi, J. Matías Di Martino, Julia R. Alonso, Ariel Fernández, Jorge L. Flores, and José A. Ferrari, "Color encoding of binary fringes for gamma correction in 3-d profiling," *Opt. Lett.*, vol. 37, no. 8, pp. 1325–1327, Apr 2012.

[12] R. I. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge University Press, ISBN: 0521540518, second edition, 2004.

[13] Mitsuo Takeda and Kazuhiro Mutoh, "Fourier transform profilometry for the automatic measurement of 3-d object shapes," *Appl. Opt.*, vol. 22, no. 24, pp. 3977–3982, Dec 1983.

[14] Song Zhang, "Recent progresses on real-time 3d shape measurement using digital fringe projection techniques," *Optics and Lasers in Engineering*, vol. 48, no. 2, pp. 149 – 158, 2010.

[15] Gastón A Ayubi, J Matías Di Martino, Jorge L Flores, and José A Ferrari, "Binary coded linear fringes for three-dimensional shape profiling," *Optical Engineering*, vol. 51, no. 10, pp. 103601–1, 2012.

[16] J.Matías Di Martino, Alicia Fernández, Gastón A. Ayubi, and José A. Ferrari, "Differential 3d shape retrieval," *Optics and Laser in Engineering*, 2014 (accepted for publication, 2/2/2014).

[17] M.P. do Carmo, *Differential Geometry of Curves and Surfaces*, Pearson Education Canada, 1976.

[18] Paul J. Besl and Ramesh C. Jain, "Invariant surface characteristics for 3d object recognition in range images," *Computer Vision, Graphics, and Image Processing*, vol. 33, no. 1, pp. 33 – 80, 1986.

[19] Gaile G. Gordon, "Face recognition based on depth maps and surface curvature," *SPIE 1570, Geometric Methods in Computer Vision*, pp. 234–247, Sept. 1991.

[20] AB Moreno, A Sánchez, JF Vélez, and FJ Díaz, "Face recognition using 3D surface-extracted descriptors," *Irish Machine Vision and . . .*, 2003.