A low cost 3D capturing system: an application for antlion pit architecture analysis

Arnold Fertin[1*] and Jérôme Casas[1]

[1]Université de Tours, IRBI UMR CNRS 6035, Parc Grandmont, 37200 Tours, France

[*]Corresponding author: arnold.fertin@etu.univ-tours.fr

1. Set-up, description and principle

The principle involves projecting the shadow of a plane with a rectilinear edge on the surface and using the deformations of the shadow to estimate the three-dimensional points corresponding to this surface. The system has two parts (Fig. 1A): 1) an apparatus with a light source (Orbitec P80818) fixed above the plane used to generate the shadow; 2) a camera (Euromex VC3031) placed above the apparatus. The apparatus is moved on ballbearings along a fixed axis on a flat surface. Displacements are made millimetre-by-millimetre above the object to be scanned. The equipment is set up such that the axis along which the apparatus moves is parallel to the X_{px} axis of the image. The edge of the plane therefore moves parallel to the Y_{px} axis of the image, making it possible to simplify implementation (Fig. 2). The images comprising the scan are coded on a grey-scale level from 0 (white) to 255 (black) (Fig. 2A). The camera captures the deformations of the shadow, which are then digitised using a graphics card (ADSTech DVDXpress) for computer processing. A point on the edge of the plane P is projected on a point S on the shadow on the surface of the object (Fig. 1). The camera and the graphics card produce a digital image of the scene on which we find the points S_{px} and P_{px} (Fig. 2A). The method used involves estimating the three-dimensional co-ordinates of S (the shadow) by reversing the process of image formation from the real scene to the digital image. These calculations lead to the generation of a 3D model of the object, as shown in Fig. 3.

2. Three-dimensional modelling of an object

2.1 Data extraction from images

The scan of an object comprises a series of images that record successive deformations of the shadow on the object. Each image is transformed into a binary image using the ImageJ program, with a threshold value giving a coded black (255) and white (0) image (Abramoff *et al.,* 2004) (Fig. 2B). The co-ordinates of the points forming the edge of the shadow and the edge of the plane were extracted from each image, using an algorithm developed in ImageJ. Each point on the edge of the projected shadow $S_{px}(Xs_{px}, Ys_{px})$ is associated with a point on the edge of the plane $P_{px}(Xp_{px}, Yp_{px})$ (Fig. 2). It should be noted that for the set-up used, $Ys_{px} = Yp_{px}$. Processing of the entire image gave a series of paired points (S_{px},P_{px}).

2.2 Transformation of pixel data into metric data

The model of camera used was of the pinhole type (Heikkilä and Silvén, 1997). This type of camera model defines a normalised image plane, corresponding to a focal plane with a focal distance fixed at 1 (Fig. 1). A point in space within the field of the camera is projected onto the plane on which the image is formed, and the CCD sensor encodes the image in pixels. The transformation therefore involves finding the metric co-ordinates of the points (S_c,P_c) on the normalised image plane based on the pixel co-ordinates of the points (S_{px},P_{px}). The transformation of a point on the digital image into a point on the image in the normalised plane requires the use of the intrinsic parameters of the camera. These parameters were estimated by calibration of the camera with the Camera Calibration Toolbox for Matlab® (Bouguet). The transformation of a given point (X_{px}, Y_{px}) from the pixel image into a point (X_n, Y_n) on the normalised plane image is given by the following linear relationship:

$$\begin{bmatrix} X_{px} \\ Y_{px} \\ 1 \end{bmatrix} = K \begin{bmatrix} X_n \\ Y_n \\ 1 \end{bmatrix} \text{ with } K = \begin{bmatrix} f_x & \beta f_x & Xo_{px} \\ 0 & f_y & Yo_{px} \\ 0 & 0 & 1 \end{bmatrix} \quad [1]$$

The matrix K is referred to as the camera matrix. It includes all the parameters intrinsic to the pinhole model camera. The point (Xo_{px}, Yo_{px}) is the principal point in the pixel image, f_x and f_y correspond to focal distances, expressed in pixels, corresponding to the axes X_{px} and $Y_{px,}$ respectively. β is the skew coefficient defining the angle between the axes X_{px} and Y_{px}. The co-ordinates, in pixels, of the points extracted from the image were converted into metric co-ordinates on the normalised image plane with equation [1].

The model of camera also defines the distortion parameters that can be used to rectify the distortion caused by the objective lens. A second transformation is used to correct the points on the normalised image field. There are no analytical solutions for the correction of lenses based on distortion coefficients. We therefore used a recursive method (Melen, 1994; Heikkilä, 2000). On the normalised image plane, the corrected co-ordinates (X_c, Y_c) are approximated based on the distorted co-ordinates $(X_{n,} Y_n)$ as follows:

$$\begin{bmatrix} X_c \\ Y_c \end{bmatrix} = \begin{bmatrix} \dfrac{X_n - \left(2p_1 X_n Y_n + p_2\left(r + 2X_n^{\,2}\right)\right)}{1 + k_1 r + k_2 r^2 + k_3 r^3} \\ \dfrac{Y_d - \left(2p_2 X_n Y_n + p_1\left(r + 2Y_n^{\,2}\right)\right)}{1 + k_1 r + k_2 r^2 + k_3 r^3} \end{bmatrix} \text{ with } r = X_n^2 + Y_n^2 \quad [2]$$

where k_1, k_2 and k_3 are the coefficients of radial distortion and p_1 and p_2 are the coefficients of tangential distortion. This calculation is repeated 20 times, with $(X_{n,} Y_n)$ replaced by the (X_c, Y_c) values from the previous iteration in each case. These two transformations were carried out in succession: conversion of the pixel co-ordinates into co-ordinates on the normalised image plane followed by correction for lens distortion.

2.3 Surface reconstruction

The corrected points on the normalised image plane $S_c(Xs_c, Ys_c)$ and $P_c(Xp_c, Yp_c)$ were calculated from the image points $S_{px}(Xs_{px}, Ys_{px})$ and $P_{px}(Xp_{px}, Yp_{px})$. The points $S_c(Xs_c, Ys_c)$ and $P_c(Xp_c, Yp_c)$ are the projections of the spatial points $S(x_S, y_S, z_S)$ and $P(x_P, y_P, z_P)$ in the camera's reference frame:

$$\begin{bmatrix} z_S Xs_c \\ z_S Ys_c \end{bmatrix} = \begin{bmatrix} x_S \\ y_S \end{bmatrix} \text{[3] and } \begin{bmatrix} z_P Xp_c \\ z_P Yp_c \end{bmatrix} = \begin{bmatrix} x_P \\ y_P \end{bmatrix} \text{ [4]}$$

The aim was to define analytically the point on the edge of the shadow $S(x_S, y_S, z_S)$ in the camera's reference frame as a function of the co-ordinates of the points $S_c(Xs_c, Ys_c)$ and $P_c(Xp_c, Yp_c)$, the angle of incidence α of the light source with the axis x of the camera's reference, and the equation of the plane in space. The equation of the plane projected in the camera's reference frame is defined as follows:

$$z_P = ax_P + by_P + c \text{ [5]}$$

expressing z_p as a function of $P_c(Xp_c, Yp_c)$ based on equations [3], [4] and [5] :

$$z_P = \frac{c}{1 - aXp_c - bYp_c} \text{ [6]}$$

The projection of the point *P* on the point *S* by the light source gives (Fig. 1B):

$$\tan\alpha = \frac{z_S - z_P}{x_P - x_S} \text{ [7]}$$

The equations [3], [4], [6] and [7] can be used to define the three-dimensional co-ordinates of the point *S* in the camera's reference frame as a function of the points S_c and P_c and the parameters of the apparatus (α, a, b, c):

$$z_S = \frac{c(1 + Xp_c \tan\alpha)}{\left((1 - aXp_c - bYp_c)(1 + Xs_c \tan\alpha)\right)} \text{ [8], } x_S = z_S Xs_c \text{ [9] and } y_S = z_S Ys_c \text{ [10]}$$

The pairs of points (S_{px},P_{px}) extracted from the series of images were treated according to the process described above, to obtain a collection of three-dimensional points located on the surface of the scanned object. Akima's linear interpolation was used to reconstruct this surface using a grid with 0.5 mm squares (Akima, 1996).

2.4 Device calibration

The apparatus was calibrated and its parameters (α, a, b, c) estimated using a rod connecting two points A and B separated by a d_{ref} of 21.182 mm. The rod was placed in N different positions in the field of the camera (N=60). At each position i the shadow of the plan was projected successively on the two points A and B of the rod. For position i, the co-ordinates in the camera's reference frame $A_i(x_{Ai}, y_{Ai}, z_{Ai})$ and $B_i(x_{Bi}, y_{Bi}, z_{Ci})$ were calculated as described above. The distance d_i between A_i and B_i is calculated as follows:

$$d_i = \sqrt{(x_{Ai} - x_{Bi})^2 + (y_{Ai} - y_{Bi})^2 + (z_{Ai} - z_{Bi})^2} \quad [9]$$

This distance d_i is a function of the parameters (α, a, b, c) and equals d_{ref} if and only if $(\alpha, a, b, c)_{estimated} = (\alpha, a, b, c)_{real}$. The method for estimating (α, a, b, c) was based on least squares minimisation:

$$RSS(\alpha, a, b, c) = \sum_{i=1}^{i=N} (d_{ref} - d_i)^2 \quad [10]$$

This estimation was carried out in the R environment (R Development Core Team 2004), using a quasi-Newtonian method (Byrd *et al.,* 1995).

3. System precision

Table 1 shows estimates of the parameters of the apparatus and their precision. The director coefficients *a* and *b* are very close to zero, because, in the set-up used, the camera is positioned with respect to both the horizontal and the projected plane.

The precision of the apparatus was estimated using the same rod used for calibration. This rod was placed in 60 positions different from those used for calibration and, for each position *i,* the distance d_i between the points A and B was calculated. The difference between d_i and d_{ref} depended on the exact repositioning of the points A and B in space. At each position, we calculated the absolute error: $\varepsilon_i^a = \left| d_{ref} - d_i \right|$ and the relative error: $\varepsilon_i^r = \frac{\left| d_{ref} - d_i \right|}{d_{ref}} \times 100$. The absolute error was ε^a=98.0737±18.5969 μm (mean ± 95% confidence interval), corresponding to a maximum error of 117 μm in the calculation of the distance between the two points. The relative error was low: ε^r=0.4630 ± 0.0878 % (mean ± 95% confidence interval), or a maximum of 0.55%. In the context of this study, the precision of the instrument corresponds to less than the diameter of a grain of sand.

4. Conclusion

This apparatus provides high-level precision at low cost (about $40) and made it possible to analyse in detail the geometry of antlion traps (Fig. 4). This technique is competitive with respect to other available techniques. Profilometric lasers are much more expensive. The simplicity of the calculations with our method makes it possible to generate a three-dimensional model with a large number of points. Photogrammetry, which requires correlation between the images of at least two cameras, may become particularly cumbersome if a large number of points are required.

The principal limitation of our technique concerns the texture of the object used. As image analysis is based on the detection of a shadow, a texture giving an image with too many shadow pixels could result in too many erroneous points. However, this could readily be corrected by dusting the object with white powder or using a more powerful white light. Image processing with contour detection algorithms could also be used to overcome these problems of “dark” textures. It should be noted that lasers also face similar problems when confronted with dark or shiny surfaces.

References

Abramoff, M.D., Magelhaes, P.J., Ram, S.J. (2004). Image Processing with ImageJ. Biophotonics International, **11** (7), 36-42.

Akima, H. (1996). Algorithm 761: scattered-data surface fitting that has the accuracy of a cubic polynomial. ACM Transactions on Mathematical Software, **22**, 362-371.

Bouguet, J.Y. Camera Calibration Toolbor for Matlab ®, http://www.vision.caltech.edu/bouguetj/calib_doc/.

Byrd, R. H., Lu, P., Nocedal, J. and Zhu, C. (1995). A limited algorithm for bound constrained optimization. SIAM J. Scientific Computing, 16, 1190-1208.

Heikkilä, J. & Silvén, O. (1997). A Four-step Camera Calibration Procedure with Implicit Image Correction. IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'97), San Juan, Puerto Rico, p. 1106-1112.

Heikkilä, J. (2000). Geometric camera calibration using circular control points. IEEE Transactions on Pattern Analysis and Machine Intelligence, **22** (10), p. 1066-1077.

Melen, T. (1994). Geometrical modelling and calibration of video cameras for underwater navigation. Dr. Ing. Thesis, Norges Tekniske Hogskole, Institute for Teknisk Kybernetikk.

R Development Core Team (2004). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Tables and figures

Parameters	Estimates	Standard deviation
α	0.9580	0.0126
a	0.0186	0.0165
b	0.0071	0.0150
c	484.6556	3.7518

Table 1: Estimates of the parameters of the apparatus. Standard errors were estimated by inversing the Hessian matrix obtained when $RSS(\alpha, a, b, c)$ was minimised.

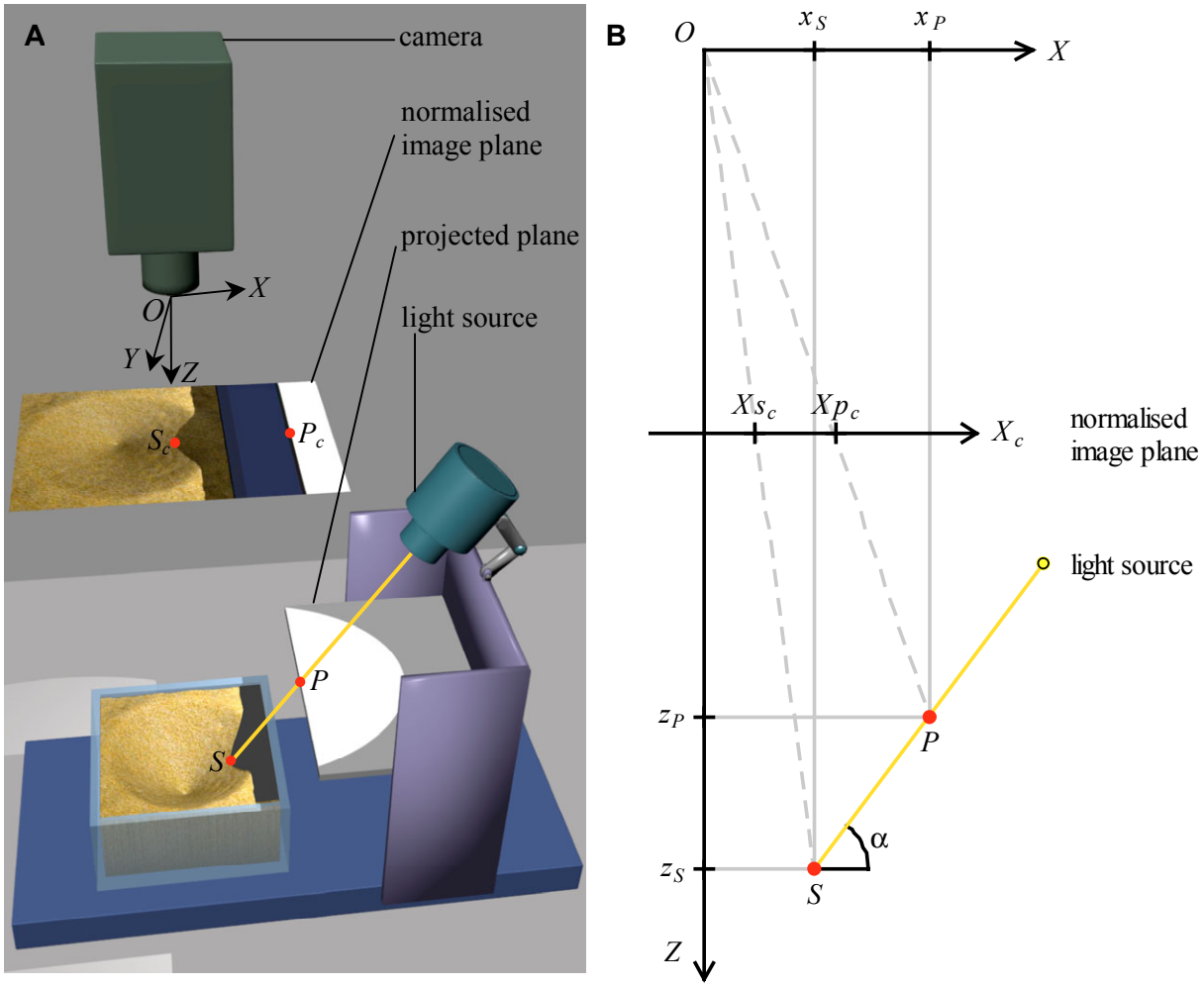

Figure 1: (A) Diagram of the set-up. The light source projects a shadow of the edge of the plane of the scene. The edge of the plane and the shadow are projected onto the normalised plane of the camera, and the resulting image is used to reconstruct the three-dimensional scene in the camera's reference frame $O(X,Y,Z)$. (B) Diagram explaining the calculation of $S(x_S, y_S, z_S)$ in the plane (X,Z) of the camera's reference frame, as $y_s=y_p$. O is the optical centre of the camera and Z is the optical axis.

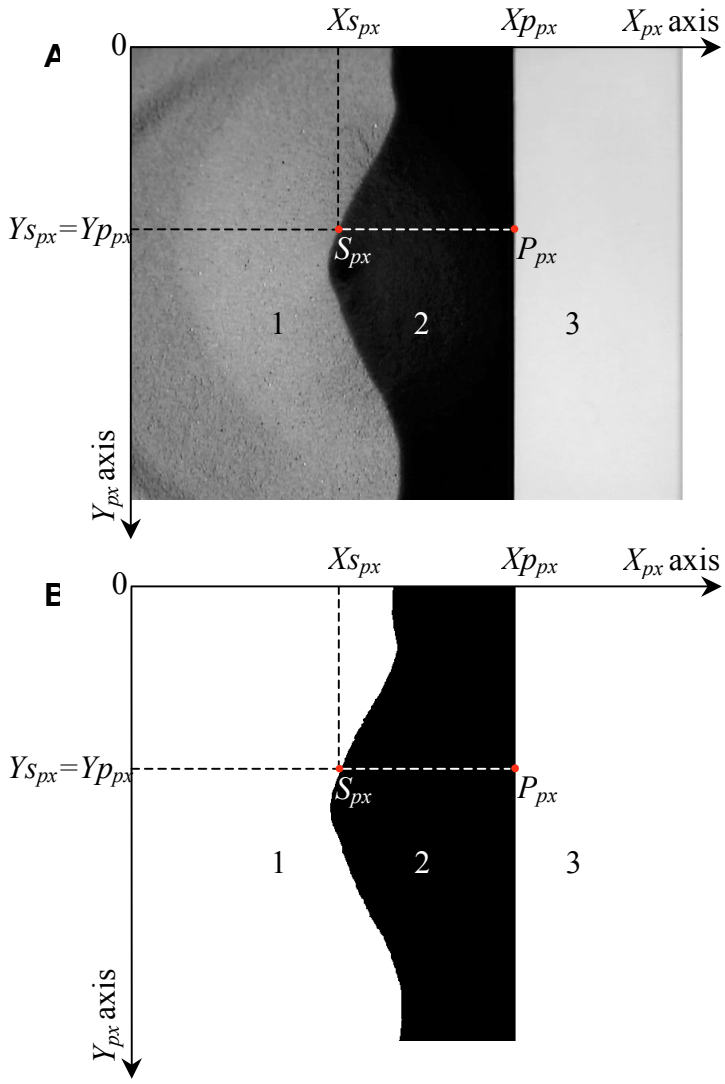

Figure 2: (A) Image extracted from the video obtained with the camera. (B) The same image after binary transformation. 1: Scanned surface, here an antlion trap; 2: Shadow; 3: Projected plane. The algorithm developed in ImageJ processes the pairs of pixel co-ordinates ($S_{px,}P_{px}$).

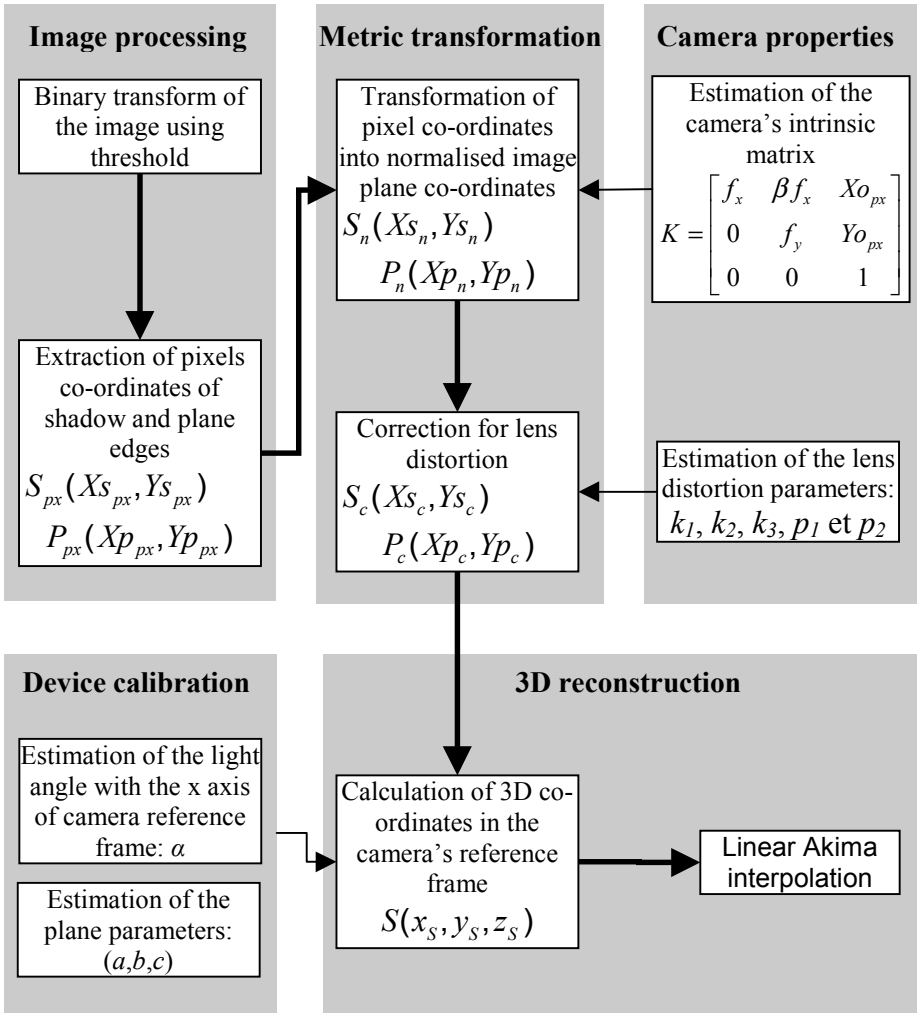

Figure 3: Diagram of the various steps in the three-dimensional modelling of an object. The thick arrows show the steps to be followed when reconstructing an object.

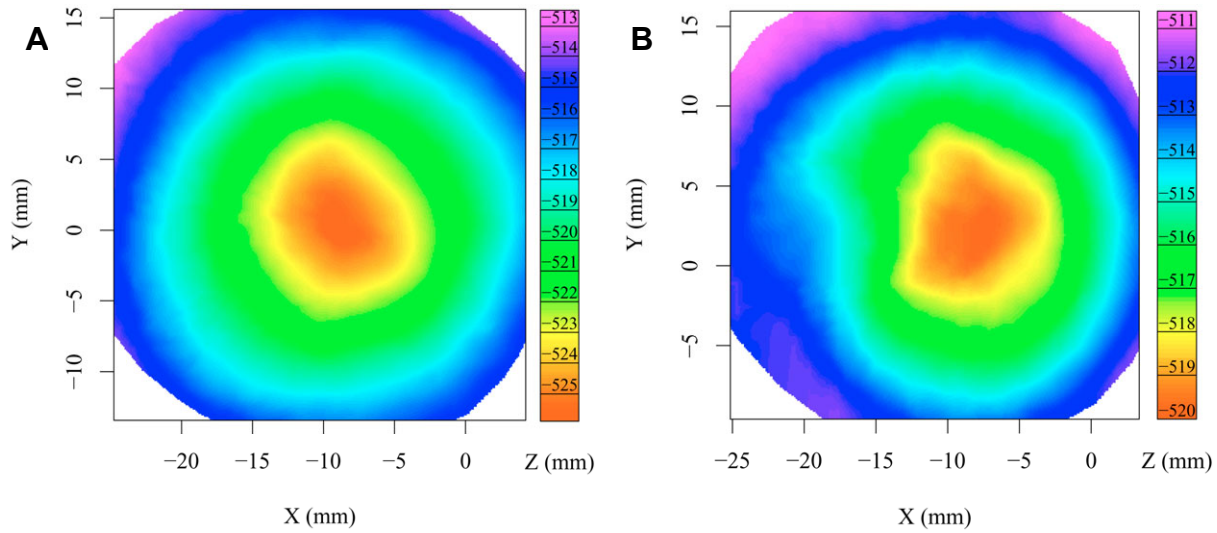

Figure 4: Examples of two antlion trap topographies. (A) Trap with a low off-centring. (B) Trap with a high off-centring.