Supplemental information for "A particle filter to reconstruct a free-surface flow from a depth camera" : Experiments on the Kinect sensor

Benoit Combès, Dominique Heitz, Anthony Guibert, Etienne Mémin

1 Introduction

Capturing the geometry of a 3D object is of great interest for many applications. Different strategies (*e.g.* stereovision, shape from deformations) handling different technologies (*e.g.* binocular system, structured light) have lead to a large variety of sensors having different characteristics (we report the reader to the article of Herbort and Wöhler (2011) for an introduction to depth sensors). In this work, we choose to use the Kinect senor. This depth sensor has the advantage to be cheap (about 150$), to provide high-level programming interfaces and not to need any specific calibration. These three qualities make it a good candidate for practical use. It is partly composed of a 30 Hertz 320×240 RGB sensor, a 30 Hertz 320×240 infrared sensor and an infrared pattern projector. Range images are obtained from the so-called light-coding technique: the projection of the infrared pattern on the object under study is captured by the infrared sensor and the analysis of this projection is used to recover the geometry of the object. This approach is close to the optical profilometry technique proposed by Cobelli, Maurel, Pagneux and Petitjeans (2009) for the measurement of water waves and in practice provides range images of 640×480 pixels each 0.03 second. To the best of our knowledge, the Kinect sensor is not well-documented and the main available information are shared by users from the web. In this work, we do not focus on the technology of the Kinect and of its subsequent limitations and facilities but use it, as best as we can, in a "black-box" way. For that purpose, we propose to investigate its accuracy to estimate range images of solid and liquid smooth surfaces. For a sake of simplicity, we restrict our study to surfaces located between 680 and 780 mm from the device. However, note that the devise is able to estimate depth data up to 13 meters leading to an observable surface of about 10m × 14m with a magnification of about 2 cm near the optical axis and 20cm at the periphery of the observable window.

2 Modeling

In this subsection, we point-out a few modeling choices and their implications.

Calibration law: It is not clear how the raw data r computed from the Kinect are related to the true distance z to the sensor. To answer this question empirically, we acquired a set of range data for known distance-to-sensor z. For this purpose, we set the camera at a given position and fix the optical axis orthogonally to the horizontal, then we fix a flat board at a given distance z and we acquire the corresponding raw data r (Fig. 1-left). This process is repeated by moving the observed flat surface along the elevation axis using calibrated spacers (Fig. 1-center). We plotted z as a function of r in Fig. 1-right. These results suggest the need to correct the raw values using a tangent law of the form $z = a\tan(br+c)+d$. We estimated the parameters (a, b, c, d) by minimising the sum of the square errors between the actual distances z and the estimated distances $a\tan(br+c)+d$ using a gradient descent algorithm. The estimation leads to a residual error of 1.13mm^2.

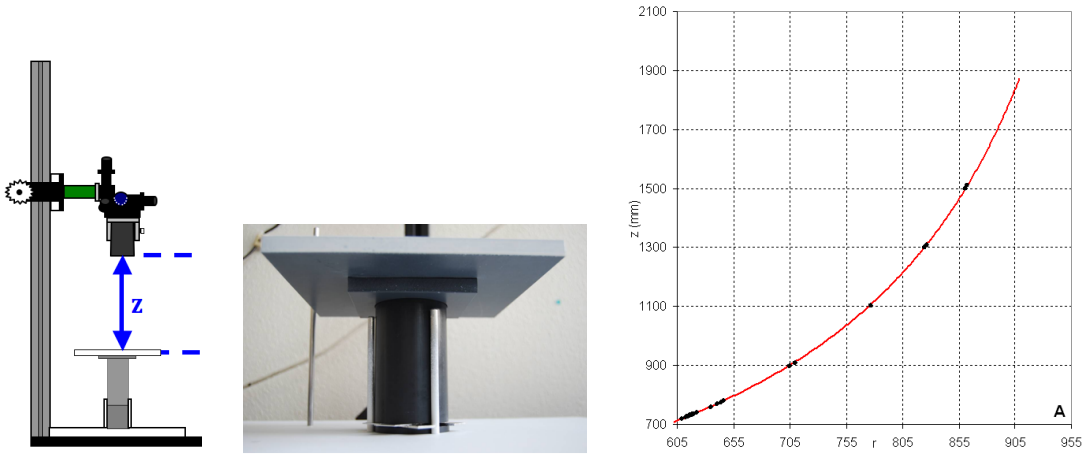

Figure 1: **Calibration law of the Kinect** From left to right: i) schema of the experimental set-up, ii) location of the metal spacers used to modify the z value and iii) plot of $z = f(r)$ (black dots) and of the tangeant calibration law (red curve).

Projection model: We modeled the geometry of the sensor as a projection model in its simplest form which depends only on a focal length that we estimated equal to 587 pixels.

Quantisation error: The raw data given by the sensor are encoded on 11-bits. We computed the mean quantisation error as the mean error between the true distance d and its nearest value given by the quantification process *i.e.* $\frac{1}{z_{max}-z_{min}}\int_{z=z_{min}}^{z_{max}} \min_{r\in N} |a\tan(br+c)+d-z|dz$ which was 0.6 mm in the range of interest (680 to 780 mm). Concerning the horizontal dimension, the magnification coefficient was about one mm per pixel near the

optical axis.

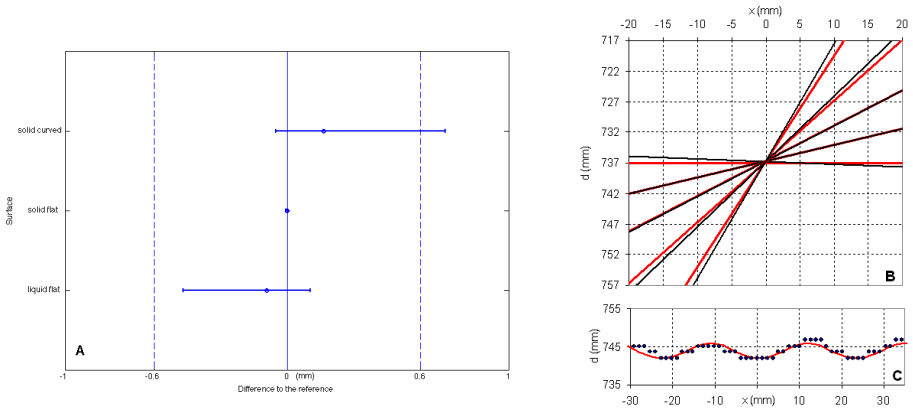

Figure 2: **Experimental characterisation of the sensor.** A: Test of equivalence between errors from flat and curved surface and from solid and liquid surfaces; the errors measured from the flat solid surface are considered as the reference. Departures of errors from curved surfaces and liquid surfaces and their 95% confidence intervals are plotted together with the quantification error bar (0.6mm). B: estimation of slopes from 3 points, red lines indicate actual slopes and black ones indicate estimated slopes. C: estimation of a sinusoidal surface, the red curves indicate the actual sinusoid and the black dots indicate the z estimated from Kinect data.

3 Experimental characterisation

3.1 Experimental characterisation with solid surfaces

Pixel-wise error: We acquired a set of solid surfaces (from flat to quickly varying sinus-like surfaces) at different elevations and then we used the calibration law to estimate the surface-sensor distances. We obtained a mean error of estimation was from 0.89 mm for flat surface to 1.05 mm for quickly varying surfaces.

The t-test with H_0: "errors coming from observations of curved and flat surfaces do not belong to the same normal population" and $p - value = 0.1$ fails and the plot box (Figure 2) illustrating an equivalence testing using the errors from the flat surface as a reference does not allow us to conclude on equivalence (or not) of the two errors (using a zone of indifference of $\pm$ quantisation error=0.6mm).

Spatial repartition of the error: To investigate the spatial repartition of the errors in the depth map, we analysed the values of the spatial autocorrelation of the measurement error (Moran's I). This analysis does not allow us to give any conclusion. Similarly, a qualitative analysis of the errors with respect to its distances to the optical axis was not

concluding. We think that this question is difficult to answer because of the discrete nature of the data r and of the high value of the quantification error (about 0.6 mm) with respect to the actual measurement error.

Shape recovering: We investigated the ability of the sensor to recover known geometrical patterns. In particular, we qualitatively noticed that the the sensor is able to recover surfaces with steep slopes and to capture wave-like patterns of small spatial periods (< 20 mm) and small amplitudes (< 2 mm). Figure 2.B and 2.C show examples of estimated point sets.

3.2 Experimental characterisation with liquid surfaces

Since water was used as a working liquid, liquid's light diffusivity was enhanced by the addition of white dye. We qualitatively noticed that errors coming from observations of solid and liquid surfaces looks comparable when the attenuation coefficient of the liquid is larger than 113 m^{-1}. Moreover, the t-test with H_0: "errors coming from observations of liquid and solid flat surfaces does not belong to the same normal population" and $p - value = 0.1$ fails and the plot box (Figure 2) illustrating an equivalence testing using the errors from the flat surface as a reference suggests an equivalence of the two errors (using a zone of indifference of $\pm$ quantisation error=0.6mm). We noticed that when observing bright surfaces (which includes the water surfaces we studied), the sensor is unable to provide a depth data for some pixels. For an acquisition of a moving water surface in laboratory, this phenomenon affected about 1% of the pixels.

References

Cobelli, P., Maurel, A., Pagneux, V. and Petitjeans, P.: 2009, Global measurement of water waves by fourier transform profilometry, *Experiments in Fluids* **46**, 1037–1047.

Herbort, S. and Wöhler, C.: 2011, An introduction to image-based 3d surface reconstruction and a survey of photometric stereo methods, *3D Research* **2**, 1–17.