

The Kinect: a low-cost, high-resolution, short-range 3D camera

Kenneth David Mankoff* and Tess Alethea Russo

Department of Earth and Planetary Sciences, University of California, Santa Cruz, Santa Cruz, California USA

Received 11 January 2012; Revised 14 September 2012; Accepted 19 September 2012

*Correspondence to: Kenneth David Mankoff, Department of Earth and Planetary Sciences, University of California, Santa Cruz, 1156 High Street, Santa Cruz, CA 95060 USA. E-mail: kdmankof@ucsc.edu

ESPL

Earth Surface Processes and Landforms

ABSTRACT: We present a novel application of the Kinect™, an input device designed for the Microsoft® Xbox 360® video game system. The device can be used by Earth scientists as a low-cost, high-resolution, short-range 3D/4D camera imaging system producing data similar to a terrestrial light detection and ranging (LiDAR) sensor. The Kinect contains a structured light emitter, an infrared camera (the combination of these two produce a distance image), a visual wavelength camera, a three-axis accelerometer, and four microphones. The cost is ~US \$100, frame rate is 30 Hz, spatial and depth resolutions are mm to cm depending on range, and the optimal operating range is 0.5 to ~5 m. The resolution of the distance measurements decreases with distance and is ≤ 1 mm at 0.5 m and ~75 mm at 5 m. We illustrate data collection and basic data analysis routines in three experiments designed to demonstrate the breadth and utility of this new sensor in domains of glaciology, stream bathymetry, and geomorphology, although the device is applicable to a number of other Earth science fields. Copyright © 2012 John Wiley & Sons, Ltd.

KEYWORDS: bathymetry; digital elevation model (DEM); sensor; structured light; geomorphology

Introduction

Measuring and monitoring surfaces and objects in three dimensions (3D) is critical to many studies in the Earth sciences. Methods for generating a digital elevation model (DEM) of a surface or a 3D scan of an object include airborne (Allouis *et al.*, 2010) and terrestrial (Hodge *et al.*, 2009b) time-of-flight light detection and ranging (LiDAR), laser scanning (Lane *et al.*, 2000; Smart *et al.*, 2004; Schaefer and Inkpen, 2010), interferometric synthetic aperture radar (InSAR/lfsAR) (Milledge *et al.*, 2009), white light interferometry (Candela *et al.*, 2011), photogrammetry (Lane, 2000, Lane *et al.*, 2000; Baily *et al.*, 2003; Brasington and Smart, 2003), GPS surveys (Chassereau *et al.*, 2011), total station surveys (Hubbard and Hubbard, 1998), and stake placement (Furbish, 1987). The Microsoft® Kinect™, a ~US \$100 input device for the Xbox 360® video game system, is capable of providing high spatial (x, y) resolution (~2–20 mm) and high distance (z) resolution (1–75 mm) 3D data. The system also offers high repeat measurement frequency (30 Hz), real-time feedback, and a simple data collection approach, while being cheaper, lighter, and smaller than equipment commonly used to collect similar data.

Affordable, commercial, off-the-shelf hardware has had unintended but useful applications as research tools in other domains in the past. For example, Hut *et al.* (2010) show that the Nintendo® Wii™ is a useful sensor for hydrologic studies, and Tribelhorn and Dodds (2007) demonstrate that the Roomba robotic vacuum can be used as a general robotics platform. Low-cost hardware also allows opportunities for science, education, and training in environments such as developing countries and schools (Welch and Dikkers, 1978).

Since its release in November 2010 the robotics and computer vision communities have made significant progress in developing software to make the Kinect a usable research instrument. The knowledge acquired through these efforts is available online through wikis, discussion groups, and web sites such as <http://openkinect.org>, which hosts the 'libfreenect' project. This paper benefits greatly from the individuals posting on those sites.

The focus of this paper is to introduce the Kinect as an Earth science research tool by describing its specifications, limitations, and example applications. We describe the hardware and software interface and discuss the quality of the data collected by the instrument, and its limitations. We then present the first published examples using the Kinect in the Earth science domain, demonstrating uses in glaciology, stream bathymetry, and geomorphology. The appendices give a step-by-step tutorial for data collection and sample code to access the raw data, supporting custom algorithm creation. A digital supplement (<https://github.com/mankoff/libfreenect/>) contains source code for calibration and registration of the raw data.

The Kinect

The Kinect (Figure 1) detects the distance from itself to objects within its field of view (FOV) by emitting a known pattern of infrared (IR) dots with a projector, and recording that pattern with an IR camera. The pattern changes with distance, expanding radially from the IR emitter point source until it is displayed on the surface of interest. Figure 2 shows the pattern of dots projected onto a flat wall with the author's hand visible.



Figure 1. Kinect with external case removed showing the three primary sensors: IR structured light projector (left), RGB camera (center left), and IR detector (center right).

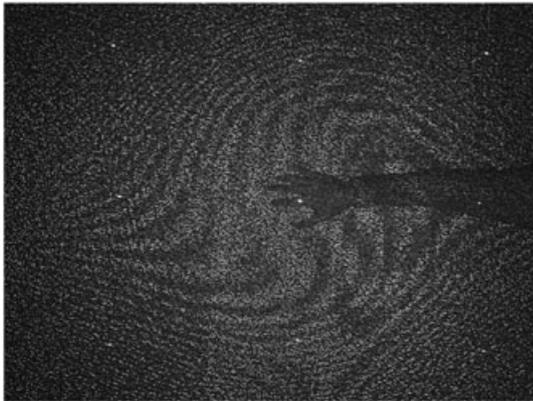


Figure 2. The infrared (IR) pattern projected by the IR emitter and recorded by the IR camera on the Kinect is seen displayed on a flat wall. The author's hand is visible, deforming the pattern. The deformation between the (known) projected and recorded pattern is used onboard the Kinect to derive distance.

The Kinect resolves the distance to the arm and the wall by comparing the projected pattern to the pattern recorded on the arm, and the pattern recorded on the wall. Using this method, the difference between the known projected and recorded patterns are used to construct a 3D distance map of the scene (e.g. patent by Freedman *et al.*, 2010).

Hardware specifications

The core of the Kinect is a chip designed by PrimeSense™, a company that pioneered the low-power high-speed 3D camera technique (PrimeSense Ltd, 2011) licensed by Microsoft and used in the Kinect. The Kinect contains a three-channel red, green, and blue (RGB) image camera, an IR camera, and an IR laser and refractor that emits a known structured light pattern at 830 nm, with the laser temperature-stabilized by a Peltier element (OpenKinect, 2011). The IR camera records to an internal sensor of 1280 × 1024 (5.2 mm) pixels with an FOV of 58 horizontal by 40 vertical degrees (Micron, 2011). The data are sent to a computer over standard USB as a 640 × 480 pixel image (PrimeSense Ltd, 2011). The Kinect operates in a continuous asynchronous mode, recording data and writing to an internal buffer at ~30 Hz, collecting over 9 million pixels per second. Basic geometric calculations show a spatial footprint of ~0.7 mm² at 0.5 m (LiDAR equivalent of 0.7 mm spot spacing, or ~204 000 points per square meter), ~1.4 mm² at 1 m, and ~7 mm² at 5 m. Due to the pattern recognition technique used to determine distance, the spatial resolution is reduced by a factor of 3, or at best approximately 2 mm.

The Kinect also contains an accelerometer, tilt motor, and a microphone array. The accelerometer (a Kionix KXS9) has a

range of ±2 g, a resolution of 819 counts per g, records data at ~200 Hz, and its purpose is to determine orientation and auto-level the Kinect with the help of the tilt motor. The microphone array records each channel at 16 kHz and can be used to determine sound source direction (OpenKinect, 2011).

The Kinect is powered by a 12 V direct current at ~1 A. The power cord, containing the AC/DC converter, can be cut and the Kinect wired directly to a 12 V battery (or eight AA batteries) for use in the field.

Software interface

A large and growing body of software aimed at communicating with the Kinect is freely available. Many programs return x, y, z point clouds or mesh surfaces in real-world units with RGB metadata and write file formats indistinguishable from commercial LiDAR devices.

These user-level programs are built on top of one of three low-level interfaces. First, the official interface is the Kinect software development kit (SDK) released by Microsoft. It requires Windows 7 and is designed for high-level user interaction such as detecting and tracking human bodies and motion. Second, the OpenNI+SensorKinect SDK is open-source code that PrimeSense has released, and is also targeted toward higher-level user interaction with the Kinect. Third, the OpenKinect community maintains 'libfreenect', a low-level open-source library that can easily be used to record the raw sensor data. Both OpenNI and libfreenect work on Windows, Linux, and Mac OS X. We use and discuss the OpenKinect libfreenect library and example programs that are distributed with that code during the remainder of this paper, as this software provides the easiest access to the raw data stream at 30 Hz.

We acquire data at the lowest possible input level using the libfreenect 'record' program, which returns raw sensor internal digital number (DN). Because we store the raw data, we can perform improved calibrations as better conversion functions are released. The 'record' program writes two 640 × 480 pixel images (one RGB, one distance) to disk at 30 Hz. The RGB file is 24-bit PPM format and the distance image is PGM (big-endian) format with DN between 0 and 2047. Appendices A and B include step-by-step instructions for collecting and processing data to a more user-friendly coordinate system and file format, and sample code to access the raw data.

Data quality

Defining the terms used throughout this section and remaining document, 'calibration' refers to the mapping from raw (pixel, pixel, DN) coordinates to world (x, y, z) coordinates, and the term 'registration' is the alignment of the RGB image to the depth image. For an object at a known distance, repeating the distance measurement will result in a distribution of measured

values centered on a mean value. We use the term ‘accuracy’ to describe the difference between a real value as measured with a tape measure and the mean Kinect-measured value of the distribution (or a specific value when not averaging data). This value is particularly relevant for absolute distance measurement. We use the term ‘precision’ to characterize the spread of the measurement around the mean value.

Calibration and registration

The Kinect includes factory-supplied parameters that we use to calibrate and register the data. In addition, a variety of third-party methods exist to calibrate a Kinect (Burrus, 2010; Herrera *et al.*, 2011; Konolige and Mihelich, 2011; Smisek *et al.*, 2011). We use the factory-supplied calibration parameters for our examples as it requires minimal additional work and a full comparison of third-party calibration choices is beyond the scope of this paper. We provide a command-line utility in the digital supplement (<https://github.com/mankoff/libfreenect/>) for Linux and OS X computers that converts the raw (pixel, pixel, DN) data to world (x, y, z) coordinates, and optionally registers the RGB data on to the depth data. One drawback of using the factory-supplied calibration is that the DN-to-distance map uses integers, limiting the resolution to 1 mm at best, even if the sensor is capable of measurements at a slightly better resolution.

Calibration results are shown in Figure 3, with the line marked ‘DN’ showing the raw sensor values mapped to distance. The DN line is cut off at the minimum reliable distance of 0.5 m, although we have acquired data as near as ~ 0.4 m for part of the scene. The Kinect can detect surfaces at distances to 10 m, but errors are large at that range so we only show data to a maximum of 5 m. The step size is the difference between one DN and the next when mapped to world- z , and is discussed in more detail below (‘Precision’). Although the minimum resolution (minimum step size) presented here is 1 mm, at low values multiple DN actually map to the same mm (for example, both DN_n and DN_{n+1} map to a single mm value). Therefore, the resolution of the sensor is actually better than 1 mm at close range, even if it is not accessible by the calibration technique we use.

After calibrating and registering the data, each data point has an x, y, z, r, g, b value with x, y, z in physical space (with 0, 0, 0 being the point of the Kinect distance camera) and a red, green, and blue value in color space.

Accuracy

To test the accuracy of the calibration with respect to an independent distance measurement, we image a wall at a known distance and compare results, using a 20×20 pixel area in

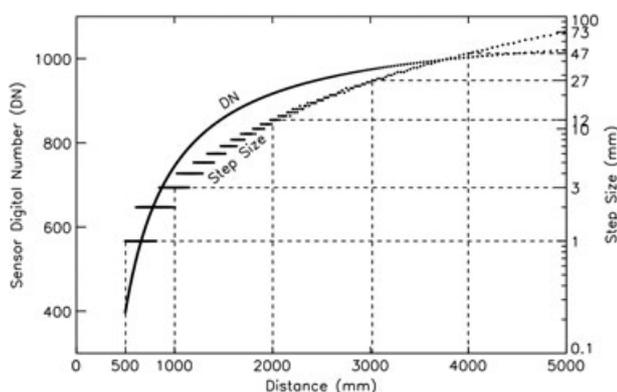


Figure 3. Calibration of digital number (DN) to distance, and resolution of distance measurement of a Kinect. Resolution is the step size between one DN and the next when converted to mm. Dashed lines highlight resolution (1–75 mm) at different distances (0.5–5 m).

the center of the image, and averaging ~ 100 frames of DN data. The average value is then passed through the calibration function to get distance in mm units. The results, presented in Table I, show that the calibration is accurate to within 1% of the independent reading, and often less than the resolution at the measured distance, for this central region of the sensor.

Next, we discuss the need for a full FOV two-dimensional ($2D, x, y$) treatment of the calibration, at least when measuring absolute distances. Figure 4 shows the average of 100 frames of DN when the Kinect is approximately perpendicular to a wall. To adjust for errors in the perpendicularity of the alignment, we convert the data to a point cloud, fit a plane to the points, and rotate the surface to perpendicular. The range of 5 DN shows there are 2D errors in the accuracy of the sensor. However, the sensor is precise and the accuracy errors are consistent in time, and Figure 5 shows that when repeating the measurement, and subtracting one set of measurements from another, the 2D error is removed. Therefore, this accuracy error can be ignored when performing change experiments, such as a difference-of-DEM (DoD). The bias pattern shown in Figure 4 is stable in time at a given distance (Figure 5), and varies slightly with changing distance (Smisek *et al.*, 2011), so these are not random errors or noise, but semi-stable biases that can be corrected in post-processing if necessary.

Given the above, we consider the accuracy of the device to be ± 1 DN when examining surface changes, and ± 3 DN

Table I. Comparison between the Kinect internal calibration parameters and a known distance. The ‘Actual’ column is a ruler-measured distance to a flat wall. The ‘Kinect’ column is a single value calculated by taking the average of a 20×20 pixel region at the center of ~ 100 frames. The error is calculated by the equation $e = \frac{|a-K|}{a}$, where a is actual and K is Kinect measured distance. The resolution is the difference between one DN and the next (in mm units) at the given range

Actual (mm)	Kinect (mm)	Error (%)	Resolution (mm)
500	501	0.20	1
1000	1002	0.20	3
2000	2003	0.15	12
3000	2993	0.23	27
4000	3998	0.05	47
5000	5050	1.00	75

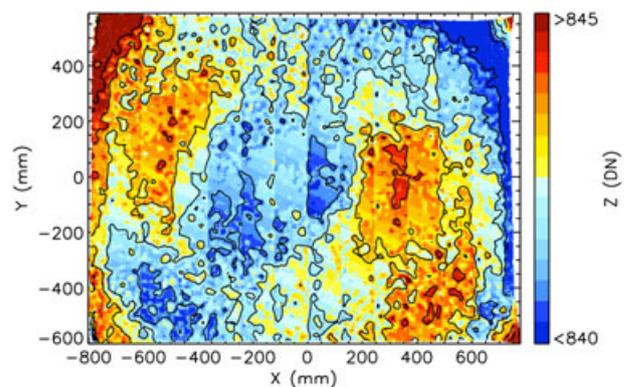


Figure 4. Two-dimensional (2D) errors in Kinect accuracy determined by imaging a flat wall. 100 images were used to increase the SNR. Contours are at intervals of 1 between 840 and 845 (DN) inclusive. This bias is stable in time for a given distance (as shown by Figure 5) and semi-stable with changing distance, so it can be quantified and mitigated in post-processing. This figure is available in colour online at wileyonlinelibrary.com/journal/espl

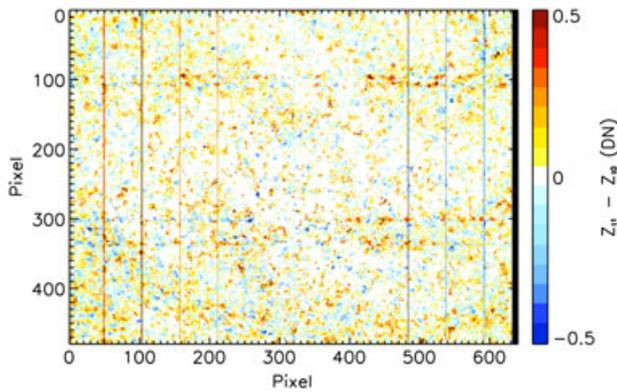


Figure 5. Difference of DEM (DoD) showing that the 2D bias shown in Figure 4 is stable in time and is therefore an issue of accuracy, not precision, and is not an issue when performing difference calculations. We imaged a flat wall from the exact same position at two different times and for each time averaged 100 frames and then subtracted mean values collected at time t_0 from the data at time t_1 . This figure is available in colour online at wileyonlinelibrary.com/journal/espl

(for the bulk of the scene, less for the center region) when measuring absolute distance, although the ± 3 DN can be improved by recognizing the noise is not random but is instead a quantifiable bias.

Herrera *et al.* (2011) present a method that captures the 2D bias shown above in their calibration equation, precluding the post-processing step suggested above. Their Kinect Calibration Toolbox (KCT) does not use the intrinsic camera parameters supplied by the manufacturer, instead determining all values extrinsically. It supports a two-dimensional calibration where each of the 640×480 pixels is treated. In addition, this code has the ability to determine registration parameters for an external camera such as an SLR mounted on a Kinect (supporting RGB data at much higher resolution).

The most accurate calibration possible for the Kinect would be a 'brute force' technique that produces a DN-to-distance lookup table (LUT) for all DN at each pixel. The method is as follows. Precisely align the Kinect perpendicular to a wall that fills its FOV. For each step (each DN or some other nonlinear spacing) from 0.5 to 5 m, record the full frame (record multiple frames and average them to increase the signal-to-noise ratio (SNR)). Acquire an independent measurement of the distance from the sensor to the wall. Record the average DN value of each x, y pixel and the known distance. The resultant LUT is $640 \times 480 \times$ the number of steps, each element storing an 8-bit floating point number (units: m).

Precision

If the phenomenon being studied is slow compared to the 30 Hz refresh rate, one can average together frames and increase the SNR. We check precision by recording the same scene for 100 frames and computing the standard deviation of each pixel across all images, showing the full frame in Figure 6 and highlighting two pixels from Figure 6 in Figure 7. The low standard deviation indicates repeat measurements frequently returning the same DN. The top of the image is farther away than the bottom of the image, as seen in Figure 8, but Figure 6 shows no degradation with distance to the target (the additional noise at the top of the image is due to the gradient of the surface, not the distance to the surface). Therefore, while accuracy is a function of location in the FOV, precision is uniform. However, since DN maps to distance nonlinearly, there is less resolution at greater distances. In addition, due to the Peltier element heating and cooling the IR projector to stabilize the wavelength, errors are larger on the left and right sides for the first ~ 30 s of data

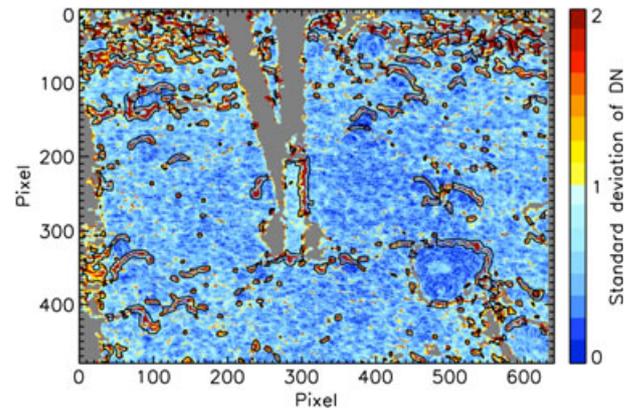


Figure 6. Standard deviation at each pixel from 100 frames of sensor digital numbers (DNs) from the Kinect. Scene is the same as Figure 8. The top of the image is farther away than the bottom and the data show there is no relationship between sensor precision and distance, although there is some additional noise due to the larger surface gradients observed at greater distances. Two pixels from this image are examined in detail in Figure 7. This figure is available in colour online at wileyonlinelibrary.com/journal/espl

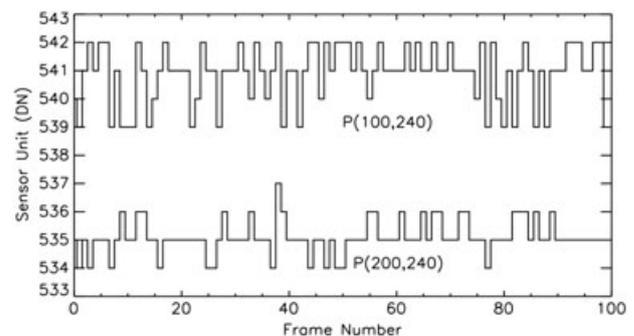


Figure 7. Noise on two sample pixels over time. The upper line, representing all the data used in pixel (100, 240) of Figure 6 has a mean DN value of 540.88, a mode of 541, and a standard deviation of 1.01. The lower line, corresponding to pixel (200, 240) of that same image, has mean 535.09, mode 535, standard deviation 0.59.

collection. Since the heater may turn on and off over time (on a timescale of minutes), the entire depth image may shift by several DN during these thermal changes.

Consistent repeat measurements with the device lead us to consider the precision of the Kinect as better than or equal to 1 DN, or equal to the distance resolution at a given distance. The 'Step Size' line in Figure 3 is simply the step size between one DN and the next when converted to mm, and the dashed lines highlight the precision (1–75 mm) at varying distances (0.5–5 m).

Issues and Limitations

The Kinect, designed to detect human motion in a standard home room, has limitations when used in other settings. There are three categories of limitations: device hardware, environmental properties, and surface properties.

Applications of the Kinect may be limited by the measurement range (0.5–5 m), FOV (up to ~ 5 m², but often smaller to improve distance resolution), accuracy (± 1 DN; 1–75 mm) or precision (± 1 DN; 1–75 mm). Thermal changes (internal or external) over time may reduce the accuracy and precision. Spatial resolution of ~ 3 pixels may cause small objects visible in the RGB frame to not be detected by the depth sensor. The primary

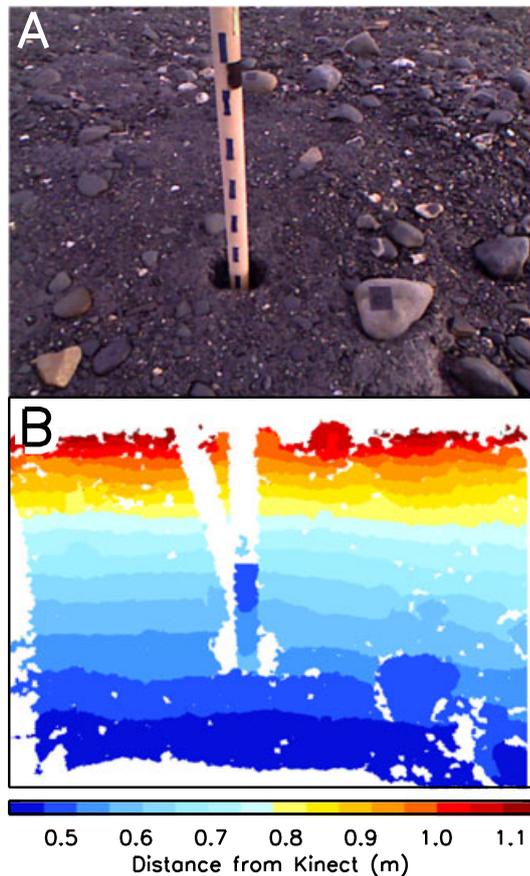


Figure 8. Scene recorded by Kinect showing debris-covered ice surface, rocks, and ablation stake. (A) is the RGB image; (B) is a single distance image. White represents no data. The left white stripe connected to the stake is the shadow cast by the IR projector and detected due to the offset between the projector and camera. The upper portion of the stake appears white where it is too close (~ 0.45 m) to the detector. This figure is available in colour online at wileyonlinelibrary.com/journal/esp1

environmental limitation is that the Kinect does not work in bright sunlight. It emits and detects IR light and cannot distinguish the returning signal from the background when an interfering IR source such as the sun saturates the sensor. Certain surfaces absorb IR and cannot be accurately detected. In our tests on glacier ice, clean smooth ice returns usable data. Highly crystallized ice and snow sometimes scatter the signal, giving few data points per image.

Large gradients or step changes in distance (for example, the upper portion of Figure 8, or the portions of that image that capture both the near stake and the far background) introduce large errors. The pattern recognition may fail locally when such gradients occur, manifesting as invalid data, the near object filling in and becoming artificially enlarged, or a small object being skipped entirely.

In a bathymetric study there are three issues: absorption, surface wave interference, and refraction. Absorption is a function of turbidity and other factors blocking IR transmission, and can be a limiting factor for subsurface studies, but we have observed underwater features at a depth of 1 m through calm, slightly turbid water before signal degradation occurred. Small amounts of turbidity limit the penetration depth of the signal, but eventually enough turbidity supports clear views of the water surface (Combès *et al.*, 2011). When waves exist on a transparent liquid, the bathymetric measurements are convoluted (Fryer and Kniest, 1985). The refraction requires a new calibration, discussed below ('Bathymetry and submerged surface measurements').

Proposed solutions

We address mitigation techniques and/or solutions to most of the limitations described above. If the target is larger than the FOV of the Kinect, we suggest two possible approaches: (1) capture multiple scenes with some overlap between the scenes, and then use industry standard LiDAR scene-stitching techniques to combine the multiple scans; or (2) Use the 'KinectFusion' algorithm (Izadi *et al.*, 2011; Newcombe *et al.*, 2011), implemented in the 'KinFu' program (part of the Point Cloud Library; Rusu and Cousins, 2011), which allows one to move the Kinect and scan an area or object, automatically stitching together each frame into one large 3D model. For very large areas, the KinFu implementation has been extended, named Kintinuous, and used to map paths more than 100 m long (Whelan *et al.*, 2012). Thermal changes introduced by the internal heater turning on and off over time, or due to external temperature changes, can be tracked by having part of the scene be a fixed object.

If sunlight may be a problem, solutions are to collect data in the shade, mornings and evenings when the sun is low, when there is thick cloud cover, or at night. If the study area is small enough, a tarp or tent with the base removed can temporarily create an area without the external IR interference. If the surface does not reflect IR well and gives sparse returns, multiple images acquired at 30 Hz may allow data gaps to be filled in. If the surface gives no returns, it may be possible to cover the surface with a material which properly reflects IR with minimal surface change errors (e.g. flour or a light fabric). Finally, when observing through a liquid with surface waves, errors will increase but more complex post-processing algorithms can be developed to retrieve bathymetry (e.g. Fryer and Kniest, 1985) and wave properties, or waves can be manually damped (e.g. with floating clear trays; Elfick and Fryer, 1984). Anything that changes the refractive index will skew the data, but this can be advantageous, for example, allowing multiple Kinect to build a 3D image of a plume of hot water entering cold water or a gas jet in air (Berger *et al.*, 2011).

Applications in Earth Science

The Kinect can be used in many environments by Earth scientists as a 3D camera with data output equivalent in format and quantity to commercial scanners that cost tens to hundreds times more (Table II). It also offers some advantages over photogrammetry, the technique traditionally used to generate small-scale digital elevation models. The Kinect advantages include real-time feedback, simple post-processing, simple data acquisition without the need for an overhead camera system, and the ability to detect surfaces that do not provide enough information for the stereo algorithms to derive distance. The Kinect also collects data at 30 Hz, which is possible with stereo video cameras, but significantly more computationally expensive to process.

While the Kinect FOV is only $\sim 5\text{ m}^2$, it still has a global context. For example, integrating data from multiple Kinects, one can cover the footprint of a high-resolution satellite sensor, useful for data verification, or a single Kinect can check the consistency of a scale model of a field site (Welch and Dikkers, 1978; Brasington and Smart, 2003).

The objective of the following examples is to demonstrate the efficacy of the Kinect outside and in adverse environmental conditions, despite its design for indoor purposes. The first example, glaciology, describes an experiment using multiple data collections over time to determine surface changes in 3D. The second example, bathymetry, illustrates streamed elevation data obtained through water and describes the required calibration. The third example, geomorphology, shows small

Table II. Comparison between the Kinect, MESA Imaging SR4000 time-of-flight camera, and Konica Minolta VIVID 9i non-contact 3D digitizer. We use data sheets for the SR4000 and VIVID 9i from the manufacturer web site, and personal communication with Konica representatives. The table is organized into three sections: sensor characteristics, device internal hardware, and external properties. Kinect precision ± 1 DN, coupled with nonlinear step size, causes the range for precision and accuracy

	Kinect	SR4000	VIVID 9i	units
Accuracy	1–75 (1 DN)	10	0.05	mm
Precision	1–75 (1 DN)	4	0.008	mm
Range	0.5–5	0.8–8	0.5–2.5	m; near–far
Field of view	58 × 40	69 × 56	70 × 66	degrees; horiz. × vert.
Output size	640 × 480	176 × 144	640 × 480	pixels
Data rate	9 216 000	760 320	122 800	points s ⁻¹
Output	DN image	<i>x, y, z</i>	STL, DXF,	
	RGB image	point distance,	OBJ, VRML,	
	or <i>x, y, z, r, g, b</i>	amplitude,	ASCII points	
		confidence		
RGB metadata	Yes	No	Yes	
Scene repeat rate	30	30	0.4	Hz
Wavelength	830	850	690	nm
Cost	100	4000	50 000	USD (approx.)
Dimensions	140 × 50 × 35	65 × 65 × 68	221 × 412 × 282	mm; l × w × h
Environment	0–40	10–50	10–40	C
Data logger needed	Yes	Yes	Yes	
Power	12; 1	12; 1	100; 0.6	V; A
Weight	500	470	15 000	g

sand waves and includes a description of the average wave feature dimensions. These example experiments illustrate a small subset of the diverse environments and material properties that are suitable for imaging with the Kinect.

Glaciology and surface change

DEMs and LiDAR-like data are useful for glaciological studies. Importantly, many studies that do not currently use 3D digital data might benefit from cost-effective access to this technology. In a recent study tracking lateral motion of supraglacial features, Irvine-Fynn *et al.* (2011) noted that a key disadvantage of the traditional photographic method was lack of depth perception. Similarly, small-scale basal roughness described by Hubbard and Hubbard (1998) was measured by total station survey, a laborious technique. Finally, Roberson (2008), Evans and Hiemstra (2005), and Glasser *et al.* (1999) have all published studies on the shape and size of mm- to cm-scale objects and structures near the termini of glaciers. Easy acquisition of digital data in three dimensions, made possible by a Kinect, might better capture these glacial processes and signatures.

We present a 4D (surface, time) ablation study. Ablation is usually measured in 2D with a stake (point, time). When debris cover is present the stake method has large errors due to the spatially heterogeneous response of the ice to thermal changes (Nakawo and Young, 1981; Nakawo and Rana, 1999; Zhang *et al.*, 2011). Thin debris transfers additional heat to the ice and accelerates melting, while thicker debris acts as an insulator and reduces melting. The critical thickness for the two modes of behavior varies with altitude, latitude, and field site, but is usually between 10 mm and 80 mm (Reznichenko *et al.*, 2010).

Experimental setup

In August 2011 we deployed a Kinect for two weeks on the surface of Matanuska Glacier, Alaska. We used a steam drill to mount a structure in the glacier, and then attached the Kinect to the structure, pointing down at a $\sim 45^\circ$ angle toward the ice. The ice was covered by a thin layer of debris less than 5 mm thick, with some pebbles and stones up to 50 mm diameter. Included in the scene was an ablation stake with 2.54 cm (1-inch)

segments marked allowing calibration between the distance-derived ablation estimates and the visual ablation as tracked by the RGB camera and the stake (Figure 8).

We protected the Kinect when not in use with a plastic bag. Each evening near 2200 h local time, when the sun was still up but near the horizon, we recorded 1 minute of data. The resultant dataset is ~ 1800 RGB and distance images per day for 13 consecutive days.

Experiment results

Data taken 24 h apart on 9 and 10 August 2011 were aligned using the stationary stake as a reference point and the early data were subtracted from the later (Figure 9). The ablation stake measurements (not shown) indicate surface lowering of ~ 0.04 m, similar to the median rate measured by the Kinect. However, the 3D data in Figure 9 show a non-uniform ablation rate. Comparing Figure 9 with Figure 8 (or rotating the data in 3D interactively on a computer; EDF R&D, 2011), it is evident that the areas corresponding to the larger rocks did not lower as much as areas corresponding to the thin-debris-covered surface. This experiment shows that the problems raised by heterogeneous debris cover in ablation zones (Nakawo and Young, 1981; Nakawo and Rana, 1999) can be solved (at small scales) in a cost-effective manner.

Bathymetry and submerged surface measurements

Many disciplines, including hydrology, geomorphology, and ecology, benefit from high-resolution mapping of river bank or bed features (Lane *et al.*, 1994; Butler *et al.*, 1998, 2002; Chandler *et al.*, 2002, 2003; Marcus and Fonstad, 2008). Review articles published in the past decade on remote sensing of rivers demonstrate the rapid advancements being made, as well as widespread interest in the field (e.g. Mertes, 2002; Gilvear and Bryant, 2003; Heritage and Hetherington, 2007; Marcus and Fonstad, 2008; Johansen *et al.*, 2010; Marcus and Fonstad, 2010). Traditional methods for measuring subaqueous river structures or monitoring riverbed changes include photogrammetry (Fryer, 1983; Elfick and Fryer, 1984; Fryer and Kniest, 1985; Lane, 2000; Butler *et al.*, 2002), near-IR LiDAR

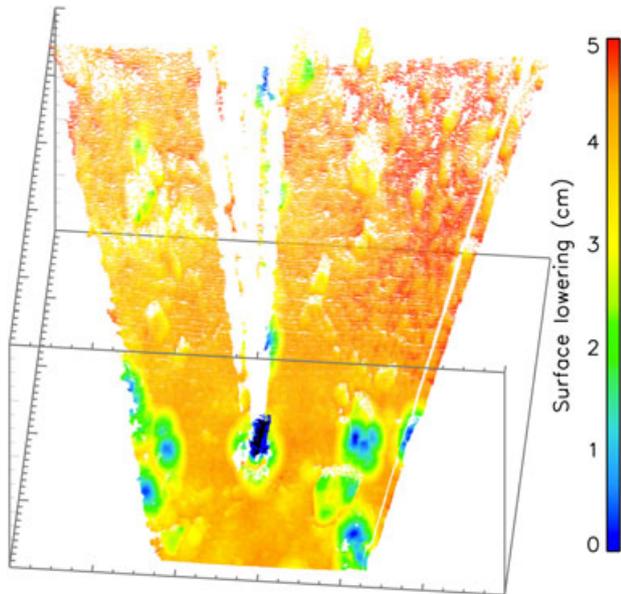


Figure 9. Surface height change from 9 August to 10 August 2011 on Matanuska Glacier, AK. Surface ablation measured by stake was approximately 0.04 m. Median scene ablation measured by differencing the distance images match this value. Areas thermally insulated by rocks (see Figure 8) indicate less ground surface lowering than the median. Wide shadow from stake is due to the projector and camera position offsets. This figure is available in colour online at wileyonlinelibrary.com/journal/espl

(airborne or ground-based) (Heritage and Hetherington, 2007; Allouis *et al.*, 2010), green LiDAR (also called bathymetric LiDAR) (Allouis *et al.*, 2010), or one-dimensional, cross-channel stream profile measurements collected with traditional surveying methods (e.g. stakes; Furbish, 1987).

Field trial

The Kinect provides a low-cost, logistically simple method for making mm- to cm-resolution 3D models of river bathymetry in appropriate environments such as shallow streams, braided rivers, and shallow lagoon or beach areas. The quality of bathymetric data will depend on several factors, including incidence angle, total dissolved solutes, turbidity, and substrate material, which are not addressed further in this discussion.

We image a section of streambed from Scott Creek, CA, with an average water depth in the area of ~7 cm (Figure 10). Environmental factors present during data collection include direct sunlight and surface ripples (both visible in Figure 10A) – two potential limitations to data collection – but we note that neither are causing significant interference in this case. Importantly, the RGB image shows a section of the water surface (near the tripod leg) that is highly reflective, and would likely not produce usable data if this image were used for photogrammetry, but the Kinect obtains bathymetric data through this visually reflective area. We do not know the cause of signal loss on the right side of the distance image.

Calibration

The DN-to-distance conversion algorithms written for the Kinect are calibrated for air, though several Earth science applications would benefit from bathymetric data. However, bathymetric observations present the two-medium problem: the IR signal is refracted at the water interface, requiring calibration of x and y values and a new DN-to- z function. As with the air calibration, each Kinect will require its own through-water calibration, in this case not factory-supplied, and therefore not accessible by existing calibration code.

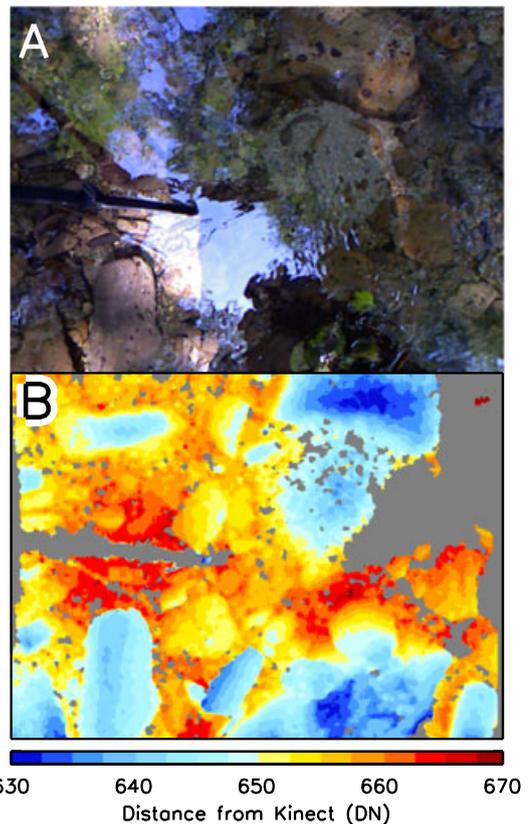


Figure 10. Bathymetric data collected with the Kinect. (A) RGB image looking down through shallow water to a submerged stream bed. (B) Average of 100 distance images of the same scene from the point of view of the distance camera shown in raw data coordinates (pixel, pixel, DN). Gray represents no data. This figure is available in colour online at wileyonlinelibrary.com/journal/espl

To calibrate the Kinect for bathymetric measurements, we suggest following the general method outlined for through-water photogrammetry presented by Butler *et al.* (2002). The Kinect should be mounted over a dry tank and the bottom surface should contain a variety of control points, perhaps painted a unique color so their locations can easily be determined by querying for that color in the RGB image and then using the (x,y) locations returned in the depth image. For the baseline, record the 2D distance array from the Kinect to the base of the dry tank, and the original x and y coordinates of the control points. Add a small amount of water to the tank and collect a sequence of raw data, then repeat for increasing water depths. For each sample, record (1) the actual depth of the water (a single value), (2) the Kinect measured distance to the base of the wet tank through the water (a 2D array), and (3) the Kinect measured x and y coordinates of the control points.

These parameters can be used with the baseline data to calibrate the Kinect for bathymetric measurements. The calibration will be applied to the apparent water depth; therefore it must first be identified by subtracting the distance through air to the water surface from each distance measurement in the 2D array. In the field, we suggest mounting the Kinect looking straight down, and a float in the scene can define the plane of the water (or, for example, the point where the tripod leg enters the water as seen in Figure 10A).

We present an example dataset highlighting the effects of water on subaqueous distance measurements (Figure 11). We began with an empty tank with the base 0.81 m below the Kinect (and a 40 mm high block in part of the scene), and added water in increments to a final depth of 0.20 m, waited about 1 min for the surface to settle, then collected one image

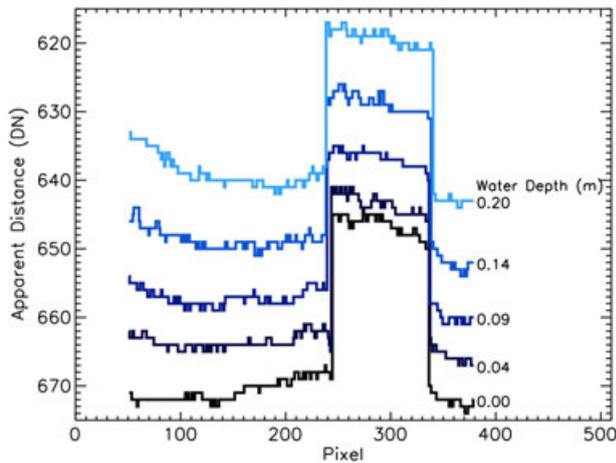


Figure 11. Transects across distance image looking down into a tank showing effect of changing water depth on observation of subaqueous surface. The unfilled tank is imaged with water depth 0 (bottom black line), and blue lines above show that same surface as imaged through water (water depth shown at right edge of lines.). The transects intersect a block surface 40 mm high. This image highlights the changing apparent depth with water depth, the changing apparent depth with pixel position (left side of plots are diverging), and the changing pixel positions due to water refraction (edge of block appears to move). This figure is available in colour online at wileyonlinelibrary.com/journal/espl

per water increment. We collected only one image because the ± 1 DN error of the Kinect was within the uncertainty of the water height measurement, and also to better familiarize the reader with more raw data, as all discussions so far have presented data with high SNR. We extracted distance values from one row of each image, crossing the block (Figure 11). The data illustrate multiple issues raised by two-medium measurements (air, water) that should be addressed by calibration: (1) The DN value to the object changes with water depth, requiring z calibration; (2) as water is added, objects appear closer and therefore occupy a larger FOV (more pixels), shown by the block appearing a few pixels wider in deeper water. This requires x and y calibration; (3) the change of x , y , and z will vary radially from the center, shown by the left edge of the lines drifting apart as water is added. This requires the x , y , and z calibration to be a function of x and y position in the FOV.

We have demonstrated in Figure 10 that the Kinect can obtain stream bathymetry data. Proper implementation of the described calibration will mitigate the issues raised by the two-medium problem, allowing the Kinect to be used for bathymetric studies.

Geomorphology and solid surface waves

Similar to bathymetric measurements, measuring and quantifying the surface properties of temporarily exposed subaqueous environments is of interest to the Earth science community (Chandler *et al.*, 2003; Smart *et al.*, 2004; Hodge *et al.*, 2009a, 2009b). In this case study we demonstrate the utility of the Kinect in the geomorphology domain and use the sensor to define the characteristic length scale of sand ripples on a beach. We again used only a single frame of data in this experiment to show that the data are usable even without collecting multiple frames and averaging to increase the SNR.

We placed the Kinect on a tripod looking roughly straight down onto a flat beach at low tide where the receding water left small ripples in the sand. We converted one frame of data to real-world x, y, z coordinates, converted the point-cloud to a DEM using points2grid (Crosby *et al.*, in review), and then rendered the DEM with a hill-shade algorithm (GDAL Development Team, 2011) (Figure 12). We then performed a 2D discrete

Fourier transform (DFT) on a smoothed subset of the DEM to find the characteristic wavelength of the ripples (Figure 13), and found the wave height to be ~ 20 mm ± 5 mm (± 1 DN) and wavelength to be ~ 40 mm in the NNE/SSW direction relative to the point of view of the Kinect. Errors can be as small as ± 1 mm if the Kinect is placed closer to the target.

Conclusion

We introduce the Kinect to the Earth science community, describe the hardware and software (providing code for data processing), and quantify the data quality. We present three examples in glaciology, bathymetry, and geomorphology, showing that the Kinect is a viable device for scientific experiments needing 3D data on scales of mm to cm over areas of a few square meters. There are a wide range of disciplines in the Earth science domain that can use this sensor for both old and new experiments in the lab and in the field. As Earth scientists explore the capabilities of this new device we expect applications and data collection at new scales and in previously unattainable environments.

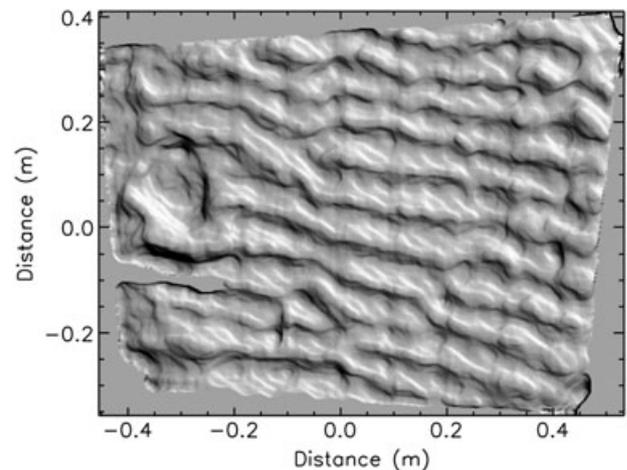


Figure 12. Plan view (x, y) looking down on a ~ 1 m² patch of beach containing small sand ripples. Figure 13 shows a discrete Fourier transform of a subset of the scene.

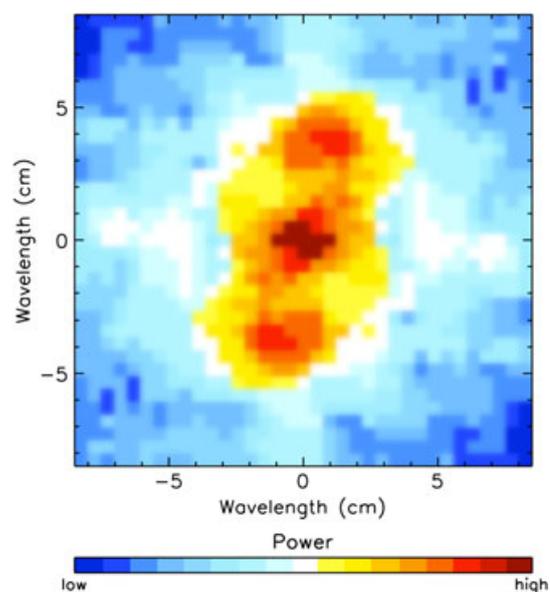


Figure 13. 2D discrete Fourier transform of a subset of Figure 12 indicating that the primary wavelength of sand ripples is ~ 40 mm in the NNE/SSW direction (relative to the point of view of the Kinect). This figure is available in colour online at wileyonlinelibrary.com/journal/espl

Acknowledgements—We thank the open source community and developers of the libfreenect software library, particularly OpenKinect at <http://openkinect.org>. We are also grateful to the three anonymous reviewers and the editors for their constructive comments and encouragement. We thank B. K. Norris and D. M. Winslow for help collecting data, C. A. Edwards for use of the fish tank and laboratory, and J. I. Walter for an examination of the accelerometer. This work was funded in part by NASA Headquarters under the NASA Earth and Space Science Fellowship Program (Grant NNX10AN83H), and the National Science Foundation Graduate Research Fellowship Program (Fellow ID 2009083666).

Appendix A

Data Collection Workflow

At this point the reader should understand how the Kinect works, possible use cases, and limitations of the device. However, collecting data may still be problematic as the device was not initially designed to be a scientific instrument, and the open source projects used to collect data have limited documentation. We therefore describe in detail the steps necessary to collect and process data with the Kinect. The following example should work on a computer running Linux or OS X.

- (1) Install necessary software. The libfreenect code can be downloaded and built from source, but it is easier to use a package manager system. On OS X, the three popular package managers (brew, macports, and fink) each include libfreenect and all necessary dependencies. After installing the package manager, install libfreenect by typing, for example, 'brew install libfreenect'. All major Linux distributions provide libfreenect via their respective package management systems.
- (2) Install the 'kinect_register' program included with the digital supplement. The latest version of this code is available at <https://github.com/mankoff/libfreenect/>.
- (3) Record the calibration parameters for later use by running the following command with the Kinect plugged in: 'kinect_register -s regdump', where '-s' stands for 'save' and 'regdump' can be any filename you choose.
- (4) Acquire data by running the 'record' program in a terminal at the command line, with one argument: the folder where to store the data. For example, 'record data' will fill the 'data' folder with the RGB, distance, and accelerometer data. Stop the record program by typing 'CTRL + C'.

From this point on the Kinect does not need to be present.

- (5) Increase the SNR. We use the following pseudo-code to average together multiple distance images and increase the SNR:


```
data = [ 640, 480, n ] % matrix: 640 x 480 x number of files
output = [ 640, 480 ]
for i = 1, n
    data[ *, *, i ] = loadPGM( ) % load a file
end
for i = 1, 640
    for j = 1, 480
        % extract one pixel from all frames
        vec = data[ i, j, * ]
        % find indices (subscripts) with good data
        good = where( vec > 0 AND vec < 2047 )
        % take the mean of the good data
        output[ i, j ] = mean( vec[ good ] )
    end
end
write(output) %output identical to 'record' format
```

- (6) Calibrate the data with the 'kinect_register' program:


```
kinect_register -a regdump file.PGM
kinect_register -a regdump file.PGM rgb.PPM
```

 Where '-a' stands for 'apply', 'regdump' is the file saved previously, and 'file.PGM' is the output from the previous step. Optionally, an RGB image can be included as shown in the second invocation. The output from this step is 4 (or 5 if RGB PPM included) new files. file.x is a 640 × 480 array of floating point numbers, each representing the x-coordinate of the data, file.y is the same for the y-coordinate, and file.z is a 640 × 480 array of integers representing distance from the Kinect. These same three matrices are available in an ASCII point-cloud form in file.ply, which contains x, y, z triplets. All files are in mm units. If the PPM file is included, then rgb.reg.PPM is the contents of rgb.PPM but registered to (aligned with) the depth data, and the file.ply contains x y z r g b sextuplets rather than x y z triplets. The ASCII PLY file is a common format for 3D point-cloud data that can be read by many different software packages.
- (7) Analyze. From this point on, data collection is complete, and the analysis will be dependent on the experiment and is beyond the scope of this paper.

Appendix B

Accessing Sensor Data

Understanding a new data format can be time consuming. To assist with reading the data we provide code (C, Python, IDL, and MATLAB) to read the DN data written by 'record' as big-endian PGM files. Since the 'kinect_register' program reads this same file format, the functions below are useful templates to write big-endian PGM files. Once a reader and writer exist, DN data can be loaded, manipulated, rewritten, and then calibrated.

```
// C
fp = fopen("file.pgm", "r");
while (getc(fp) != 'n' // skip header line
uint16_t data [640*480];
fread (data, sizeof (uint16_t), 640*480, fp) ; // read the data
fclose(fp) ;

# Python
import numpy as np
infile = open('file.pgm','r')
header = next(infile)
infile.seek(len(header))
data = np.fromfile(infile, dtype=np.uint16).reshape((480, 640))

; ; IDL
openr, lun, "file.pgm", /get_lun
header = {P5:BYTARR(2), width:BYTARR(4), height:BYTARR(4), $
maxV:BYTARR(7)}
readu, lun, header
data = intarr( string(header.width), string(header.height) )
readu, lun, data
free_lun, lun

% MATLAB
data = imread('file.pgm');
data = swapbytes(data);
```

References

- Allouis T, Bailly J, Pastol Y, Le Roux C. 2010. Comparison of LiDAR waveform processing methods for very shallow water bathymetry using Raman, near-infrared and green signals. *Earth Surface Processes and Landforms* **35**: 640–650.

- Baily B, Collier P, Farres P, Inkpen R, Pearson A. 2003. Comparative assessment of analytical and digital photogrammetric methods in the construction of DEMs of geomorphological forms. *Earth Surface Processes and Landforms* **28**: 307–3203.
- Berger K, Ruhl K, Magnor M. 2011. The capturing of turbulent gas flows using multiple Kinects. In Proceedings of the 1st IEEE Workshop on Consumer Depth Cameras for Computer Vision, Barcelona, Spain, 12 November 2011. IEEE. Available: <https://www.youtube.com/watch?v=MkuCYgd3g5s> [2 October 2012].
- Brasington J, Smart RMA. 2003. Close range digital photogrammetric analysis of experimental drainage basin evolution. *Earth Surface Processes and Landforms* **28**: 231–247.
- Burrus N. 2010. Kinect calibration. Available: <http://nicolas.burrus.name/index.php/Research/KinectCalibration> [10 October 2011].
- Butler JB, Lane SN, Chandler JH. 1998. Assessment of DEM quality for characterizing surface roughness using close range digital photogrammetry. *The Photogrammetric Record* **16**: 271–291.
- Butler JB, Lane SN, Chandler JH, Porfiri E. 2002. Through-water close range digital photogrammetry in flume and field environments. *The Photogrammetric Record* **17**: 419–439.
- Candela T, Renard F, Schmittbuhl J, Bouchon M, Brodsky EE. 2011. Fault slip distribution and fault roughness. *Geophysical Journal International* **187**: 959–968.
- Chandler JH, Ashmore P, Paola C, Gooch M, Varkaris F. 2002. Monitoring river-channel change using terrestrial oblique digital imagery and automated digital photogrammetry. *Annals of the Association of American Geographers* **92**: 631–644.
- Chandler JH, Buffin-Bélanger T, Rice S, Reid I, Graham DJ. 2003. The accuracy of a river bed moulding/casting system and the effectiveness of a low-cost digital camera for recording river bed fabric. *The Photogrammetric Record* **18**: 209–223.
- Chassereau JA, Bell JM, Torres R. 2011. A comparison of GPS and lidar salt marsh DEMs. *Earth Surface Processes and Landforms* **36**: 1770–1775.
- Combès B, Guibert A, Memin E, Heitz D. 2011. Free-surface flows from Kinect: Feasibility and limits. In Proceedings of the Forum on Recent Developments in Volume Reconstruction Techniques Applied to 3D Fluid and Solid Mechanics (FVR 2011), Chasseneuil, France.
- Crosby CJ, Krishnan S, Arrowsmith JR, Kim HS, Colunga J, Alex N, Baru B. in review. Points2Grid: An Efficient Local Gridding Method for DEM Generation from Lidar Point Cloud Data: Geosphere special issue on Applications of Lidar in the Earth Sciences. Available: <http://www.opentopography.org/index.php/Tools/otforge/points2grid>.
- EDF R&D. 2011. CloudCompare (GPL software), Telecom ParisTech Version 2.3. Available: <http://www.danielgm.net/cc/> [2 October 2012].
- Elfick MH, Fryer JG. 1984. Mapping in shallow water. *International Archives of Photogrammetry and Remote Sensing* **25**: 240–247.
- Evans DJA, Hiemstra JF. 2005. Till deposition by glacier submarginal, incremental thickening. *Earth Surface Processes and Landforms* **30**: 1633–1662.
- Freedman B, Shpunt A, Machline M, Arieli Y. 2010. Depth mapping using projected patterns, 13 May 2010. US Patent App. US 2010/0118123 A1.
- Fryer JG. 1983. A simple system for photogrammetric mapping in shallow water. *The Photogrammetric Record* **11**: 203–208.
- Fryer JG, Kniest HT. 1985. Errors in depth determination caused by waves in through-water photogrammetry. *The Photogrammetric Record* **11**: 745–753.
- Furbish DJ. 1987. Conditions for geometric similarity of coarse stream-bed roughness. *Mathematical Geology* **19**: 291–307.
- GDAL Development Team. 2011. GDAL – Geospatial Data Abstraction Library, Version 1.8.1. Open Source Geospatial Foundation.
- Gilvear D, Bryant R. 2003. Analysis of aerial photography and other remotely sensed data. In *Tools in Fluvial Geomorphology*, Kondolf GM, Piégay H (eds). Wiley: Chichester; 135–170.
- Glasser NF, Bennett MR, Huddart D. 1999. Distribution of glaciofluvial sediment within and on the surface of a high arctic valley glacier: Marthabreen, Svalbard. *Earth Surface Processes and Landforms* **24**: 303–318.
- Heritage GL, Hetherington D. 2007. Towards a protocol for laser scanning in fluvial geomorphology. *Earth Surface Processes and Landforms* **32**: 66–74.
- Herrera DC, Kannala J, Heikkilä J. 2011. Accurate and practical calibration of a depth and color camera pair. In Proceedings of the 14th International Conference on Computer Analysis of Images and Pattern, Seville, Spain, 29–31 August 2011; 437–445.
- Hodge R, Brasington J, Richards K. 2009a. Analysing laser-scanned digital terrain models of gravel bed surfaces: linking morphology to sediment transport processes and hydraulics. *Sedimentology* **56**: 2024–2043.
- Hodge R, Brasington J, Richards K. 2009b. In situ characterization of grain-scale fluvial morphology using terrestrial laser scanning. *Earth Surface Processes and Landforms* **34**: 954–968.
- Hubbard B, Hubbard A. 1998. Bedrock surface roughness and the distribution of subglacially precipitated carbonate deposits: implications for formation at Glacier de Tsanfleuron, Switzerland. *Earth Surface Processes and Landforms* **23**: 261–270.
- Hut R, Weijs S, Luxemburg WMJ. 2010. Using the Wiimote as a sensor in water research. *Water Resources Research* **46**:W12601.
- Irvine-Fynn TDL, Bridge JW, Hodson AJ. 2011. In situ quantification of supraglacial cryoconite morphodynamics using time-lapse imaging: an example from Svalbard. *Journal of Glaciology* **57**: 651–657.
- Izadi S, Kim D, Hilliges O, Molyneaux D, Newcombe R, Kohli P, Shotton J, Hodges S, Freeman D, Davison A, Fitzgibbon A. 2011. KinectFusion: real-time 3D reconstruction and interaction using a moving depth camera. In Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology, ACM; 559–568.
- Johansen K, Phinn S, Witte C. 2010. Mapping of riparian zone attributes using discrete return LiDAR, QuickBird and SPOT-5 imagery: Assessing accuracy and costs. *Remote Sensing of Environment* **114**: 2679–2691.
- Konolige K, Mihelich P. 2011. Technical description of Kinect calibration. Available: http://www.ros.org/wiki/kinect_calibration/technical [9 October 2012].
- Lane SN. 2000. The measurement of river channel morphology using digital photogrammetry. *The Photogrammetric Record* **16**: 937–961.
- Lane SN, Richards KS, Chandler JH. 1994. Developments in monitoring and modelling small-scale river bed topography. *Earth Surface Processes and Landforms* **19**: 349–368.
- Lane SN, James TD, Crowell MD. 2000. Application of digital photogrammetry to complex topography for geomorphological research. *The Photogrammetric Record* **16**: 793–821.
- Marcus W, Fonstad M. 2008. Optical remote mapping of rivers at sub-meter resolutions and watershed extents. *Earth Surface Processes and Landforms* **33**: 4–24.
- Marcus W, Fonstad M. 2010. Remote sensing of rivers: the emergence of a subdiscipline in the river sciences. *Earth Surface Processes and Landforms* **35**: 1867–1872.
- Mertes LAK. 2002. Remote sensing of riverine landscapes. *Freshwater Biology* **47**: 799–816.
- Micron. 2011. 1/2-Inch Megapixel CMOS Digital Image Sensor. Technical report. Available: http://download.micron.com/pdf/datasheets/imaging/mt9m001_1300_mono.pdf [20 October 2011].
- Milledge DG, Lane SN, Warburton J. 2009. The potential of digital filtering of generic topographic data for geomorphological research. *Earth Surface Processes and Landforms* **34**: 63–74.
- Nakawo M, Rana B. 1999. Estimate of ablation rate of glacier ice under a supraglacial debris layer. *Geografiska Annaler: Series A, Physical Geography* **81**: 695–701.
- Nakawo M, Young G. 1981. Field experiments to determine the effect of a debris layer on ablation of glacier ice. *Annals of Glaciology* **2**: 85–91.
- Newcombe RA, Izadi S, Hilliges O, Molyneaux D, Kim D, Davison AJ, Kohli P, Shotton J, Hodges S, Fitzgibbon A. 2011. KinectFusion: real-time dense surface mapping and tracking. In Proceedings of the 10th IEEE symposium on mixed and augmented reality, IEEE.
- OpenKinect. 2011. OpenKinect. <http://openkinect.org> [2 October 2012].
- PrimeSense Ltd. 2011. The PrimeSense[®] Reference Design 1.08. Available: http://www.primesense.com/files/FMF_2.PDF [accessed on July 10 2011].
- Reznichenko N, Davies T, Shulmeister J, McSaveney M. 2010. Effects of debris on ice-surface melting rates: an experimental study. *Journal of Glaciology* **56**: 384–394.
- Roberson S. 2008. Structural composition and sediment transfer in a composite cirque glacier: Glacier de St. Sorlin, France. *Earth Surface Processes and Landforms* **33**: 1931–1947.
- Rusu RB, Cousins S. 2011. 3D is here: Point Cloud Library (PCL). In IEEE International Conference on Robotics and Automation (ICRA), Shanghai, China, 9–13 May 2011.

- Schaefer M, Inkpen R. 2010. Towards a protocol for laser scanning of rock surfaces. *Earth Surface Processes and Landforms* **35**: 147–423.
- Smart G, Aberle J, Duncan M, Walsh J. 2004. Measurement and analysis of alluvial bed roughness [Mesure et analyse de la rugosité de lit d'alluvion]. *Journal of Hydraulic Research* **42**: 227–237.
- Smisek J, Jancosek M, Pajdla T. 2011. 3D with Kinect. In Proceedings of the 1st IEEE Workshop on Consumer Depth Cameras for Computer Vision, Barcelona, Spain, 12 November 2011.
- Tribelhorn B, Dodds Z. 2007. Evaluating the Roomba: a low-cost, ubiquitous platform for robotics research and education. In 2007 IEEE International Conference on Robotics and Automation; 1393–1399.
- Welch R, Dikkers K. 1978. Educational and research aspects of non-metric, close range analogue photogrammetry. *The Photogrammetric Record* **9**: 537–547.
- Whelan T, McDonald JB, Kaess M, Fallon MF, Johannsson H, Leonard JJ. 2012. Kintinuous: spatially extended KinectFusion. In RSS Workshop on RGB-D: Advanced Reasoning with Depth Cameras, Sydney, Australia.
- Zhang Y, Fujita K, Liu S, Liu Q, Nuimura T. 2011. Distribution of debris thickness and its effect on ice melt at Hailuoguo glacier, southeastern Tibetan Plateau, using in situ surveys and ASTER imagery. *Journal of Glaciology* **57**: 1147–1157.