Subaerial gravel size measurement using topographic data derived from a UAV-SfM approach

Amy S. Woodget[1] and Robbie Austrums[2]

[1]Institute of Science and Environment, University of Worcester, Henwick Grove, Worcester, WR2 6AJ, UK, [2]Independent Geospatial Consultant, Richmond Road, Worcester, UK

Abstract

Accurate and reliable methods for quantifying grain size are important for river science, management and in various other sedimentological settings. Remote sensing offers methods of quantifying grain size, typically providing; (a) coarse outputs (c. 1m) at the catchment scale where individual grains are at subpixel level, or; (b) fine resolution outputs (c. 1mm) at the patch scale. Recently, approaches using unmanned aerial vehicles (UAVs) have started to fill the gap between these scales, providing hyperspatial resolution data (<10cm) over reaches a few hundred metres in length, where individual grains are at suprapixel level. This 'mesoscale' is critical to habitat assessments. Most existing UAV-based approaches use 2D textural variables to predict grain size. Validation of results is largely absent however, despite significant differences in platform stability and image quality obtained by manned aircraft versus UAVs. Here, we provide the first quantitative assessment of the accuracy and precision of grain size estimates produced from a 2D image texture approach. Furthermore, we present a new method which predicts subaerial gravel size using 3D topographic data derived from UAV imagery. Data is collected from a small gravel-bed river in Cumbria, UK. Results indicate that our new

topographic method gives more accurate measures of grain size (mean residual error -0.0001m). Better results for the image texture method may be precluded by our choice of texture measure, the scale of analysis or the effects of image blur resulting from an inadequate camera gimbal. We suggest that at our scale of assessment, grain size is more strongly related to 3D variation in elevation than to the 2D textural patterns expressed within the imagery. With on-going improvements, our novel method has potential as the first grain size quantification approach where a trade-off between coverage and resolution is not necessary or inherent.

Introduction

The mapping and quantification of fluvial grain (or substrate) size is important in the study of fluvial process, within both river science and management. Grain size data are a key input to hydraulic models, and are essential for quantifying sediment entrainment, transfer and deposition. Traditional approaches to grain size mapping typically use qualitative classification schemes such as the Wentworth Scale (Wentworth, 1922), or quantitative methods, such as in-situ or laboratory based physical measurement of individual grains, including areal, grid, transect or volumetric sampling (Wolman, 1954; Hey and Thorne, 1983; Church et al., 1987; Rice and Church, 1996). Data collection of this type is never spatially continuous, only sometimes spatially referenced, and rarely covers large spatial areas with great detail. Furthermore, traditional approaches can be labour-intensive, time consuming and often make assumptions about the representativeness of the spatially discontinuous samples over larger areas (Leopold, 1970; Verdú et al., 2005). The finer grain material is often under-sampled by a grid-by-number approach (Wolman, 1954; Church et al., 1987) and the removal of samples for volumetric analyses in the laboratory can destroy the local patches of habitat that they are aiming to investigate (e.g. freeze coring; Milan, 1996).

Since the 1970s, alternative methods of grain size quantification have made use of remote sensing technologies, fuelled by the need for less subjective approaches, which are non-invasive, reduce the time and effort spent in the field or laboratory and provide more continuous spatial coverage at a range of scales. Ongoing advances in digital photogrammetry, digital image analysis and surveying technologies mean that there is now an evolving body of remote sensing research for grain size quantification which makes use of imagery and/or elevation data. An overview is provided in Table 1. These studies evidence the trade-off between resolution (i.e. level of detail) and coverage (i.e. extent of survey) which often afflicts remote sensing methods. Table 1 also highlights that there exist a variety of different ways for obtaining grain size from imagery or digital elevation data. However, no single technique has yet proved its value for the rapid quantification of grain size at the mesoscale; that is, with centimetric spatial resolution over channel lengths from c. 50m to a few hundred metres. However, such outputs would be of great value for contributing to scientific understanding of fluvial mesohabitats and their applied management (Frissell et al., 1986; Newson and Newson, 2000).

In recent years, dramatic development in the technology and applicability of unmanned aerial vehicles (UAVs) has provided an alternative approach for quantifying fluvial grain size. UAVs are sometimes also known as ‘unmanned aerial systems’ (UAS), ‘remotely piloted aircraft systems’ (RPAS) or drones. Within this letter, we focus on the use of small (< 7kg) UAVs used in conjunction with novel ‘structure from motion’ digital photogrammetry (SfM) to derive fully orthorectified and georeferenced aerial imagery and topographic data. Readers are referred to Smith et al., (2015) and Eltner et al., (2016) for further detail on these developments.

To date, very few published studies have applied UAVs and SfM for quantifying grain size specifically. Those who have made progress in this area have adapted the image texture methods of Carbonneau et al., (2004), designed originally for use on imagery acquired from manned aircraft. Tamminga et al., (2015) acquired 5cm resolution imagery from a small, rotary-winged UAV over a 1km stretch of the Elbow River in Canada. Imagery was processed using digital photogrammetry software EnsoMOSAIC (MosaicMill Ltd, Finland) to create an orthophoto. Image texture, in the form of standard deviation of spectral values, was computed from this imagery, using a 1m^2 moving window. Grain size calibration data were acquired using close-range photo-sieving for 30 small sample plots (1m^2), where the B axes of 50 clasts were measured automatically using a Matlab routine. The resulting relationship between image texture and grain size gave a strong empirical correlation ($R^2 = 0.82$), which was subsequently used to estimate grain size over the entire area of interest. Whilst the UAV imagery itself was of hyperspatial resolution (5cm), the nature of their approach means that Tamminga et al., (2015) were only able to produce grain size predictions at a much coarser 1m spatial resolution. Furthermore, they present no associated quantitative error assessment of their predictions.

A similar approach was taken by de Haas et al., (2014) as part of a study exploring the evolution of alluvial fan surfaces. UAV imagery was collected at a resolution of 4-6cm and processed using SfM and the texture approach of Carbonneau et al.,

(2012) to produce grain size outputs at 0.7m resolution of an area covering 0.745km^2. Relative motion blur was found to affect UAV image quality, and was attributed to a combination of cloudy conditions (which reduced light levels and therefore necessitated increased exposure times) and wind gusts. Blurred parts of the resulting orthophoto artificially reduced image texture outputs and adversely affected the calibration with grain size. As a result, such areas were excluded from the calibration. Validation of the model using independent grain size data is not presented by de Haas et al., (2014) which again prohibits an understanding of the accuracy of this texture approach and limits any real comparison against existing techniques.

These papers highlight a need for robust and quantitative testing of grain size estimations produced using UAVs and SfM. In addition, the development and evaluation of alternative approaches which are less affected by spectral issues are of interest. For example, the development of topographic analysis methods for grain size estimation using terrestrial laser scanner data (e.g. Heritage and Milan, 2009; Brasington et al., 2012) may be applicable to UAV imagery, as topographic data in the form of dense point clouds are one of the outputs from SfM. Westoby et al., (2015) applied a UAV and SfM derived point cloud roughness approach to grain size quantification of an Antarctic moraine, but were unable to obtain a strong calibration relationship (R^2 = 0.225) between the standard deviation of elevation (i.e. roughness) and patch-scale D_{50} measures (i.e. grain size). They report a mean grain size estimation error of -2.90mm based on only five validation points, and do not report the precision of their results. Woodget et al., (2016) provide an initial pilot study in a fluvial setting, where topographic point cloud roughness data were successfully used for grain size prediction (R^2 = 0.7712, mean error = -0.01mm, precision = 16.4mm). We build on these results within this letter, using different and more comprehensive ground validation data. Our aim is to provide a quantitative assessment of the accuracy and precision of grain size predictions made using (a) an image texture approach and (b) a topographic (point cloud roughness) approach, based on imagery acquired using a small UAV and processed using SfM.

Site location

We selected a c.120m long reach of Coledale Beck, a gravel-bed river located near Braithwaite, Cumbria for this research. The chosen reach comprises a meandering pool-riffle system, with a bed composed predominantly of cobbles and boulders. The channel features a number of large unvegetated point bars and opposing steep, undercut banks. Variable subaerial grain sizes and a safe and accessible location for UAV flying made this a suitable site. Furthermore, the sediment dynamics of Coledale Beck are of interest due to their downstream impacts on Bassenthwaite Lake. The lake is designated as a National Nature Reserve and a Site of Special Scientific Interest, partly due to its rare vendace (*Coregonus vandesius*) fish population. The spawning grounds of this species are particularly sensitive to changes in the quantity and quality of sediment within the lake. Increasing siltation of the lake is thought to be partially responsible for the significant decline and subsequent extinction of the vendace population (Orr and Brown, 2004). As a result, methods capable of mapping and monitoring the evolution of sediment distribution within inflowing streams hold potential for habitat evaluation and informing management strategies.

Data acquisition and processing

Site set-up

Prior to data collection at Coledale, we established four permanent markers at the outer extents of the area of interest, using wooden stakes and circular survey markers. All subsequent data collected using a Leica Builder 500 total station (expected accuracy c. 1.5mm) were referenced to these markers using an arbitrary local co-ordinate system.

UAV survey

We flew a Draganflyer X6 rotary-winged UAV over the site at an altitude of c. 30m above ground level. Flight control was entirely manual due to the lack of an autopilot function. The UAV was mounted with a small, consumer grade digital camera (Panasonic Lumix DMC-LX3) held in a 1-axis brushless gimbal. The survey was conducted in July 2013 during dry, bright and calm weather conditions. We distributed 25 ground control points (GCPs) prior to the UAV survey, ensuring they were positioned to represent adequately the variation in topography across the site. The GCPs were constructed from thin, black PVC sheeting, marked in a cross pattern with white paint and, once positioned, were surveyed using the total station relative to the local co-ordinate system (using the permanent markers). The relatively short battery life on the UAV (c. 6 minutes) meant that three flights were required to cover the site with sufficient redundancy for subsequent processing using SfM. We acquired a total of 88 convergent images from the UAV, of which we discarded 24 due to blurring or unsuitable coverage. The use of imagery collected at convergent view angles, in conjunction with the use of well distributed GCPs, helps to reduce the risk of systematic 'doming' or 'dishing' errors within the resulting topographic data, which can occur as a consequence of inadequate self-calibrations of the camera lens models within the subsequent SfM process (Chandler et al., 2005; Wackrow and Chandler, 2011; Javernick et al., 2014; James and Robson, 2014; Woodget et al., 2015; Eltner et al., 2016). Whilst the small scale of the ground truth validation plots we use here (see subsequent section on 'Ground truth data') means that the effects of poor camera self-calibrations on our results are likely to be minimal, it is worth establishing good practice in this regard, especially if multiple applications of the data are intended.

Structure from motion digital photogrammetry

We imported the 64 chosen images into Agisoft's PhotoScan Professional digital photogrammetry software, and processed them to create a c. 1cm resolution orthophoto, a c. 2cm resolution digital elevation model (DEM), and dense 3D point cloud, all referenced to the local co-ordinate system using the GCPs and permanent markers. For further detail on the SfM process, readers are referred to Fonstad et al., (2013), Smith et al., (2015) and Eltner et al., (2016).

Ground truth data

For ground truthing purposes, we established 23 grain size sample plots along four exposed bars at Coledale Beck (Figure 1). Each plot measured 40cm x 40cm. This plot size was sufficiently large as to encompass the largest clasts within the field site, but sufficiently small to ensure substrate size was as uniform as possible within the plot itself. For each plot, a scaled, close-range photograph (e.g. Figure 1c) was acquired using a handheld camera. These photographs were then georeferenced in GIS to the site coordinate system, using a total station survey of each plot's four corners. Within these plots, a sample of clasts was selected for measurement using a 5cm x 5cm regular grid. Clasts falling beneath each grid node had their A- and B-axis dimensions measured from the scaled photograph, unless they were deemed unsuitable for measurement. Unsuitable clasts were those which were too small to measure at a scale of 1:1, those which were largely obscured by other clasts, and those which were not included fully within the photograph. Based on these data, we computed grain size statistics for each plot, including the mean, D_{50} (grain size of the 50^{th} percentile, or the median) and the D_{84} (grain size of the 84^{th} percentile). We did not collect any ground truth data in submerged areas and therefore our subsequent analyses are valid for subaerial gravel surfaces only.

Data analysis

Image texture

We used the technique developed by Carbonneau et al., (2004) to compute image texture from the orthophoto output. This empirical approach aims to establish a statistical correlation between a given measure of image texture and grain size. We computed image texture using a Matlab (Mathworks Inc.) routine on the red band of the imagery (this is an arbitrary choice and the method would also work on other bands). A square moving window with a kernel size of 41 pixels was passed over the image at intervals of five pixels (the routine requires a kernel size of an uneven number). A kernel size of 41 pixels is roughly equivalent to a kernel width of 41cm and was selected based on *a priori* knowledge that maximum clast sizes at Coledale Beck rarely exceed 40cm. We did not test other window sizes for the purposes of this short communication, however, we intend to explore this in subsequent research. We chose the interval size of five pixels as a compromise between detail and processing time. As a result, texture outputs are produced at 5cm resolution, but this could be altered as necessary. Within each kernel step, a measure of image texture is calculated and assigned to the central pixel. Image texture can be measured using a number of different metrics; in this case, we calculated the ‘negative image entropy’. This is a measure of image texture calculated using a grey level co-occurrence matrix (GLCM), i.e. a grey-tone spatial dependence probability distribution matrix first advocated by Haralick et al., (1973). The matrix provides the probabilities of all pairwise (*i, j*) combinations of pixel grey levels occurring within the specified moving window. The outputs are a function of the angular relationship between a single pixel and its neighbours (*V*), and the distance between them (the inter-pixel sampling distance, *D*). Negative image entropy provides a measure of randomness or the disorder of pixel values and is calculated according to Equation 1;

$$Negative\ Entropy = \sum_{i,j} P_{i,j}\,(-\log P_{i,j})$$

Equation 1

(after Haralick 1979)

Where *P* is the co-occurrence matrix of the image within each step of the moving window, based on the number of times that cells with grey levels *i* and *j* occur in two pixels separated by set distance *D* and direction *V*, divided by the total number of pixel pairs. We chose to use negative image entropy to compute image texture because the logarithmic component of algorithm (Equation 1) normalises extremes, thereby enhancing small variations in texture. Dugdale et al., (2010) suggested that entropy is therefore an appropriate measure to use where grain sizes are relatively small, as they are at our site, because small grain sizes tend to produce poorly defined light-dark boundaries. Other image texture operators are available however, and will be explored further in future.

The output is a map of negative image entropy, where higher values are returned for more textured or heterogeneous parts of the image and lower values for smoother or more homogeneous areas (Figure 3a). This image texture map was then imported into GIS to permit statistical comparison with the ground-truthing sample plots using linear regression.

Topographic point cloud roughness

We exported the dense point cloud of the Coledale site from PhotoScan Pro (Agisoft LLC) to the open source CloudCompare software (www.danielgm.net/cc/), and assessed the need for detrending, filtering and smoothing of the cloud. Detrending was found to be unnecessary but filtering and smoothing were required to reduce noise within the cloud (Figure 2). This noise can introduce roughness to the point cloud which does not result directly from grain size and therefore must be removed. A filtering and smoothing procedure was written in-house. We filtered the cloud by taking the mean of the interquartile range in elevation within 6mm x 6mm cells and smoothed the cloud by averaging the elevation values of each point by considering the elevation of all other points within a 2.5cm radius moving window. We performed a visual sensitivity check on the filtering cell size and smoothing window size, to ensure that sufficient noise was removed whilst preserving as much of the topographic detail within the cloud as possible.

Next, we used CloudCompare's inbuilt roughness tool to compute roughness values for each point in the smoothed and filtered cloud (CloudCompare, 2016). Roughness is defined as the shortest distance between each point in the cloud and the ordinary least squares best fitting plane computed on the nearest neighbours of that point, which fall within a spherical kernel of a specified size. This means that for each point in the cloud, a different ordinary least squares best fitting plane is generated, and thus a single roughness value is computed for each and every point within the cloud. The only case where this does not occur is when less than four points are present within the kernel, because a minimum number of three points are required to compute the least squares best fitting plane in addition to the one point for which roughness is being calculated. We found that only c. 0.0003% of kernels featured fewer than four points, with kernels comprising a maximum of 11,910 points and an average of 5668 points. A kernel radius of 20cm was chosen (i.e. a kernel width of 40cm), again based on *a priori* knowledge of typical grain sizes at Coledale Beck and to be comparable with the ground surface areas covered by the image texture interrogation window (41cm x 41cm) and validation plots (40cm x 40cm). Lastly, we created a raster of roughness outputs by averaging the roughness values computed for points in the cloud within 3cm pixels (Figure 3b). Sensitivity testing showed that rasterisation of the roughness data at smaller pixel sizes produced holes in the data where point density was low. A pixel size of 3cm therefore provided a good compromise for maximising resolution and minimising interpolation. We exported the raster to ArcGIS (ESRI, Inc.) and computed roughness statistics on a plot by plot basis for subsequent linear regression against the ground truth data.

Jack knife analysis

Linear regressions of image texture and roughness with grain size for each of our sample plots provide calibration relationships for predicting grain size over the wider area of interest. Validation is also required to assess the accuracy and precision of grain size estimates. We validated our calibration relationships using a jack knife approach (Quenouille 1949; Tukey 1958), an iterative method which excludes one ground truth plot at a time, and uses the linear regression equation based on the remaining plots to predict grain size for the excluded plot. We compared the measured grain size for each plot to the equivalent predicted grain size, to assess the strength of the predictive relationship. Measured grain sizes were also subtracted from the predicted grain sizes on a plot by plot basis to obtain residual error values. The average and standard deviation of the residuals for all plots are taken to represent the overall accuracy and precision of grain size estimates.

Results

Calibration and validation relationships for grain size predictions using image texture and roughness approaches are presented in Tables 2-3 and Figures 4-5. We found that maximum negative entropy correlated against average A axis length (Figure 4a) and average roughness values correlated against D_{84} of the B axes (Figure 4b) produced the strongest calibration relationships, as indicated by the co-efficients of determination in Table 2. Our results demonstrate that using the data for this site and at this scale, the point cloud roughness approach to grain size estimation gives both stronger calibration and validation relationships, as indicated by the slope and R^2 values in Table 3. Furthermore, Table 3 shows that grain sizes predicted using the roughness method are more than an order of magnitude more accurate than those predicted using the image texture method, as indicated by the mean of residual errors. Precision, represented by the standard deviation of residual errors, is greater than 0.01m for both approaches.

Discussion

Within this paper we have, for the first time, quantified the accuracy and reliability of an image texture *and* a topographic point cloud roughness approach to grain size quantification using UAV imagery and digital photogrammetry. The high resolution, quantitative, objective, spatially continuous, spatially explicit results are computed easily and have potential to aid our understanding of sediment dynamics and habitat heterogeneity at the mesoscale within a riverscape style framework (Fausch et al., 2002). However, our results raise three important and interlinked questions;

(1) Why does our image texture approach not produce calibration relationships of similar strength to those reported by others (e.g. de Haas et al., 2014, Tamminga et al., 2015)?

Weak calibration and validation relationships between image texture and grain size, and poor residual errors, may be a consequence of various factors, including (a) the use of an inappropriate texture operator, (b) the use of an inappropriate scale of analysis (i.e. kernel size and interval step), and/or (c) because image texture is also influenced by factors other than grain size. We have not explored variations in (a) or (b) for the purpose of this short communication, instead basing our choice of operator and scale of analysis on the findings of others (e.g. Carbonneau et al., 2004; Dugdale et al., 2010) and *a priori* knowledge of grain sizes at this site. However, the successful application of an image texture approach, based on UAV imagery or otherwise, will require further investigation of factors (a) and (b). This is especially true given the scale-dependent nature of the image based texture method.

In terms of (c), other factors influencing image texture might include the use of blurred imagery, the effects of local topographic shadowing and the presence of vegetation or water. Relative motion blur, a consequence of (i) increased exposure times resulting from cloudy conditions and (ii) wind gusts, are noted by de Haas et al., (2014) as a significant problem in predicting grain sizes using image texture. They note that quantitative correction of relative motion blur could not be conducted because their fixed-wing UAV was not equipped with the accelerometers necessary to provide correction data. The UAV used by de Haas et al., (2014) lacked a gimbal altogether (P. Carbonneau, *pers. comm.*), making image acquisition significantly more rudimentary than when using the 3-axis stabilisation mounts often available today. As a result, the approach of de Haas et al., (2014) is to side-step the issue by excluding any blurred sections of the orthophoto from further analysis, to achieve strong calibrations with grain size (R^2 = 0.82). However, such manual interventions can be time consuming and may result in inadequate site coverage or necessitate extra field time. Furthermore, the issue of image blurring remains unaddressed. Tamminga et al., (2015) find that shadows also disrupt calibration relationships by introducing high texture values in areas of pronounced topographic relief and vegetation, which in turn result in erroneously high grain size predictions. However, the 3-axis stabilised gimbal used on their Aeryon Scout UAV helps to reduce image blur, permitting another strong calibration with grain size (R^2 = 0.82). In this paper, we use a basic 1-axis camera gimbal on our UAV, which was flown in calm wind conditions. Whilst efforts were made to remove blurred images before photogrammetric processing, areas of blurring are evident on the resulting orthophoto (Figure 6), which is then used to develop the empirical calibration with grain size. Alongside the minor influence of vegetation presence within some of the ground truth plots, we expect that this gimbal is a key reason for the poorer calibration with grain size than reported elsewhere. However, further dedicated testing is required to prove this, and subsequently to reduce the incidence of blurring or improve our ability to detect and eliminate it from images. Sieberth et al., 2013 and Sieberth et al., 2016, provide some initial work on blur detection and removal.

(2) Why does our topographic approach using point cloud roughness perform so much better than our image texture approach (Table 3)?

Our topographic (point cloud roughness) method was conceived out of a need to move away from the adverse effects of blurred UAV imagery. Given that exactly the same UAV imagery is used as input for both texture and roughness approaches though, we might expect the roughness approach to be adversely affected by blur too. The SfM-photogrammetry process computes indirect measures of elevations using UAV image parallax, to create a point cloud. Thus, where image quality is poor (e.g. due to blurring) or lacking in texture (e.g. spectrally homogeneous areas) then greater amounts of noise (i.e. erroneous point matches) are likely to be observed within the point cloud. More generally, we would expect other factors to influence the point cloud roughness-grain size relationship, including;

- Presence of vegetation – where topographic variation in the point cloud is not a result of variation in grain size.

- Interstitial spaces between large clasts which are occupied by smaller clasts - where topographic variation is high within the extent of the kernel but grain size is low.
- Complex levels of topographic variation over short distances – where features such as footprints introduce variation which does not result from grain size and cannot be removed easily by detrending.
- Packing and imbrication of clasts – where partially buried clasts do not produce the same topographic signature as exposed clasts of equivalent size, a well-known issue for a number of grain size quantification methods (e.g. Church et al., 1987; Sime and Ferguson, 2003; Heritage and Milan, 2009; Picco et al., 2013).

Despite these complicating factors, we are still able to predict grain sizes with exceptionally low mean residual errors (<1mm). This may be because the listed factors do not have a significant impact in the location of our ground truth plots, or that their effect is instead observed in the less encouraging precision metric (standard deviation >10mm). We also believe that the smoothing and filtering procedures described earlier are partly responsible for this success of our topographic point cloud roughness approach. However, the generic nature of the two different methods we have tested here also deserves attention. According to Buscombe (2016), roughness can be defined as *"a measure of the statistical variation in the distribution of topographic relief of a surface"*, and texture as *"the frequency of change and arrangement of roughness"* (p.93). In other words, we might consider topographic roughness (i.e. point cloud roughness) to be a function of variation in all three dimensions, whilst image texture relates to variation solely in the horizontal dimension. Thus, at the mesoscale level of assessment we consider here, our results suggest that grain size is more strongly related to variation in 3D topographic relief, than it is to the horizontal arrangement of roughness as expressed by the image texture. Whether this pattern holds true at different scales of assessment is uncertain, and deserves further research. Texture may prove to be a better predictor of horizontal patterns, such as the rate of change in grain size or bedforms, or of grain shape, orientation, inclination, spacing or clustering.

(3) Which remote sensing approach is "best" for quantifying fluvial grain size?

The simple answer to this question is that it depends on the application at hand. The accuracy and precision of our results for our novel topographic (point cloud roughness) method indicates that they are roughly in line with or better than other remote sensing approaches for grain size quantification (Table 1), including other UAV based approaches (e.g. Westoby et al., 2015). The spatial resolution of our outputs is also finer than those approaches with similar mean accuracy levels (Table 1). However, we note that the slope of the observed versus predicted relationship for point cloud roughness (0.777, Figure 5) is lower than those reported by Carbonneau et al., (2004) and Carbonneau et al., (2005b) for the use of an image texture approach on imagery of a different scale acquired from a manned aircraft (Table 1). We anticipate that platform stability and image clarity may be responsible for this difference. Ultimately, the choice of the "best" method for quantifying fluvial substrate size will be determined by the specific requirements of a given application, including the required scale, spatial coverage, accuracy, precision, data acquisition and processing times and costs. At present, our point cloud roughness approach is best suited to studies requiring coverage of up to c. 1km channel length with spatial resolutions of a few centimetres, where multiple flight passes can be undertaken in order to acquire convergent imagery for SfM processing (whereas the texture approach can be conducted on a single image). With rapid and on-going developments in UAV, gimbal, sensor and software technology as well as associated processing algorithms, we anticipate that covering larger areas with greater detail and at lower costs will only become more practicable with time.

Future work

Future research should aim to reduce the impact of image blur on both image texture and point cloud roughness approaches. For example, we intend to compare the results obtained using different camera gimbals and conduct sensitivity analysis to determine the optimal kernel sizes and operators for calculating image texture and point cloud roughness. Further consideration of scale and quantification of the range of grain sizes which can be predicted accurately and reliably is also of importance. For instance, the use of the 2.5cm radius smoothing kernel means that reliable prediction of grain sizes smaller than 5cm is compromised at present. A reduction of image blur should reduce point cloud noise and thereby permit a smaller smoothing kernel size to be used and enable prediction of smaller grain sizes. Additionally, we might obtain different results by using imagery of different resolutions over different spatial scales. Such enhanced research is necessary to help us fully understand the potential for upscaling and transferability of this method to different fluvial settings and other environments, including submerged areas.

Conclusion

Within this letter, we have provided an initial quantitative assessment of two different approaches to subaerial gravel size measurement using UAV imagery processed with SfM digital photogrammetry. We flew a rotary-winged UAV over a gravel bed river in the English Lake District and processed the resultant imagery into an orthophoto, DEM and point cloud. We developed an empirical relationship between grain size validation data and (a) a measure of image texture and (b) topographic roughness of the SfM point cloud. Our error assessment reveals poor calibration and validation results for the texture approach, as well as poor accuracy and precision of grain size estimates. We suspect this may result from the use of blurred imagery caused by an inadequate camera gimbal, the use of a suboptimal texture operator or window size, or that the texture method is not well suited to studies at the mesoscale. Conversely, point cloud roughness is much better correlated with grain size at this scale of assessment and produces much lower mean errors. Whilst smoothing and filtering of the point cloud has permitted very accurate grain size estimations on a plot-by-plot basis, precision is weaker, highlighting the need for improvements to the reliability of this roughness method. The use of either technique requires careful consideration of (a) potential error sources and (b) the appropriate scales at which each method can be applied. With further work in these areas, the methods we have presented here have potential to be of value to a range of research and management applications, both within fluvial systems and beyond.

Acknowledgments

This research was funded by the University of Worcester (PhD studentship for Amy Woodget) and a student research award from the Geological Remote Sensing Group, a special interest group of The Geological Society of London and the Remote Sensing and Photogrammetry Society. We are very grateful to Patrice Carbonneau for the image texture processing, to Fleur Visser, Richard Johnson, James Atkins, Andy Skellern, Jeff Warburton, Carl Greenman and Milo Creasey for field assistance, and to the landowners for permitting us to access the site. Constructive comments from Dan Buscombe and Normand Bergeron were gratefully received on an earlier version of the manuscript.

References

Adams, J. 1979. Gravel size analysis from photographs. Journal of the Hydraulics Division, Proceedings of the American Society of Civil Engineers, 105, HY10: 1247-1255

Brasington, J., Vericat, D. and Rychov, I. 2012. Modeling river bed morphology, roughness, and surface sedimentology using high resolution terrestrial laser scanning. Water Resources Research 48 W11519, doi: 10.1029/2012WR012223

Buscombe, D. 2008. Estimation of grain size distributions and associated parameters from digital images of sediment. Sedimentary Geology 210: 1-10

Buscombe, D. 2013. Transferable wavelet method for grain-size distribution from images of sediment surface and thin section, and other natural granular patterns. Sedimentology 60 (7): 1709-1732.

Buscombe, D. 2016. Spatially explicit spectral analysis of point clouds and geospatial data. Computers and Geosciences 86: 92-108

Buscombe, D. and Masselink, G. 2009. Grain-size information from the statistical properties of digital images of sediment. Sedimentology 56: 421-438

Buscombe, D. Rubin, D.M. and Warrick, J.A. 2010. A universal approximation of grain size from images of noncohesive sediment. Journal of Geophysical Research 115 F02014 doi: 10.1029/2009JF001477

Butler, J.B., Lane, S.N, and Chandler, J.H. 2001. Automated extraction of grain-size data from gravel surfaces using digital image processing. Journal of Hydraulic Research 39 (4): 519-529

Carbonneau, P.E., Lane, S.N. and Bergeron, N. 2004. Catchment-scale mapping of surface grain size in gravel bed rivers using airborne digital imagery. Water Resources Research 40, W07202, doi:10.1029/2003WR002759

Carbonneau, P.E., Bergeron, N.E. and Lane, S.N. 2005a. Texture-based segmentation applied to the quantification of superficial sand in salmonids river gravels. Earth Surface Processes and Landforms 30: 121-127

Carbonneau, P.E., Bergeron, N. and Lane, S.N. 2005b. Automated grain size measurements from airborne remote sensing for long profile measurements of fluvial grain sizes. Water Resources Research 41, W11426, doi:10.1029/2005WR003994

Carbonneau, P.E., Fonstad, M.A., Marcus, W.A., and Dugdale, S.J. 2012. Making riverscapes real. Geomorphology 137 (1): 74-86

Centre for Ecology and Hydrology. 2001. Further validation of PHABSIM for the habitat requirements of salmonid fish. Final project report to Environment Agency (W6-036) and CEH (C00962)

Church, M.A. McLean, D.G. and Wolcott, J.F. 1987. River bed gravels: Sampling and Analysis, In Thorne, C.R., Bathurst, J.C. and Hey, R.D. (Eds) Sediment Transport in Gravel-bed Rivers, John Wiley and Sons, Chichester

CloudCompare. 2016. User Manual (version 2.6.1) Available online: www.cloudcompare.org/doc/qCC/CloudCompare%20v2.6.1%20-%20User%20manual.pdf (accessed 12.12.2016)

De Haas, T., Ventra, D., Carbonneau. P. and Kleinhans, M.G. 2014. Debris flow dominance of alluvial fans masked by runoff reworking and weathering. Geomorphology 217: 165-181

Dugdale, S.J., Carbonneau, P.E. and Campbell, D. 2010. Aerial photosieving of exposed gravel bars for the rapid calibration of airborne grain size maps. Earth Surface Processes and Landforms 35: 627-639

Eltner, A., Kaiser, A., Castillo, C., Rock, G., Neugirg, F. and Abellán, A. 2016. Image-based reconstruction in geomorphometry – merits, limits and developments. Earth Surface Dynamics 4: 359-389

Entwistle, N.S. and Fuller, I.C. 2009. Terrestrial laser scanning to derive the surface grain size facies character of gravel bars. In Heritage, G.L. and Large, A.R.G. (Eds) Laser Scanning for the Environmental Sciences, Wiley-Blackwell, London

European Commission. 2000. Directive 2000/60/EC of the European Parliament and of the Council of 23rd October 2000: Establishing a framework for Community action in the field of water policy. Official Journal of the European Communities, Brussels, 22.12.2000, L327: 1-72

Evans, L.J. and Norris, R.H. 1997. Prediction of benthic macroinvertebrate composition using microhabitat characteristics derived from stereo photography. Freshwater Biology 37: 621-633

Fausch, K.D., Torgersen, C.E., Baxter, C.V. and Hiram, L.W. 2002. Landscapes to riverscapes: bridging the gap between research and conservation of stream fishes. BioScience 52 (6): 483-498

Fonstad, M.A., Dietrich, J.T., Courville, B.C., Jensen, J.L. and Carbonneau, P.E. 2013. Topographic structure from motion: a new development in photogrammetric measurement. Earth Surface Processes and Landforms 38 (4): 421-430

Frissell, C.A., Liss, W.J., Warren, C.E. and Hurley, M.D. 1986. A hierarchical framework for stream habitat classification: viewing streams in a watershed context. Environmental Management 10 (2): 199-214

Garcia, A., Jorde, K., Habit, E., Caamano, D. and Parra, O. 2011. Downstream environmental effects of dam operations: changes in habitat quality for native fish species. River Research and Applications 27: 312-327

Goodwin, P., Jorde, K., Meier, C. and Parra, O. 2006. Minimizing environmental impacts of hydropower development: transferring lessons from past projects to a proposed strategy for Chile. Journal of Hydroinformatics 8: 253–270

Graham, D.J., Reid, I. and Rice, S.P. 2005a. Automated sizing of coarse-grained sediments: image- processing procedures. Mathematical Geology 37(1): 1-28

Graham, D.J., Rice, S.P. and Reid, I. 2005b. A transferable method for the automated grain sizing of river gravels. Water Resources Research 41, W07020, doi:10.1029/2004WR003868

Habit, E., Belk, M.C. and Parra, O. 2007. Response of the riverine fish community to the construction and operation of a diversion hydropower plant in central Chile. Aquatic Conservation: Marine and Freshwater Ecosystems 17: 37-49

Haralick, R.M., Shanmugam, K. and Dinstein, I. 1973. Textural features for image classification. IEEE Transactions on Systems, Man and Cybernetics SMC-3 (6): 610-621

Haralick, R.M. 1979. Statistical and structural approaches to texture. Proceedings of the IEEE 67 (5): 786-804

Heritage, G.L. and Milan, D.J. 2009. Terrestrial laser scanning of grain roughness in a gravel-bed river. Geomorphology 113: 4-11

Hey, R.D. and Thorne, C.R. 1983. Accuracy of surface samples from gravel bed materials. Journal of Hydraulic Engineering 109 (6): 842-851

Hodge, R., Brasington, J. and Richards, K. 2009. In situ characterization of grain-scale fluvial morphology using Terrestrial Laser Scanning. Earth Surface Processes and Landforms 34: 954-968

Ibbeken, H. and Schleyer, R. 1986. Photo-sieving: a method for grain size analysis of coarse-grained, unconsolidated bedding surfaces. Earth Surface Processes and Landforms 11: 59-77

Keeley, E.R. and Slaney, P.A. 1996. Quantitative measures of rearing and spawning habitat characteristics for stream-dwelling salmonids: guidelines for habitat restoration. Province of British Columbia, Ministry of Environment, Lands and Parks, and Ministry of Forests. Watershed Restoration Project Report 4

Leopold, L.B. 1970. An improved method for size distribution of stream bed gravel. Water Resources Research 6 (5): 1357-1366

McEwan, I. K., Sheen, T.M., Cunningham, G.J. and Allen, A.R. 2000. Estimating the size composition of sediment surfaces through image analysis. Proceedings of the Institution of Civil Engineers – Water and Maritime Engineering 142: 189–195

Milan, D.J. 1996. The application of freeze-coring for siltation assessment in a recently regulated stream. Hydrologie dans les pays celtiques, Rennes, France 8-11 July 1996. Ed. INRA Paris (Les Colloques 79)

Milan, D.J. and Heritage, G.L. 2012. LiDAR and ADCP use in gravel-bed rivers: Advances since GBR6. In Church, M., Biron, P. and Roy, A. (Eds) Gravel-bed Rivers: Processes, Tools, Environments, Wiley-Blackwell, Chichester

Newson, M.D. and Newson, C.L. 2000. Geomorphology, ecology and river channel habitat: mesoscale approaches to basin-scale challenges. Progress in Physical Geography 24 (2): 195-217

Orr, H. and Brown, D. 2004. Bassenthwaite Lake Geomorphological Study Findings: Summary Report, Environment Agency Publication Reference ScNW0904BIGO-E-P

Picco, L., Mao, L., Cavalli, M., Buzzi, E., Rainato, R. and Lenzi, M.A. 2013. Evaluating short-term morphological changes in a gravel-bed braided river using terrestrial laser scanner. Geomorphology 201: 323-334

Quenouille, M.H. 1949. Approximate tests of correlation in time-series. Journal of the Royal Statistical Society Series B 11: 68-84

Rice, S. and Church, M. 1996. Sampling surficial fluvial gravels: the precision of size distribution percentiles estimates. Journal of Sedimentary Research 66 (3): 654-665

Rubin, D.M. 2004. A simple autocorrelation algorithm for determining grain size from digital images of sediment. Journal of Sedimentary Research 74 (1): 160-165

Rychov, I., Brasington, J. and Vericat, D. 2012. Computational and methodological aspects of terrestrial surface analysis based on point clouds. Computers and Geosciences 42: 64-70

Sieberth, T., Wackrow, R. and Chandler, J.H. 2013. Automation isolation of blurred images from UAV image sequences. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, XL-1/W2: 361-366

Sieberth, T., Wackrow, R. and Chandler, J.H. 2016. Automatic detection of blurred images in UAV image sets. ISPRS Journal of Photogrammetry and Remote Sensing 122: 1-16

Sime, L. C. and Ferguson, R. I. 2003. Information on grain sizes in gravel-bed rivers by automated image analysis. Journal of Sedimentary Research 73 (4): 630-636

Smith, M.W., Carrick, J.L. and Quincey, D.J. 2015. Structure from motion photogrammetry in physical geography. Progress in Physical Geography 40 (2): 247-275

Tamminga, A., Hugenholtz, C., Eaton, B. and LaPointe, M. 2015. Hyperspatial remote sensing of channel reach morphology and hydraulic fish habitat using an unmanned aerial vehicle (UAV): A first assessment in the context of river research and management. River Research and Applications 31 (3): 379-391

Tukey, J.W. 1958. Bias and confidence in not-quite large samples. Annals of Mathematical Statistics 29: 614

Verdú, J.M., Batalla, R.J. and Martinez-Casasnovas, J.A. 2005. High-resolution grain-size characterisation of gravel bars using imagery analysis and geo-statistics. Geomorphology 72: 73-93

Wentworth, C.K. 1922. A scale of grade and class terms for clastic sediments. Journal of Geology 30: 377-392

Westoby, M.J., Dunning, S.A., Woodward, J., Hein, A.S., Marrero, S.M., Winter, K. and Sugden, D.E. 2015. Sedimentological characterization of Antarctic moraines using UAVs and Structure-from-Motion photogrammetry. Journal of Glaciology 61 (230): 1088-1102

Wise, D.H. and Molles, M.C. 1979. Colonisation of artificial substrate by stream insects: influence of substrate size and diversity. Hydrobiologia 65 (1): 69-74

Wolman, M.G. 1954. A method of sampling coarse river-bed material. Transactions of the American Geophysical Union 35 (6): 951-956

Woodget, A.S., Visser, F., Maddock, I.P. and Carbonneau, P. 2016. Quantifying fluvial substrate size using hyperspatial resolution UAS imagery and SfM-photogrammetry. Extended Abstract, 11th International Symposium on Ecohydraulics, Melbourne, Australia, 7-12 February

Table 1. An overview of remote sensing methods for quantifying fluvial grain sizes.

Method	Theory	Typical survey extent	Typical spatial resolution of grain size estimates	Typical accuracy of grain size estimates	Typical precision of grain size estimates	Typical slope (obs v. pred)	Limitations	References
Close-range photo-sieving	Manual or automated analyses of photos acquired from tripod-mounted cameras to measure individual grains	Patch level (microscale)	< 1cm	<0.25 phi	0.11-0.25 phi	1.25	Segmentation approach can result in over-segmentation of some grains (leading to an underestimation of true grain sizes) and under-segmentation of others (resulting in an overestimation of true grain sizes)	Adams, 1979; Ibbeken and Schleyer, 1986; Butler et al., 2001; Sime and Ferguson, 2003; Graham et al., 2005a; Graham et al., 2005b
Statistical image analysis	Use of the frequency (spectral) content of images to quantify grain sizes	Patch level (microscale)	<1mm	<3mm	<3mm	0.77-1.12	Extensive site-specific look-up data required for calibration by some approaches (indicated by *) and scaling is required by all	Rubin, 2004*; Buscombe, 2008*; Buscombe and Masselink, 2009*; Buscombe et al., 2010; Buscombe and Rubin, 2012; Buscombe 2013
Image textural analysis	Computed image textural variables are correlated with field measures from small patches	Reach to catchment level	c. 1m	3-8mm	13.9-29mm	1.03-1.23	Labour intensive and time consuming collection of field data required for calibration purposes	Carbonneau et al., 2004; Carbonneau et al., 2005a; Carbonneau et al., 2005b; Verdú et al., 2005
Terrestrial laser scanning	(i) Roughness (standard deviation) of laser-derived point clouds or (ii) segmentation of grey-level images derived from DEMs are used to estimate grain sizes	Patch (microscale) to reach level	c. 5cm	c. 1mm	2.34cm	0.5261	Requires significant field and processing efforts to cover large areas (including de-trending)	McEwan et al., 2000; Entwistle and Fuller, 2009; Heritage and Milan, 2009; Hodge et al., 2009; Brasington et al., 2012; Milan and Heritage, 2012; Rychov et al., 2012

Table 2. Co-efficients of determination (R^2 values) for the regression of a range of grain size metrics with maximum image texture and average point cloud roughness. The strongest calibration relationship for each method is highlighted in bold text.

Grain size metric		**Image texture (maximum)**	**Point cloud roughness (average)**
A axis	D84	0.3963	0.7881
	D50	0.3812	0.5095
	D mean	**0.4787**	0.7265
B axis	D84	0.2985	**0.7987**
	D50	0.3765	0.7032
	D mean	0.4400	0.7615

Table 3. Comparison of calibration, validation and residual errors between the image texture and point cloud roughness approaches to grain size quantification.

Method		**Image texture (maximum)**	**Point cloud roughness (average)**
Grain size metric		**Mean of A axis**	**D84 of B axis**
Calibration	R^2	0.4787	0.7987
	Slope	0.0005	12.349
	Intercept	-0.3064	-0.0029
Validation	R^2	0.2169	0.7554
	Slope	0.4393	0.777
	Intercept	0.0246	0.0117
Residual errors	Mean (m)	-0.0032	-0.0001
	Standard deviation (m)	0.0262	0.0184