Geomorphology 205 (2014) 102–119

Contents lists available at ScienceDirect

Geomorphology

journal homepage: www.elsevier.com/locate/geomorph

Delineation of river bed-surface patches by clustering high-resolution spatial grain size data

Peter A. Nelson [a,*], Dino Bellugi [b,c], William E. Dietrich [b]

[a] Department of Civil and Environmental Engineering, University of Genova, via Montallegro 1, 16145 Genova, Italy
[b] Department of Earth and Planetary Science, University of California, Berkeley, CA 94720, USA
[c] Department of Computational Science and Engineering, University of California, Berkeley, CA 94720, USA

ARTICLE INFO

Article history:
Received 30 June 2011
Received in revised form 4 June 2012
Accepted 6 June 2012
Available online 18 June 2012

Keywords:
Fluvial geomorphology
Patchiness
Agglomerative clustering
Spectral clustering
k-Means clustering
Fuzzy clustering

ABSTRACT

The beds of gravel-bed rivers commonly display distinct sorting patterns, which at length scales of ~0.1 − 1 channel widths appear to form an organization of patches or facies. This paper explores alternatives to traditional visual facies mapping by investigating methods of patch delineation in which clustering analysis is applied to a high-resolution grid of spatial grain-size distributions (GSDs) collected during a flume experiment. Specifically, we examine four clustering techniques: 1) partitional clustering of grain-size distributions with the *k*-means algorithm (assigning each GSD to a type of patch based solely on its distribution characteristics), 2) spatially-constrained agglomerative clustering ("growing" patches by merging adjacent GSDs, thus generating a hierarchical structure of patchiness), 3) spectral clustering using Normalized Cuts (using the spatial distance between GSDs and the distribution characteristics to generate a matrix describing the similarity between all GSDs, and using the eigenvalues of this matrix to divide the bed into patches), and 4) fuzzy clustering with the fuzzy *c*-means algorithm (assigning each GSD a membership probability to every patch type). For each clustering method, we calculate metrics describing how well-separated cluster-average GSDs are and how patches are arranged in space. We use these metrics to compute optimal clustering parameters, to compare the clustering methods against each other, and to compare clustering results with patches mapped visually during the flume experiment.

All clustering methods produced better-separated patch GSDs than the visually-delineated patches. Although they do not produce crisp cluster assignment, fuzzy algorithms provide useful information that can characterize the uncertainty of a location on the bed belonging to any particular type of patch, and they can be used to characterize zones of transition from one patch to another. The extent to which spatial information influences clustering leads to a trade-off between the quality of GSD separation between patch types and the spatial coherence of patches. Methods incorporating spatial information during the clustering process tended to produce a finite number of types of patches. As methods improve for collecting high-resolution grain size data, the approaches described here can be scaled up to field studies to better characterize the grain size heterogeneity of river beds.

1. Introduction

The beds of gravel bed rivers frequently display considerable spatial heterogeneity in the grain size and sorting. Variations in the surface composition of river beds occur at all scales, but once they reach a scale of ~0.1 − 1 channel widths, they are referred to as textural patches or facies (e.g., Bluck, 1971; Bridge and Jarvis, 1976; Forbes, 1983; Dietrich and Smith, 1984; Kinerson, 1990; Wolcott and Church, 1991; Lisle and Madej, 1992; Paola and Seal, 1995; Sambrook Smith and Ferguson, 1995; Crowder and Diplas, 1997; Buffington and Montgomery, 1999; Dietrich et al., 2005; Yarnell et al., 2006; Nelson et al., 2009, 2010). Sorted areas of river beds that are temporally and spatially stable ("forced" patches (Nelson et al., 2010)) remain spatially persistent through time, despite passing considerable sediment load (see discussion in Dietrich et al., 2005). These patches emerge as a consequence of the complex interaction between bed topography, the flow field, and the local sediment transport field, wherein topographically-forced local divergences in boundary shear stress are compensated by local divergences in bedload transport (Dietrich and Smith, 1984; Dietrich, 1987), which under low excess stress conditions, commonly encountered in gravel bed rivers, are achieved through selective transport and local adjustment of bed-surface grain size (Clayton and Pitlick, 2007, 2008; Nelson et al., 2010).

Bed-surface patches at this scale affect physical and biological processes through the influence on local near-bed flow fields and rates of

* Corresponding author.
E-mail address: peteranelson@gmail.com (P.A. Nelson).

0169-555X/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.geomorph.2012.06.008

sediment transport. A characteristic local grain size (e.g., the D_{84}, the grain size for which 84% of the sediment is finer) is often used to parameterize local roughness in hydrodynamic calculations for flow and boundary shear stress (e.g. Leopold and Wolman, 1957; Dietrich and Whiting, 1989; Wilcock, 1996). This hydrodynamic feedback between grain size and boundary shear stress has been shown to affect particle mobility (e.g., Venditti et al., 2010) and the formation of bedload sheets (Seminara et al., 1996). Current mixed-grain-size bedload transport algorithms (Parker, 1990; Wilcock and Crowe, 2003) require a bed-surface grain-size distribution for input, in part because they employ empirical "hiding functions" to determine the critical stress for mobility for each grain size as a function of the local bed-surface grain-size distribution. Bed-surface patchiness has been invoked as a potential cause of downstream fining (Paola and Seal, 1995) and as an important source of error in one-dimensional calculations of bedload transport (Ferguson, 2003). Numerical simulations of channel morphodynamics with mixed-grain-size sediment transport have shown that bed-surface patches interact with the evolving bed and, through the effect on the flow field, can profoundly influence channel evolution and steady state bed morphology (Nelson, 2010; Nelson et al., 2011). Patchiness also has biological implications, because many aquatic organisms prefer microhabitats (Cummins and Lauff, 1969; Rabeni and Minshall, 1977; Reice, 1980) or spawning grounds (Kondolf and Wolman, 1993; Overstreet et al., 2010; Riebe et al., 2010) consisting of particular grain sizes, and field studies have demonstrated connections between morphological units and spawning activity (e.g., Moir and Pasternack, 2008; Senter and Pasternack, 2011).

Although the grain-size distribution of the surface of a stream bed varies continuously through geographic space, it can be advantageous to classify locally-similar regions of the bed into patches of grain size. For example, the accuracy of flow models or calculations of mixed-grain-size sediment transport (Parker, 1990; Wilcock and Crowe, 2003) where bed-surface heterogeneity has been discretized into representative types of patches should exceed that of calculations that assume a width-averaged or reach-averaged bed-surface grain-size distribution (e.g., Ferguson, 2003; Nelson et al., 2009), but the patch-based calculations will be simpler to implement than calculations requiring continuously-varying, spatially-distributed information of grain sizes. Additionally, patch-based representations of heterogeneity are frequently used in landscape and fluvial ecology (e.g., Cooper et al., 1997; Gustafson, 1998; Wu et al., 2000; Turner et al., 2001; Ahlqvist and Shortridge, 2010), and representation of bed-surface heterogeneity as categorical patch data may allow geomorphic characteristics to be more readily incorporated into existing frameworks of ecological or landscape heterogeneity. For instance, Yarnell et al. (2006) used FRAGSTATS (McGarigal et al., 2002), software popularly used to compute indices and metrics used in landscape ecology, to measure the Shannon's Diversity Index (SHDI) for grain size patch maps of river channels and flumes (e.g., Kinerson, 1990; Lisle et al., 1993) and compared the SHDI with relative sediment supply q^* (Dietrich et al., 1989) to show that increased sediment supply tends to increase habitat heterogeneity.

Although numerous studies document bed patches or facies (e.g., Nelson et al., 2009, and references therein), the techniques used to delineate the boundaries of patches tend to be visual, semi-quantitative, and site-specific (e.g., Kinerson, 1990; Lisle and Madej, 1992), which can pose potential problems of repeatability, transferability, and precision (e.g., Poole et al., 1997). To make maps of patches more objective and reliable, techniques have been proposed to combine visual identification of facies and quantitative measurement of grain sizes (Kondolf and Li, 1992; Buffington and Montgomery, 1999). A possibly better approach for classification and delineation of patches would be the development of algorithms that sweep over high-resolution grain-size data and automatically delineate boundaries of patches based on some classification or clustering criteria. Whereas some efforts have been made to apply moving-window techniques to grid-based sediment samples (Crowder and Diplas, 1997), traditionally such purely data-driven approaches to the delineation of patches have not been feasible because of the extreme effort required to obtain grain-size data at sufficient resolution.

Recently, however, significant progress has been made in developing methods to measure bed-surface grain size and roughness characteristics from photographs and remotely sensed data. Photographic methods that use statistics such as semivariograms of image texture (Verdú et al., 2005) or the autocorrelation (Rubin, 2004; Barnard et al., 2007; Warrick et al., 2009), fractal dimension (Buscombe and Masselink, 2009), or spectral decomposition (Buscombe et al., 2010) of image intensity to estimate grain sizes have been applied to close-up photos of sediment (Barnard et al., 2007) or to images captured from aerial platforms (Carbonneau et al., 2004, 2005); other image processing methods seek to outline and measure individual grains in an image (Butler et al., 2001; Sime and Ferguson, 2003; Graham et al., 2005a, 2005b, 2010). Similar methods have been developed to extract grain sizes (McEwan et al., 2000) and roughness properties (Katul et al., 2002; Aberle and Nikora, 2006; Heritage and Milan, 2009; Aberle et al., 2010) from high-resolution digital elevation models.

Whatever the technique used, the ultimate output of a procedure of this kind will be a dataset of grain sizes of high spatial resolution, in the form of either a field of point measurements of grain sizes or an array of grain-size distributions. For the purposes of this study, we now ask: how can we draw boundaries on this field of grain sizes to delineate a meaningful set of patches?

Here, we explore how clustering methods, when applied to a high-resolution spatial grid of grain-size distributions, can be used to automatically delineate bed-surface patches. Clustering is of great interest to computer scientists and statisticians concerned with pattern recognition, data mining, image segmentation, and machine learning (among other areas). In a general sense, clustering is the unsupervised classification of patterns (observations, data items, or feature vectors) into groups (clusters) (Jain et al., 1999). Jain (2009) provides the following operational definition of clustering: "Given a *representation* of n objects, find K groups based on a measure of *similarity* such that the similarities between objects in the same group are high while the similarities between objects in different groups are low." This is precisely the objective of river bed-surface patch delineation: a facies map of a river bed should divide the bed into patches such that the sediment comprising an individual patch is as homogeneous as possible whereas the grain-size distributions of adjacent patches are as different as possible.

In our analysis, we seek answers to the following questions: (1) Are patches produced from data-driven, unsupervised clustering methods better defined than those produced from visual mapping? (2) Is a trade-off between grain size separation and spatial coherency of patches reflected in the degree to which clustering methods incorporate spatial information? And (3) how certain can we be that a particular bed location belongs to just one patch? We use partitional, hierarchical, spectral, and fuzzy clustering methods to generate maps of patches from a grain size dataset collected during a flume experiment in which channel topography and sorting patterns on the bed surface were extensively documented. From the resulting patch maps, we calculate metrics that describe the spatial arrangement of patches and the similarities and differences between the grain-size distributions of different types of patches. Our results suggest that when these metrics are used to guide the choice of appropriate parameters, clustering methods generate better separation of grain-size distributions between patches than visual mapping. They also show that the inclusion of spatial constraints in the clustering process results in delineation of a finite number of types of patches whose characteristic grain-size distributions are less well-separated, but whose spatial arrangement is more coherent, than patches produced by methods that neglect spatial location.