A refined automated grain sizing method for estimating river-bed grain size distribution of digital images

Chang-Han Chung, Fi-John Chang*

Department of Bioenvironmental Systems Engineering, National Taiwan University, Taipei 10617, Taiwan, ROC

1. Introduction

Grain size distribution is fundamentally important to fluvial environment such as sediment transport, river channel behavior, bed-load flux and ecological habitat (Hoey and Ferguson, 1994; Mao and Surian, 2010; Parker, 1991; Rice and Church, 1998). Natural fluvial materials are of heterogeneous sizes and shapes; consequently it is a great challenge to quantify the grain size distribution. Many field measurement methods of river-bed materials have been established, such as pebble-counting, volume-by-number, grid-by-number and area-by-number methods (Bunte and Abt, 2001; Kellerhals and Bray, 1971; Rice and Haschenburger, 2004). The conventional field measurement of grain sizes, however, is labor-intensive and time-consuming. It is highly desirable to develop measurement techniques that provide reliable estimation of grain size distribution while simultaneously reduce the time spent in both field and laboratory.

Digital images (captured by digital cameras) and the development in computer software have opened up the possibility of automated image analysis and thus facilitated the progress toward grain size distribution. The history of the automated grain sizing (AGS) is closely related to the continuous improvement in photographic equipments. The early emulsion-based photography was introduced to reduce the time spent on fieldwork, despite that the photographic analysis was time-consuming (Adams, 1979; Rice and Church, 1998). Even though Ibbeken and Schleyer (1986) and Ibbeken et al. (1998) presented more effective approaches to photographic analysis, those approaches still required a large amount of time for extracting necessary information from photographs in the laboratory. Recent advances in image processing techniques can automatically characterize grain sizes from digital images and has shown promise as a viable method for measuring gravels and fluvial sediments (Butler et al., 2001; Graham et al., 2005a; Chang and Chung, 2012). In general, there are two major approaches to image-based analysis (Graham et al., 2010). The first approach calculates the spectral characteristics or semivariance structure of an image to provide ensemble grain size, such as the $D_{50}$ in an area (Adams et al., 2007; Buscombe, 2008; Buscombe and Masselink, 2009; Carbonneau, 2005; Verdu et al., 2005). The second approach is based on image segmentation principles to obtain grain-size measurements in an image and thereby to provide its grain size distribution (Graham et al., 2005b; Sime and Ferguson, 2003). This study focuses on automatically obtaining grain size information from captured images with reference to image segmentation methods such as automated grain sizing (AGS) by Graham et al. (2005b) and Strom et al. (2010).

The AGS methods possess several advantages such as a reduction in processing time, non-intrusive sampling and the avoidance of operator bias. The sizes of natural river-bed materials, however,
vary substantially. According to Bunte and Abt (2001), grain sizes can be mainly classified into four categories: sand (0.063–2 mm); gravels (2–64 mm); cobbles (64–256 mm); and boulders (larger than 256 mm). Previous studies (Butler et al., 2001; Chang and Chung, 2012; Sime and Ferguson, 2003) could deal with various categories of grain sizes including small gravels up to cobbles, while Graham et al. (2005a) did investigate a sample composed of grains including boulders. We intend to develop a viable tool that can precisely estimate grain sizes in a great range from gravels to boulders (16–512 mm) on account of the minimum resolvable grain size of the captured images in which small grains less than 16 mm and/or sandy areas can be automatically recognized as the background.

Early studies showed that binary thresholds could be applied to gray-level images for deriving binary images composed of white colored grains and black colored sandy areas (background). Moreover, many studies indicated that the binary threshold is a crucial parameter to the AGS procedure (Butler et al., 2001; Chen et al., 2008; Graham et al., 2005a; Strom et al., 2010) and its value for each digital image needs to be properly determined. Consequently, how to adequately identify the value of the binary threshold for each digital image is a vital task and would one-step further make the AGS procedure more applicable and less laborious. The captured images of a natural river bed might be composed of: (a) gravels or (b) gravels and large sandy areas (Fig. 1). The various features of gray levels raise difficulties in determining a proper binary threshold \( T \) to covert the image into a binary image (black and white). An inappropriate binary threshold \( T \) would improperly distinguish gravels from sandy areas, as shown in Fig. 1c. An expert could determine an appropriate binary threshold through manual identification based on the features of gray levels in a captured image, which is, however, a long and cumbersome procedure. Artificial intelligence (AI) techniques have the great ability of learning from experiences and/or extracting the key features from data, which indicates AI techniques can be very useful to learn experiences from manual identification with respect to proper binary thresholds. Artificial neural networks (ANNs), a branch of AI, have promising performance in various fields such as hydrological

![Fig. 1](image_url). Captured images and its appropriate/mistaken binary thresholds \( T \) to convert into binary images from different feature of gray level: (a) the image composes mainly grains with an appropriate \( T \), (b) the image contains large sandy areas with an appropriate \( T \), and (c) the image contains large sandy areas with an inappropriate \( T \).
systems (Chen and Chang, 2009; Modarres, 2009; Piotrowski and Napierkowski, 2013) and the scour around hydraulic structures (Azamathulla et al., 2010, 2012; Zakaria et al., 2010). Furthermore, neuro-fuzzy theories mimic the human thinking process for learning similar experiences to make optimal decisions, which is well recognized for their outstanding abilities in system control (Chaves and Kojiri, 2007; Pintong et al., 2009). In this study, we implement a rule-based counterpropagation fuzzy neural network (CFNN) for learning and inferring the relationship between the features of a gray-level image (inputs) and its corresponding binary threshold (output).

The watershed transform algorithm (WST) is one of the most commonly used image segmentation methods in image processing (Butler et al., 2001; Graham et al., 2005a, 2005b; Sime and Ferguson, 2003; Strom et al., 2010). Unfortunately, over-segmentation frequently occurs in the segmentation procedure (Belaid and Mourou, 2009). The h-minima transform was reported to improve the over-segmentation problem of the WST by use of a threshold value (Graham et al., 2005a, 2005b). Chen et al. (2008) investigated the optimal threshold value for the h-minima transform and demonstrated its effectiveness in an AGS method.

In this study, a refined automated grain sizing method (R-AGS) (fusing the CFNN into AGS) is proposed to automatically and precisely estimate grain size distribution through systematically providing two key parameters for each captured digital image, including (1) an adaptive binary threshold for recognizing sandy areas as the background and (2) a threshold for the h-minima transform to control the over-segmentation of the WST.

2. Methodology

To automatically and precisely obtain the grain size distribution, the R-AGS is developed by extracting the information contained within digit images. This method combines (1) a neuro-fuzzy network for determining a suitable binary threshold in each captured image and (2) an automated grain sizing (AGS) procedure mainly comprising the discrete wavelet transform (DWT) and watershed transform algorithm (WST). The counterpropagation fuzzy neural network (CFNN) is trained to produce a binary threshold. The DWT can down-sample the probabilities of occurrence of gray levels and is implemented to reduce the input dimension for the neural-fuzzy network. The WST is applied to segmenting the individual grains in a given binary image. Related methods used in this study are addressed as follows.

2.1. Image pre-processing

In the field, each digital image was captured over a sample patch of a 1 m² square frame laid on the investigated river-bed. In each captured image with resolutions approximately 1000 by 1000 pixels, gravels disposed outside the frame were trimmed during image pre-processing. Meanwhile, spatial distortion is inevitable while taking photographs. To overcome such a problem, a bilinear interpolation presented by Graham et al. (2005b) was employed to map four reference points onto a rectangle in a correction sense. Next, colorful images were converted into gray-level ones because the color information is not helpful for precise object recognition (Graham et al., 2005b), in addition, the colorful images would consume more computational time than gray-level ones (Chang and Chung, 2012).

2.2. Obtaining adaptive binary thresholds: using the CFNN and DWT

The architecture of the counterpropagation fuzzy neural network (CFNN) is shown in Fig. 2. In the input layer, the features of a gray-level image can be analyzed by calculating the probability of occurrence for each gray level \( g_k \) as follows:

\[
P(g_k) = \frac{n_k}{n}, \quad k = 0, 1, 2, \ldots, 255
\]  

where \( n \) is the total number of pixels in the image (in this case \( n = 1000 \times 1000 \)), \( n_k \) is the number of pixels with gray level \( g_k \) and \( P(g_k) \) has 256 dimensions (0–255 gray levels). A high dimensional input signal could easily cause an over-fitting problem, consequently it is necessary to effectively identify the most important input combination and reduce the input dimension.

The discrete wavelet transform (DWT) can decompose a signal into low-frequency components that represent the optimal approximation and high-frequency components that represent the detailed information or noise of the original signal (Mallat, 1989). The DWT is an available tool to "down-sample" the probabilities of occurrence of gray levels and has been widely applied to image and signal processing (Lewis and Knowles, 2002; Partal and Cigizoglu, 2008; Rioul and Vetterli, 2002). In this study, we implement the DWT to reduce the input dimension. The original 256-dimensional input signal derived from a gray-level image is down-sampled into two 128-dimensional signals, where noises are filtered by a high-pass filter, and approximations are retained by a low-pass filter. The down-sampling procedure is repeated five times starting from 256 dimensions, and an 8-dimensional signal produced from a low-pass filter in Round 5 will be determined as an input signal to the CFNN after a model selection procedure (Fig. 2 and Table 1).

The input signal \((X_1, X_2, \ldots, X_8)\) would be transformed into an output, i.e. the binary threshold \(T\). The connections between the input layer and the hidden layer are denoted as "w" which is the "if" statement of the control rule-base, while the connections between the hidden layer and the output layer are denoted as "π" which is the "then" statement of the control rule-base. Thus, the statement of each rule is defined as: "if the input signal \((X_1, X_2, \ldots, X_8)\) is w, then the output binary threshold \(T\) is π". The number of rules is determined by the distance \(A\) between w and the input signal. Generally speaking, the number of rules increases when the value of \(A\) decreases. The pattern matching is conducted by the Gaussian function that calculates the distances between the input signal and rule nodes to assign a degree of membership to each rule. Finally, the output (binary threshold, T) is the weighted average of the output of each rule, and the gray-level image can then be converted into a binary image based on T.

2.3. Image segmentation and minor-axis measurement

An excellent grain-size analysis highly depends upon a successful image segmentation of grains; consequently adopting a suitable segment method is crucial. The watershed transform algorithm (WST) is a classical method for image segmentation. When being used alone, the WST, however, often leads to over-segmentation problems, which could even cause the result useless. Recent solutions to mitigate over-segmentation are: (1) the h-minima transform (Graham et al., 2005a, 2005b; Soille, 2003); (2) marker setting (Chang and Chung, 2012); and (3) region merging (Haris et al., 1998). The h-minima transform is relatively simpler than the other two methods and is implemented in this study, where the given threshold is denoted as H.

After the image segmentation procedure is completed, the Hotelling transform is used to transform the axes of the segmented objects in parallel with the x and y axes of a coordinate plane, and then to measure the maximum lengths of gravels in the directions x and y, separately. The axis with a shorter length is defined as the minor axis of the gravel. These results will be used to plot the cumulative curves in the grain-size analysis.
Because the uniqueness of images or videos in image processing, it requires intensive computation to obtain and/or search the optimal solution for each image or video. Fortunately, MATLAB with relevant image processing tools can provide a friendly environment to researchers of interest. The implementation details of the bilinear interpolation, DWT, region filling, WST and $h$-minima transform can be found in MATrix LABoratory (MATLAB 2008b).

3. Application

3.1. Data collection

Image data of the Lanyang River bed were collected from the First River Management Office, Water Resources Agency, Taipei, Taiwan. The Lanyang River is located in northeastern Taiwan and has a drainage area of approximately 978 square kilometers. The Lanyang River is about 73 km in length with an average channel slope of 1/55, which has created an alluvial fan over the Lanyang Plain. In this study, a total of 130 images were captured in the Lanyang River bed in 2007 (First River Management Office, 2007). A frame of one square meter was placed over the river bed to ensure photographs were taken consistently and vertically by using an inexpensive camera (DX7590). Twenty-three river sections, where their elevations decrease approximately from 400 to 200 meter, were investigated, and approximately four to eight images of the river bed materials were captured at each section. The means of the 50 percentile of grain sizes and elevations for the 23 sections are shown in Fig. 3. It appears the 50 percentile of grain sizes by using manual image analysis slightly decreases from upstream to midstream (from Section 1 to Section 23). The collected images can be typically classified into two types:

(1) Type A (without obvious sandy areas): the image is almost entirely composed of grains larger than 16 mm whereas sandy areas are too small to be recognized as the background.

(2) Type B (with obvious sandy areas): the image is composed of grains including obvious sandy areas, where sandy areas need to be recognized as the background.

The numbers of images in Types A and B are 103 and 27, respectively.

3.2. Manual image analysis

The minimum resolvable grain size of a digital image is highly related to image resolution (Graham et al., 2005a). In this study, the image resolution is 1000 by 1000 pixels and the image scale is 1 mm per pixel; consequently the minimum resolvable grain size is approximately 16 mm. We aim to estimate the grain size distribution of digital images subject to the condition that grains with sizes less than 16 mm and/or sandy areas are automatically recognized as the background. The length of the minor axis of each grain in an image was measured manually in this study, referred from Buscombe (2008). The minor axis of each grain in each image was drawn with a red line by MS paint, and all the red lines in an image were extracted by using the corresponding color threshold (Fig. 4a and b). It would take approximately an hour for recognizing and measuring the minor axes of grains in each image. The red lines (minor axes) in each image were numerically counted to obtain the cumulative grain size distribution curve by using a conversion equation ($\psi = \log_2$ (the minor axis; unit: mm)). The grain size of 16 mm can be converted into that of 4 psi.

For configuring the CFNN and assessing its performance, the images captured at the same river sections were selected and allocated into training, validation and testing subsets, respectively. Therefore, the 130 images were divided into training, validation and testing subsets, where the numbers of images in these three subsets are shown in Table 1. It appears the 50 percentile of grain sizes by using manual image analysis slightly decreases from upstream to midstream (from Section 1 to Section 23). The collected images can be typically classified into two types:

(1) Type A (without obvious sandy areas): the image is almost entirely composed of grains larger than 16 mm whereas sandy areas are too small to be recognized as the background.

(2) Type B (with obvious sandy areas): the image is composed of grains including obvious sandy areas, where sandy areas need to be recognized as the background.

The numbers of images in Types A and B are 103 and 27, respectively.

### Table 1
Performances of the CFNN with different input combinations.

<table>
<thead>
<tr>
<th>Dimension of input combination</th>
<th>RMSE (T: gray level)</th>
<th>Number of Fuzzy rules</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Training</td>
<td>Validation</td>
</tr>
<tr>
<td>64</td>
<td>33</td>
<td>36</td>
</tr>
<tr>
<td>32</td>
<td>31</td>
<td>31</td>
</tr>
<tr>
<td>16</td>
<td>27</td>
<td>28</td>
</tr>
<tr>
<td>8</td>
<td>14</td>
<td>15</td>
</tr>
<tr>
<td>4</td>
<td>38</td>
<td>39</td>
</tr>
<tr>
<td>2</td>
<td>36</td>
<td>39</td>
</tr>
</tbody>
</table>

The minimum resolvable grain size of a digital image is highly related to image resolution (Graham et al., 2005a). In this study, the image resolution is 1000 by 1000 pixels and the image scale is 1 mm per pixel; consequently the minimum resolvable grain size is approximately 16 mm. We aim to estimate the grain size distribution of digital images subject to the condition that grains with sizes less than 16 mm and/or sandy areas are automatically recognized as the background. The length of the minor axis of each grain in an image was measured manually in this study, referred from Buscombe (2008). The minor axis of each grain in each image was drawn with a red line by MS paint, and all the red lines in an image were extracted by using the corresponding color threshold (Fig. 4a and b). It would take approximately an hour for recognizing and measuring the minor axes of grains in each image. The red lines (minor axes) in each image were numerically counted to obtain the cumulative grain size distribution curve by using a conversion equation ($\psi = \log_2$ (the minor axis; unit: mm)). The grain size of 16 mm can be converted into that of 4 psi.

For configuring the CFNN and assessing its performance, the images captured at the same river sections were selected and allocated into training, validation and testing subsets, respectively. Therefore, the 130 images were divided into training, validation and testing subsets, where the numbers of images in these three subsets are shown in Table 1.
subsets are 70, 30 and 30 accordingly. The results of manual image analysis in these subsets are shown in Fig. 4c–e, which reveals a wide range of grain size conditions. The results indicate three subsets are suitably divided owing to the similar grain size distribution of these three subsets.

3.3. R-AGS construction

Fig. 5 provides the explicit flowchart of developing the R-AGS, which can be organized into 4 parts (main process and three stages). Main process: the operating procedure of the AGS. Stage
1: obtain a target binary threshold $T^*$ (for the DWT and CFNN) and an $H^*$ (for the $h$-minima transform) associated with the maximum AR value. Stage 2: construct an optimal model (DWT + CFNN) for estimating a binary threshold $T$ for making a binary image; and assign a suitable $H$ for controlling the over-segmentation of the WST. Stage 3: determine a binary threshold $T$ of the input image for the AGS by the constructed model (DWT + CFNN) to estimate grain size distribution of the image.

### 3.3.1. AGS operation

The operating procedure of the AGS can be organized into six steps and is performed sequentially on each image (Fig. 5). Step 1: a color image is converted into a gray-level one. Step 2: the image background is subtracted from the gray-level image by using an adaptive binary threshold ($T$), which is determined by the CFNN, for properly differentiating sandy areas from grains. Step 3: filling the interior holes of each grain in the binary image by using the region-filling technique. Step 4: segmenting individual grains in the binary image by using the WST, which requires a threshold ($H$) for the $h$-minima transform to control over-segmentation. Step 5: measuring the minor axis of each grain by using the Hotelling transform. Step 6: plotting number-cumulative curves based on classified grain sizes.

### 3.3.2. Searching suitable parameters

The target binary threshold $T^*$ and the suitable threshold $H^*$ for the $h$-minima transform in each image can be obtained by calculating the accurate rate (AR) shown as follows:

$$\text{AR} \ (%) = \left[ 1 - \frac{\sum_{i=1}^{n} (t_{ob}(i) - t_{es}(i))}{\sum_{i=1}^{n} t_{ob}(i)} \right] \times 100$$

where $t_{ob}$, $t_{es}$, and $n$ are the observation, estimation and the total number of grain size levels (from 4 to 9 psi with an increment of 0.5 psi). A higher AR value indicates a better precision of the estimation results between observed values (manual image analysis) and estimated values (the AGS procedure).

The AGS estimates the grain sizes by trying different $T$ and $H$ pairs for each image. $T$ is searched within the range of [50, 250] with an interval of 5, while the $H$ is searched within the range of [1, 15] with an interval of 1. Therefore, a total of 615 (=41×15) AR values are calculated for each image, and thus $T^*$ and $H^*$ are therefore obtained in an association with the maximum AR value for each image.

### 3.3.3. Determining the optimal CFNN and assigned $H$ for the $h$-minima transform

To construct a suitable CFNN, we have tried different input combinations. An appropriate input combination will allow the network to not only adequately map onto the desired output (binary threshold $T$) but also avoid a loss of important information. The probability of occurrence of gray levels has 256 dimensions in each

![Fig. 5. Flowchart of the R-AGS. Right-handed figures are the results related to Steps (1)–(6) of the AGS.](image)

![Fig. 6. Averaged AR values simulated in training and validation subsets.](image)
captured image, which easily causes over-fitting problems (due to large input dimension) when constructing the estimation model. Six input combinations (64, 32, 16, 8, 4 and 2 dimensions) down-sampled by using the DWT are used as the CFNN inputs for determining the optimal network structures (i.e. six network structures). For each input combination, a number of networks with different numbers of rules (hidden nodes) are calibrated based on training subsets, then the optimal network structure is determined based on validation subsets. The optimal CFNN structure that contains the most suitable input combination, rule number and output is determined among those six CFNNs with respect to the smallest RMSE. In addition, a total of 61500 (=41 \( T \times 15 \times H \times 100 \) images in training and validation subsets) AR values are calculated in Stage 1, the averaging AR values of those 100 images in corresponding to the specific values of \( T \) and \( H \) are shown in Fig. 6. A suitable \( H \) for the \( h \)-minima transform to control the over-segmentation of the WST is assigned based on Stage 1 result (i.e., the midpoint of the region composed of high AR values).

3.4. Comparative model and performance

The methods presented by (Chang and Chung, 2012) and (Graham et al., 2005a, 2005b), denoted as M1 and M2 models accordingly, are conducted for comparison purpose. M1 introduces multilevel-thresholding and is recognized as a suitable tool for improving the over-segmentation of the WST; however, M1 focuses on the grain distribution for grain sizes ranging within 25.4–215 mm without considering sandy areas. M2 is a pivotal method for automatically measuring grain sizes (16–256 mm) of river-bed materials from digital images. Both methods require setting more than five parameters. The Hotelling transform with the same comparative standard is used to measure the minor axes in the R-AGS, M1 and M2 models. The estimation information of digital images from river-bed materials can be used to plot cumulative grain size distribution curves. This study adopts seven most commonly used grain-size percentiles in sediment topological studies, which are 5, 16, 25, 50, 75, 84 and 95. Errors of percentiles are calculated in terms of root mean square error (RMSE) as Eq. (3), and the AR is also calculated to assess the performance of the estimation results associated with the R-AGS, M1 and M2 models.

\[
RMSE = \sqrt{\frac{\sum_{i=1}^{n} (t_{oi}(i) - t_{re}(i))^2}{n}}
\]

Table 2

<table>
<thead>
<tr>
<th>Grain size (psi)</th>
<th>O(^a)</th>
<th>R-AGS</th>
<th>M1</th>
<th>M2</th>
<th>O(^a)</th>
<th>R-AGS</th>
<th>M1</th>
<th>M2</th>
<th>O(^a)</th>
<th>R-AGS</th>
<th>M1</th>
<th>M2</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>4573</td>
<td>5233</td>
<td>5423</td>
<td>5749</td>
<td>1417</td>
<td>1706</td>
<td>1839</td>
<td>1997</td>
<td>1850</td>
<td>1928</td>
<td>2267</td>
<td>2457</td>
</tr>
<tr>
<td>4.5</td>
<td>4051</td>
<td>4227</td>
<td>5740</td>
<td>4250</td>
<td>1024</td>
<td>897</td>
<td>1645</td>
<td>1045</td>
<td>1112</td>
<td>1042</td>
<td>1636</td>
<td>1254</td>
</tr>
<tr>
<td>5</td>
<td>2586</td>
<td>2602</td>
<td>4497</td>
<td>2856</td>
<td>615</td>
<td>527</td>
<td>960</td>
<td>685</td>
<td>608</td>
<td>655</td>
<td>952</td>
<td>764</td>
</tr>
<tr>
<td>5.5</td>
<td>1431</td>
<td>1435</td>
<td>2580</td>
<td>1755</td>
<td>319</td>
<td>306</td>
<td>427</td>
<td>400</td>
<td>299</td>
<td>347</td>
<td>409</td>
<td>427</td>
</tr>
<tr>
<td>6</td>
<td>680</td>
<td>718</td>
<td>1149</td>
<td>935</td>
<td>4877</td>
<td>4731</td>
<td>6850</td>
<td>5660</td>
<td>5532</td>
<td>5565</td>
<td>7461</td>
<td>6649</td>
</tr>
<tr>
<td>Total grains</td>
<td>13321</td>
<td>14215</td>
<td>19389</td>
<td>15545</td>
<td>173</td>
<td>439</td>
<td>264</td>
<td>–</td>
<td>64</td>
<td>416</td>
<td>295</td>
<td>–</td>
</tr>
</tbody>
</table>

| AR              | –      | 93%  | 54% | 83% | –      | 85%  | 59% | 84% | –      | 94%  | 65% | 80% |
| RMSE            | –      | 306  | 1324| 577 | –      | –    | –   | –   | –      | 64   | 416 | 295 |

AR = \( \frac{\sum_{i=1}^{n} (t_{oi}(i) - t_{re}(i))^2}{n} \) – \( \frac{\sum_{i=1}^{n} (t_{oi}(i))^2}{n} \)

\( n \) is the number of images in validation subset.

<table>
<thead>
<tr>
<th>Grain size (mm)</th>
<th>O(^a)</th>
<th>R-AGS</th>
<th>M1</th>
<th>M2</th>
<th>O(^a)</th>
<th>R-AGS</th>
<th>M1</th>
<th>M2</th>
<th>O(^a)</th>
<th>R-AGS</th>
<th>M1</th>
<th>M2</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>4573</td>
<td>5233</td>
<td>5423</td>
<td>5749</td>
<td>1417</td>
<td>1706</td>
<td>1839</td>
<td>1997</td>
<td>1850</td>
<td>1928</td>
<td>2267</td>
<td>2457</td>
</tr>
<tr>
<td>4.5</td>
<td>4051</td>
<td>4227</td>
<td>5740</td>
<td>4250</td>
<td>1024</td>
<td>897</td>
<td>1645</td>
<td>1045</td>
<td>1112</td>
<td>1042</td>
<td>1636</td>
<td>1254</td>
</tr>
<tr>
<td>5</td>
<td>2586</td>
<td>2602</td>
<td>4497</td>
<td>2856</td>
<td>615</td>
<td>527</td>
<td>960</td>
<td>685</td>
<td>608</td>
<td>655</td>
<td>952</td>
<td>764</td>
</tr>
<tr>
<td>5.5</td>
<td>1431</td>
<td>1435</td>
<td>2580</td>
<td>1755</td>
<td>319</td>
<td>306</td>
<td>427</td>
<td>400</td>
<td>299</td>
<td>347</td>
<td>409</td>
<td>427</td>
</tr>
<tr>
<td>6</td>
<td>680</td>
<td>718</td>
<td>1149</td>
<td>935</td>
<td>4877</td>
<td>4731</td>
<td>6850</td>
<td>5660</td>
<td>5532</td>
<td>5565</td>
<td>7461</td>
<td>6649</td>
</tr>
<tr>
<td>Total grains</td>
<td>13321</td>
<td>14215</td>
<td>19389</td>
<td>15545</td>
<td>173</td>
<td>439</td>
<td>264</td>
<td>–</td>
<td>64</td>
<td>416</td>
<td>295</td>
<td>–</td>
</tr>
</tbody>
</table>

| AR              | –      | 93%  | 54% | 83% | –      | 85%  | 59% | 84% | –      | 94%  | 65% | 80% |
| RMSE            | –      | 306  | 1324| 577 | –      | –    | –   | –   | –      | 64   | 416 | 295 |

AR = \( \frac{\sum_{i=1}^{n} (t_{oi}(i) - t_{re}(i))^2}{n} \) – \( \frac{\sum_{i=1}^{n} (t_{oi}(i))^2}{n} \)

\( n \) is the number of images in validation subset.

Table 3

<table>
<thead>
<tr>
<th>Subset</th>
<th>Training</th>
<th>Validation</th>
<th>Testing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>R-AGS</td>
<td>M1</td>
<td>M2</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.14</td>
<td>0.21</td>
<td>0.22</td>
</tr>
<tr>
<td>RMSE</td>
<td>4.48</td>
<td>6.95</td>
<td>8.53</td>
</tr>
<tr>
<td>Improvement rate (%)</td>
<td>–</td>
<td>36</td>
<td>47</td>
</tr>
</tbody>
</table>

The numbers of images in training, validation and testing are 70, 30 and 30, respectively.

\( ^a \) Manual image analysis.

The methods presented by (Chang and Chung, 2012) and (Graham et al., 2005a, 2005b), denoted as M1 and M2 models accordingly, are conducted for comparison purpose. M1 introduces multilevel-thresholding and is recognized as a suitable tool for improving the over-segmentation of the WST; however, M1 focuses on the grain distribution for grain sizes ranging within 25.4–215 mm without considering sandy areas. M2 is a pivotal method for automatically measuring grain sizes (16–256 mm) of river-bed materials from digital images. Both methods require setting more than five parameters. The Hotelling transform with the same comparative standard is used to measure the minor axes in the R-AGS, M1 and M2 models. The estimation information of digital images from river-bed materials can be used to plot cumulative grain size distribution curves. This study adopts seven most commonly used grain-size percentiles in sediment topological studies, which are 5, 16, 25, 50, 75, 84 and 95. Errors of percentiles are calculated in terms of root mean square error (RMSE) as Eq. (3), and the AR is also calculated to assess the performance of the estimation results associated with the R-AGS, M1 and M2 models.

\[
RMSE = \sqrt{\frac{\sum_{i=1}^{n} (t_{oi}(i) - t_{re}(i))^2}{n}}
\]
where \( t_{ob}, t_{es}, \) and \( n \) are the observation, estimation and the total number of data accordingly.

4. Results and discussion

This study proposes a novel approach (R-AGS) for estimating grain sizes ranging from 4 psi to 9 psi (16–512 mm) in digital images. A total of 615 pairs of \( T \) and \( H \) values are assigned to perform the AGS process to calculate AR values for each image. Fig. 6 shows the averaged AR results of 100 images (70 images from the training subset and 30 images from the validation subset), in which 615 pairs of \( T \) and \( H \) values are derived for each image. Higher AR values occur as the binary threshold \( T \) falls within 160–190 gray levels and the threshold \( H \) falls within 1–5. Moreover, we find the \( H \) is less sensitive to the AR as it is within [1,5], therefore the suitable value assigned to the \( H \) can be determined as 3. Table 1 presents the most suitable input dimension for the CFNN network. The optimal input dimension is “8” and the correspondent number of fuzzy rules is “7” owing to its smallest RMSE values in both training and validation subsets. Fig. 7 shows the comparison results between the target binary threshold \( T \) (identified from 615 AR values derived from the AGS for each digital image) and the CFNN output \( T \) in the training subset (70 images) and validation subset (30 images), respectively. It shows that the target binary thresholds are well-fitted by CFNN outputs in both training (CC = 0.92) and validation (CC = 0.88) subsets. The precise estimation results indicate that the constructed CFNN model can be used to automatically provide a suitable binary threshold \( T \) for a given digital image. We then use the constructed CFNN to estimate the \( T \) value for each image in the testing subset.

After \( T \) and \( H \) are determined by the CFNN and AR accordingly, the R-AGS is ready to be used to estimate the grain size distribution.

![Fig. 8. Performance of the R-AGS, M1 and M2 in two types of images.](image-url)
for a given image. Table 2 illustrates the numbers of grains in sizes ranging between 4–9 psi with an interval of 0.5 psi by using manual image analysis, R-AGS, M1 and M2, respectively. The total numbers of gravels (from 4 to 6 psi) observed (by manual image analysis) in the training, validation and testing subsets are 13321, 4877 and 5532, respectively. In contrast, the total numbers of grains observed (by manual image analysis) in cobble or boulder size (from 6.5 to 9 psi) in the training, validation and testing subsets are only 383, 220 and 181, respectively. Because the total number of grains of sizes from 4 to 6 psi is much larger than that of sizes from 6.5 to 9 psi (23730 versus 784), AR values for gravels of these two categories are computed separately for assessing model performance. The results indicate the R-AGS and M2 have similar performance for grain sizes less than 6 psi (64 mm), but M2 produces biases for grain sizes beyond 6.5 psi (90.5 mm). M1 overestimates grain numbers in all cases but provides more stable performance than M2 in cobble or boulder size. The error sources of M1 and M2 would be the misrecognition of sandy areas as large-sized grains. In addition, Table 3 also demonstrates the R-AGS is superior to M1 and M2 with respect to RMSE in both psi and mm senses. The improvement rates of the R-AGS over M1 and M2 in an mm sense are all larger than 20% in the testing subsets. In sum, the R-AGS can estimate the grain size distribution much more precisely than M1 and M2.

Fig. 8 presents the comparison results of the proposed R-AGS and comparative models M1 and M2 with respect to Types A and B images. The R-AGS, M1 and M2 produce similar performance for the Type A image (without obvious sandy areas), which demonstrates that these three models perform equally well in this case. For the Type B image, neither M1 nor M2 can properly recognize sandy areas as the background, where sandy areas are regarded as larger grains.

Fig. 9 presents the cumulative curves by summarizing the results in the training and testing subsets, respectively. Again, it
clearly shows that the R-AGS outperforms M1 and M2 for grain sizes within 4 to 6 psi (in gravel size) and also gains significantly improvements in estimation accuracy for grain sizes within 6.5–9 psi (in cobbles or boulder size). Fig. 10 shows the scatter plot at seven specified percentiles (from 5% to 95%) of the grain size distribution by manual image analysis and the R-AGS in the testing subset (30 images). It appears that all the 30 measured and estimated results at each of the seven specified percentiles match quite well (close to the identity line in the scatter plot) and maintains a high correlation coefficient ($R^2 = 0.95$).

In general, measuring the grain size distribution is labor-intensive and time-consuming. The proposed R-AGS method would require only a few minutes for suitably estimating the grain size distribution of the digital image. Consequently, the automated grain-size measurements could be achieved more efficiently and practically.

5. Conclusions

Quantification of the grain size distribution of fluvial gravels is a critical task in the studies of river channel behavior in hydraulics, geomorphology, and ecology. Recent AGS methods have shown advantages such as time reduction, non-intrusive sampling and the avoidance of operator bias. This study moves one step toward the practical application of automated grain-size measurements. The R-AGS is introduced to effectively distinguish grains from sands and automatically estimate the grain size distribution. To ensure the proposed R-AGS is robust as being applied to complicated natural river-bed images, the constructed CFNN adequately provides the binary threshold ($T$) for each image and the suitable threshold ($H$) assigned for the $h$-minima transform can effectively control the over-segmentation of the WST. The proposed R-AGS outperforms two comparative methods in the estimation of grain size distribution. Moreover, it is easy to re-construct the CFNN model by updating and/or adding rule nodes when investigating image samples that are significantly different from those in this study; consequently its applicability and practicability could be expanded. In summary, the proposed R-AGS not only automatically and adaptively provides grain size distributions for variously captured river-bed images but has the potential to replace complex and time-consuming manual sampling methods.

Acknowledgements

The authors sincerely appreciate the First River Management Office, WRA, Taiwan, for provide valuable river-bed images and Professor Chiou-Shann Fuh in Department of Computer Science and Information Engineering, National Taiwan University, for giving valuable suggestions. This project was partially sponsored by the National Science Council, Taiwan, ROC (Grant No.: 100-2313-B-002-011-MY3).

References
