

Universität Stuttgart



Lehrstuhl für Wasserbau

und Wassermengenwirtschaft

Prof. Dr.-Ing. Silke Wieprecht

# Surficial Sediment Size Distribution Through The Use of Digital Imaging and Autocorrelation Analysis Master Thesis

**Sptember 2008**

---

Institut für Wasserbau

Pfaffenwaldring 61 \* D-70550 Stuttgart (Vaihingen)

Tel.: +49 (0)711 685 64752 \* Fax: +49 (0)711 685 64746

E-mail: <vorname.name>@iws.uni-stuttgart.de

<http://www.iws.uni-stuttgart.de>

## Abstract

Grain size analysis is a process by which the proportion of material of each grain size present in a given soil sample is determined. The objectives of a grain size analysis are to accurately measure individual particle sizes to determine their frequency distribution, and calculate a statistical description that adequately characterize the sample.

Information on grain size distribution has many applications in the field of fluvial engineering, such as sediment transport and deposition, and fluvial habitat modeling.

Traditional grain size analysis techniques which have been in use are time consuming, effort intensive and expensive. Not only the grain size analysis methods, but sampling techniques in use today also have difficulties due to lateral, longitudinal and depth variability of sediment size and distribution. Most of the techniques are complicated and require professional experience.

This study aimed at finding a sampling and analysis technique that achieves a satisfactory characterization of grain size while simultaneously reducing the time spent in both the field and the laboratory. During this study a new method of size distribution of river bed materials through digital imaging and autocorrelation analysis has been used. 75 % of the samples analyzed showed strong correlation to the sieve analysis results. Overall average error of around 10 % was observed. Moreover, the method proved to be cheaper and faster than the traditional methods by more than 100 %.

Keywords: grain-size distribution, mechanical sieving, digital image analysis, autocorrelation analysis

---

---

# Contents

<b>1</b>	<b>Introduction</b>	<b>10</b>
1.1	General . . . . .	10
1.2	Objectives . . . . .	11
1.3	Thesis Organization . . . . .	11
<b>2</b>	<b>Traditional Grain-Size Analysis and Sampling Techniques</b>	<b>14</b>
2.1	Mechanical Sieving . . . . .	14
2.2	Settling Tubes . . . . .	15
2.3	Visual Grain-size Estimates . . . . .	16
<b>3</b>	<b>Size Distribution from Digital Image Analysis - Previous Research</b>	<b>19</b>
3.1	Sime and Ferguson (2003) . . . . .	20
3.2	Graham, Stephan and Reid (2005) . . . . .	22
<b>4</b>	<b>Size Distribution from Digital Images Using Autocorrelation Analysis</b>	<b>26</b>
4.1	Spatial Autocorrelation Basics . . . . .	26
4.2	Spatial Autocorrelation Calculation - First Algorithm . . . . .	27
4.3	Grain-Size Distribution Calculation - Second Algorithm . . . . .	29
<b>5</b>	<b>Experimental Procedures</b>	<b>34</b>
5.1	Experimental Equipment . . . . .	34
5.2	Experimental Procedures . . . . .	37
5.2.1	First Stage . . . . .	38
5.2.2	Second Stage . . . . .	38

---

---

5.2.3	Third Stage . . . . .	40
<b>6</b>	<b>Results</b>	<b>46</b>
6.1	First Stage . . . . .	46
6.1.1	100 dpi Resolution . . . . .	46
6.1.2	300 dpi Resolution . . . . .	48
6.1.3	600 dpi Resolution . . . . .	48
6.2	Second Stage . . . . .	49
6.3	Third Stage . . . . .	50
<b>7</b>	<b>Analysis and Discussion</b>	<b>60</b>
7.1	Effect of Image Resolution on Calculated Grain-Size Distribution . . . . .	60
7.2	Effect of Lighting Conditions and Camera Height on Calculated Grain-Size Distribution	60
7.3	Effect of Calibration Method on Calculated Grain-Size Distribution . . . . .	62
7.4	Correlation Strength of Image and Sieve Analysis Results . . . . .	63
7.5	Sources of Error . . . . .	64
7.5.1	Errors Associated with Projected Area and Volume / Weight Comparison .	64
7.5.2	Errors Associated with Sampling . . . . .	65
7.5.3	Errors Associated with Calibration and Validation Inconsistency . . . . .	65
7.6	Image Collection . . . . .	65
<b>8</b>	<b>Conclusions and Recommendations</b>	<b>69</b>
8.0.1	Conclusions . . . . .	69
8.0.2	Recommendations and Future Work . . . . .	70
<b>A</b>	<b>Appendix A</b>	<b>71</b>
<b>B</b>	<b>Appendix B</b>	<b>72</b>
	<b>Bibliography</b>	<b>73</b>

---

## List of Figures

Figure 1.1	Substrate map of a river reach . . . . .	12
Figure 2.1	Mechanical sieve apparatus (Retsch 2008) . . . . .	15
Figure 2.2	Settling Tube (Poppe 2001) . . . . .	17
Figure 3.1	Image of a gravel bed surface (Sime 2003) . . . . .	21
Figure 3.2	Edge maps generated from fig. 3.1 using different edge detection methods (Sime 2003) . . . . .	21
Figure 3.3	The stages followed by Graham, Stephan and Reid . . . . .	23
Figure 3.4	Photographic procedure followed by Graham et al. (Graham 2005) . . . . .	24
Figure 3.5	Image processing procedure (a) a color image, (b) the same image after image processing (Graham 2005) . . . . .	25
Figure 4.1	Autocorrelation curves of different grain-size fractions. Pixels in larger grains are more similar for a longer distance than pixels in smaller grains (Rubin 2004) . . . . .	28
Figure 4.2	Diagram illustrating how the first algorithm works: offset plaquettes moving at single offset distances and each time being compared to the reference plaquette . . . . .	30
Figure 4.3	A schematic view of the stages of grain-size distribution determination of a soil sample from digital images using autocorrelation analysis . . . . .	33
Figure 5.1	Wooden frame used for holding soil samples during scanning . . . . .	35
Figure 5.2	Set of sieves mounted on a shaker (left), Drying oven (right) . . . . .	36
Figure 5.3	Digital bench scale (left), Underwater digital camera (right) . . . . .	36
Figure 5.4	Metal box sampler . . . . .	37

---

Figure 5.5	Images from first stage scans, (1) class-1 (<0.315 mm), (2) class-2 (0.315 mm - 2.5 mm), (3) class-3 (2.5 mm - 10 mm), (4) class-4 (>10 mm), (5) class-5 (<0.315 mm -2.5 mm), (6) class-6 (2.5 mm - >10 mm) . . . . .	39
Figure 5.6	Images from second stage scans, (1) class-1 (<0.315 mm - 1 mm), (2) class-2 (1.0 mm - 2.5 mm), (3) class-3 (2.5 mm - 6.3 mm), (4) class-4 (6.3 mm - 12.5 mm), (5) class-5 (12.5 mm - 20 mm), (6) class-6 (>20 mm) . . . . .	41
Figure 5.7	underwater river bed images, (1) field image 1 (downstream of Schneitbach creek), (2) field image 2 (downstream of Schneitbach creek), (3) field image 3 (upstream of Scheitbach river), (4) field image 4 (upstream of Schneitbach creek)	42
Figure 5.8	Images from third stage taken at 30FI, (1) class-1 (<0.315 mm - 1 mm), (2) class-2 (1.0 mm - 2.5 mm), (3) class-3 (2.5 mm - 6.3 mm), (4) class-4 (6.3 mm - 12.5 mm), (5) class-5 (12.5 mm - 20 mm), (6) class-6 (>20 mm), (7) class-7, (8) class-8, (9) class-9 . . . . .	43
Figure 6.1	Comparison of actual and computed GSD for grains <0.315 mm (class-1), images taken at a resolution of 100 dpi. The algorithm was unable to detect the grains of the sample due to reasons explained in section 6.1.1. . . . .	47
Figure 6.2	Comparison of actual and computed GSD for grains between 0.315 - 2.5 mm (class-2), images taken at a resolution of 100 dpi. The algorithm was unable to detect the grains of the sample due to reasons explained in section 6.1.1. . . . .	48
Figure 6.3	Comparison of actual and computed GSD for grains >10 mm (class-4), images taken at a resolution of 100 dpi . . . . .	49
Figure 6.4	Comparison of actual and computed GSD for grains <0.315 mm (class-1), images taken at a resolution of 300 dpi . . . . .	50
Figure 6.5	Comparison of actual and computed GSD for grains between 0.315 - 2.5 mm (class-2), images taken at a resolution of 300 dpi . . . . .	51
Figure 6.6	Comparison of actual and computed GSD for grains <0.315 - 2.5 mm (class-5), images taken at a resolution of 300 dpi . . . . .	52

---

---

Figure 6.7	Comparison of actual and computed GSD for grains between 0.315 - 2.5 mm (class-2), images taken at a resolution of 600 dpi . . . . .	53
Figure 6.8	Comparison of actual and computed GSD for grains between 2.5 - >10 mm (class-6), images taken at a resolution of 600 dpi . . . . .	53
Figure 6.9	Comparison of actual and computed GSD for field image 1 in stage two (scanned calibration images). Very large error and very small correlation between the actual and computed results observed due to reasons explained in section 6.2 .	54
Figure 6.10	Comparison of actual and computed GSD for field image 2 in stage two (scanned calibration images). Very large error and very small correlation between the actual and computed results observed due to reasons explained in section 6.2 .	54
Figure 6.11	Comparison of actual and computed GSD for field image 3 in stage two (scanned calibration images). Very large error and very small correlation between the actual and computed results observed due to reasons explained in section 6.2 .	55
Figure 6.12	Comparison of actual and computed GSD for field image 4 in stage two (scanned calibration images). Very large error and very small correlation between the actual and computed results observed due to reasons explained in section 6.2 .	55
Figure 6.13	Comparison of actual and computed GSD for field image 1 using 40NFI calibration images . . . . .	56
Figure 6.14	Comparison of actual and computed GSD for field image 2 using 30NFI calibration images . . . . .	56
Figure 6.15	Comparison of actual and computed GSD for field image 3 using 30NFI calibration images. Very large error and very small correlation between the actual and computed results observed due to shadowed (underexposed) lighting conditions.	57
Figure 6.16	Comparison of actual and computed GSD for field image 4 using 30NFI calibration images. Very large error and very small correlation between the actual and computed results observed due to shadowed (underexposed) lighting conditions.	57
Figure 6.17	Comparison of actual and computed GSD for sample class-7 using 30NFI calibration images . . . . .	58

---

---

Figure 6.18 Comparison of actual and computed GSD for sample class-8 using 30FI calibration images . . . . .	58
Figure 6.19 Comparison of actual and computed GSD for sample class-9 using 30NFI calibration images . . . . .	59
Figure 7.1 Comparison of scanned soil samples using errors for different image resolutions	61
Figure 7.2 Comparison of field soil samples using errors for different correlation image resolutions . . . . .	61
Figure 7.3 Comparison of field soil samples using errors for different lighting conditions and camera heights . . . . .	62
Figure 7.4 Comparison of field soil samples using errors for different Calibration Methods .	63
Figure 7.5 Overall correlation strength of image and sieve analysis results . . . . .	64
Figure 7.6 A $2x$ by $2y$ ground area captured by placing the camera at a height of $h$ mm from the ground . . . . .	67
Figure 7.7 Side views of the captured ground area . . . . .	67
Figure 7.8 Graph showing the relation between camera height and smallest grain-size of interest for different camera resolutions . . . . .	68

---

---

## List of Tables

Table 4.1	Autocorrelation values of calibration fractions . . . . .	32
Table 5.1	Summary of grain-size classes, image resolutions, and lighting conditions used for calibration during the three stages of the study . . . . .	44
Table 5.2	Summary of soil samples, image resolutions, and lighting conditions used for validation during the three stages of the study . . . . .	45
Table 6.1	Summary of total experiments performed for this study . . . . .	46

---

# 1. Introduction

## 1.1. General

Grain-size analysis is a process by which the proportion of material of each grain-size present in a given soil sample is determined. The objectives of a grain-size analysis are to accurately measure individual particle sizes to determine their frequency distribution, and to calculate a statistical description that adequately characterizes the sample (USGS 2001).

Grain-size is a fundamental property of materials from which much information can be obtained with regards to their behavior (Syvitski 1991; Berg 1939). In particular the dynamical mechanisms of transport and deposition of the constituent grains of sediments is usually drawn from their size distribution (McLaren 1981). Bridge (1981) explains how grain-size distribution of river bed sediments can be interpreted hydraulically and used to describe sediment transport and deposition. In his report, he underlines that important information about the transport and deposition processes can be extracted from statistical grain-size distribution parameters such as mean, median, kurtosis and skewness values.

Information on grain-size distribution of river beds has also another important application in the field of fluvial habitat modeling. Habitat modelers use parameters such as water depth, flow conditions, and substrate maps to produce habitat maps of river reaches (Eberstaller 2006). While water depth and river flow conditions can be accurately calculated using hydraulic models, river substrate maps, as shown in fig. 1.1, are produced based on grain-size distribution of river bed materials.

Most grain-size analysis is carried out in the laboratory. Approximate estimation using a grain-size comparator is commonly undertaken for sand sizes in the field, but any instrumental work is laboratory based. The exception is measurement of the size of gravels. For coarse stream gravels

---

it is usually impracticable to transport sufficiently large samples back to the lab. Therefore, large diameter steel plate sieves are used in the field with a simple weighing balance (Syvitski 1991).

Traditional techniques of particle size analysis namely: particle counting, sieving, and settling tubes also are still in use though often using automated instruments. New ways of grain-size distribution determination such as by digital image analysis have also been developed (Syvitski 1991). These methods are simple and rapid and avoid extensive field work and the need for a large sample to be transported to the laboratory for analysis (Carbonneau 2004).

This paper briefly mentions the traditional methods of grain-size analysis and sampling techniques which have been in use, discusses previous advances in grain-size distribution through digital imaging, and explains in detail, using practical field experiments, a method which uses a simple auto-correlation algorithm to generate grain-size distribution of samples from digital images.

## 1.2. Objectives

The objectives of this thesis are:

- to review previously developed methods of grain-size distribution from digital images, and to identify their applications in the field of fluvial ecological modeling.
- to identify a simple, rapid and accurate method which can be used across a diverse range of sediment types and be easily applicable in preparing fluvial substrate maps for CASiMiR software, a fluvial habitat modeling software developed at The Department of Hydraulic Engineering of The University of Stuttgart.
- to compare size-distribution results of digital image analysis and mechanical sieving and identify the strengths, weaknesses and level of accuracy of the image analysis method.

## 1.3. Thesis Organization

This thesis is organized in 8 chapters. Traditional size-distribution and sampling techniques are reviewed in Chapter 2. Chapter 3 presents a general overview of previously developed methods

---

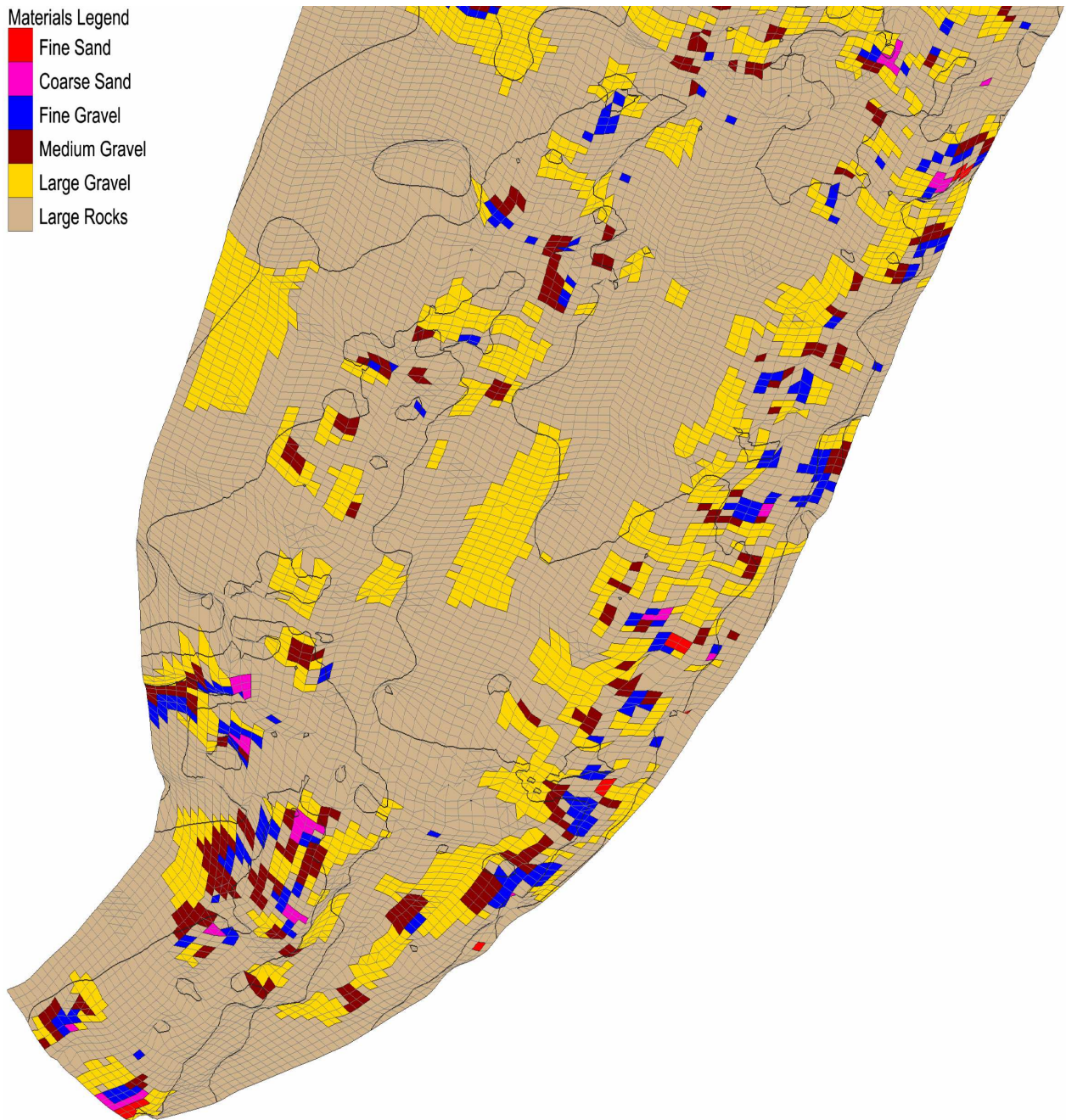


Figure 1.1: Substrate map of a river reach

of size-distribution determination from digital images. The method employed in this thesis - size-distribution using digital imaging and autocorrelation analysis - is thoroughly explained in chapter 4. Chapters 5, 6 and 7 present the experimental procedures followed during this study and their results. Conclusions, recommendations and future research works that need to be carried out to further develop this method are also explained in the last chapter of this thesis.

---

## 2. Traditional Grain-Size Analysis and Sampling Techniques

Traditional grain-size analysis techniques which have been in use in the scientific world are time consuming, effort intensive, and expensive especially when used for large scale tasks like preparing morphological maps of river beds. Usually large soil samples are collected in the field, transported to the laboratory and analyzed using either mechanical sieving or settling tubes (further explained in the coming sections).

Not only the grain-size analysis methods, but sampling techniques in use today also have difficulties due to the lateral, longitudinal and depth variability of sediment size and distribution (Latulippe 2001). Most of the techniques are complicated and require professional experience (Bunte 2001). This chapter briefly discusses a surficial sampling method through visual grain-size estimates.

### 2.1. Mechanical Sieving

Mechanical sieving is a procedure for determining the size distribution of particles in a sediment sample by passing them through stacks of nested sieves with square openings of known diameter (Cheel 2005). As stated in Sahu (1965), mechanical sieving dates back to the early days of mechanical analysis. Sieves were used to separate sands, for the first time, in 1704.

The basic principle of sieving technique is that a representative sample of known weight of particles is passed through a set of sieves of known mesh sizes. As shown in fig. 2.1, the sieves are arranged in downward decreasing mesh diameters and are mechanically vibrated for a fixed period of time. The weight of particles retained on each sieve is measured and converted into a percentage of the

---

total sample. This method is sufficiently accurate for most purposes as long as the materials have similar densities.



Figure 2.1: Mechanical sieve apparatus (Retsch 2008)

## 2.2. Settling Tubes

The technique of settling tubes utilizes the settling of grains through a column of liquid, usually water, from the same starting point at the top. A sediment sample is introduced into a clear column of water for the settling velocity and grain-size to be determined. This method is used for analysis of sands (0.0625 - 2 mm) (Syvitski 1991).

Settling tubes, as shown in fig. 2.2, are the devices used to determine the fall velocity distribution of particles settling out individually. It is assumed that settling particles are not affected by other settling particles and not slowed down by upflow of displaced fluid. This is achieved by introducing

---

only coarse sediment particles, typically in the sand size range, in low concentration, at a common level near the top of a wide settling tube (Syvitski 1991). At the bottom of the tube, the falling particles are sensed, and the frequency distribution of the sample as a function of settling duration is obtained. Depending on the sensing method used, the distribution can be in terms of number of particles, particle volume, particle weight, or projected area. Weight is sensed most of the time. The frequency distribution based on settling duration is converted into the fall velocity distribution using the particles' settling distance (Syvitski 1991).

The grain-size distribution can then be determined using equations relating settling velocity to grain-size. As stated in Syvitski (1991), for the settling of a sphere of any density in any Newtonian fluid and gravitational field, Stoke's Equation may be used to calculate its settling velocity as shown below:

$$w_s = [4/3g \cdot d_s(\rho_s - \rho_f)/C_d \cdot \rho_f]^{0.5} \quad (2.1)$$

where  $w_s$  is the particle's fall or settling velocity,  $g$  is the acceleration due to gravity,  $d_s$  is the particle's sedimentation diameter,  $\rho_s$  is the density of the particle, and  $\rho_f$  is the density of the fluid.  $C_d$  is the particle's drag coefficient and is defined by:

$$C_d = F/[A \cdot \rho_f(w_s^2/2)] \quad (2.2)$$

where  $F$  is the force of resistance exerted by the fluid on the falling particle, and  $A$  is the cross-sectional area of the particle. In this form, tables or nomograms of  $C_d$  versus  $w_s$  must be used, with particle density or fluid density held constant (Syvitski 1991).

### 2.3. Visual Grain-size Estimates

This method involves estimating the grain-size distribution of substrate particles by using visual estimation. Usually, broad grain-size classes are created and visual estimates of these size classes in the sample are done (Kondolf 2003). An example of the size class categorization (similar to

---



Figure 2.2: Settling Tube (Poppe 2001)

the one shown in fig. 1.1) can be the one used by Platts (1983). He differentiated between large boulders (> 610 mm), small boulders (> 305 mm), cobbles (> 76 mm), gravel (> 4.8 mm), large fines (> 0.83 mm), and small fines (< 0.83 mm).

Visual size estimates have been in use by fisheries biologists and engineers in the areas of fluvial ecology to produce substrate maps of river beds for the purpose of habitat modeling (Kondolf 2003). The commonly used way of producing fluvial substrate maps is to visually define boundaries (patches) whose grain-size distribution are approximately similar in the river bed (Latulippe 2001). This is followed by assigning one of the grain-size classes to each patch. An example of this method is shown in fig. 1.1.

Although this method is good for preliminary survey, that is to get familiar with general bed formation of a river during the early stages of a project, it has many drawbacks which limit its use. Good visual estimates require trained and experienced operators as untrained operators can easily introduce large errors. Moreover, the method is relatively subjective and it is difficult to achieve consistent results from different operators.

---

### 3. Size Distribution from Digital Image Analysis - Previous Research

It is necessary and highly desirable that a sampling and analysis technique that achieves a satisfactory characterization of grain size while simultaneously reducing the time spent in both the field and the laboratory is developed (Graham 2005). Recent developments in digital image processing softwares has provided the means for the development of an alternative method for soil grain-size distribution analysis based on digital images (Ghalib 1999).

A number of approaches have been introduced for this and they can generally be categorized into two groups. The first group includes deterministic methods, which use edge detection or grayscale thresholding for component segmentation and component pixel counting. These methods have successfully determined not only the grain-size but also the size distribution. However, these methods have one major limitation. The soil grains have to be detached from one another, otherwise overlapping or touching grains will be interpreted as single large components (Shin 2004).

Because of the limitations of deterministic methods, another group of approaches for analyzing soil images has been introduced. This second group includes statistical methods based on image texture. Texture can be defined as certain patterns which are repeated in an image. In general, these methods represent an image by a set of index values. The index values are more easily obtained from a transformed image than from the original (Shin 2004).

According to Shin (2004), textural analysis (statistical) methods usually involve three main tasks:

1. Image transformation into a form that can be more easily interpreted such as grayscale;
2. Description of the image using index values based on textural features; and

3. Establishment of relationships between the indices and grain-size.

The main advantage of image textural analysis techniques is that they overcome the problem of touching and overlapping grains.

The method introduced and used for this study is, as explained in the next chapter, a textural analysis technique and uses spatial autocorrelation of the intensity of pixels in an image to calculate the grain-sizes.

This chapter presents two deterministic image analysis methods, the first method developed by Sime (2003) uses edge detection approach while the second method developed by Graham (2005) uses a grayscale thresholding approach as explained in the next two sections.

### 3.1. Sime and Ferguson (2003)

Sime and Ferguson investigated the application of automated image analysis to determine the grain-size distribution in gravel-bed rivers. They stated that most of photo-sieving techniques which have been so far developed follow the same procedure of performing the task. The method they developed involves three stages (Sime 2003):

1. detect possible edges between grains, from their distinctively dark appearance or the rapidity of tonal variation across them;
2. clean up the initial edge map to minimize errors and ensure that edges are connected to define discrete grains;
3. and then determine properties of interest for the separate grains.

#### ***Edge Detection***

Sime and Ferguson first converted images to grayscale and then used two different methods to identify the edges of grains: thresholding either the image itself, or a first- or second- derivative map obtained by filtering the image. Image thresholding works on the assumption that dark areas are gaps between grains; derivative methods work on the assumption that the zones of most rapid tonal change are between grains, or at edges between grains and sand or shadow (Sime 2003).

---

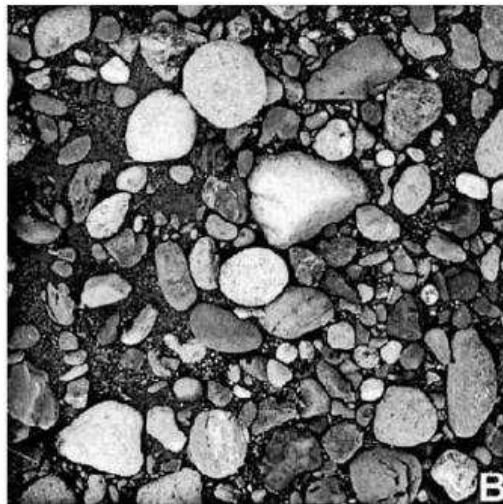


Figure 3.1: Image of a gravel bed surface (Sime 2003)

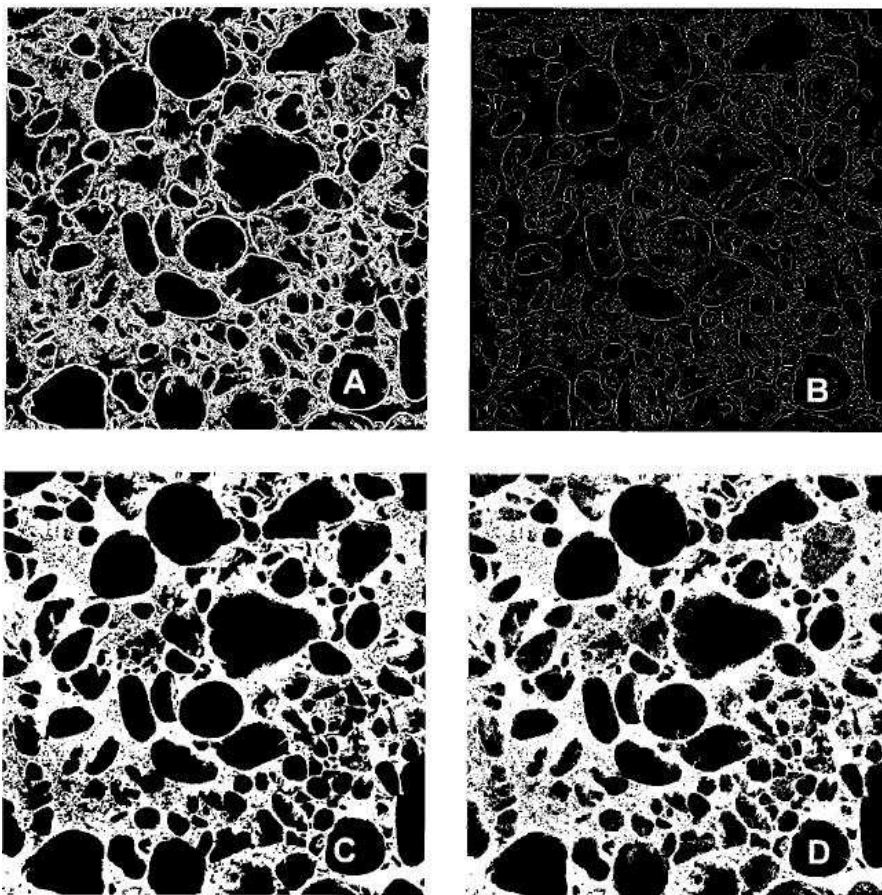


Figure 3.2: Edge maps generated from fig. 3.1 using different edge detection methods (Sime 2003)

### *Edge Tidying*

Edges of grains identified by any of the above methods may not be completely segmented. Secondary processing is therefore required to ensure complete linking of true edges and minimal survival of false edges which might interfere with subsequent edge linking. A good way to do this is watershed segmentation, in which the edge map is expanded through successive episodes of dilation (addition of pixels adjacent to those already identified as edges) until no grain space is left. The map is then eroded (the dilations are reversed) but with the condition that no merging of separate grains is allowed. This should mean that grains which were incompletely separated by edge spurs are now fully separated (Sime 2003).

### *Measuring grain-sizes*

The final stage of Sime and Ferguson's image processing was feature extraction such as measurement of grain-size, shape, or orientation from a fully connected edge map. The method concentrates on the most widely used end product, the grain-size distribution based on  $b$ -axis diameters of grains (Sime 2003). The  $b$ -axis can be identified by first finding the long axis  $a$ , then the short axis  $c$  - the thickness, and finally the intermediate axis  $b$  that is perpendicular to both the  $a$  and  $c$ -axes (see Figure 4). The  $b$ -axis is the axis that governs whether a particle will fit through a sieve mesh of a given size, so measuring it gives results most comparable to a standard sieve test (Collins 2007).

## 3.2. Graham, Stephan and Reid (2005)

Graham, Stephan and Reid divided their automated grain sizing procedure into four key stages (Graham 2005):

1. Image collection,
  2. Image pre-processing,
  3. Image processing and analysis, and
  4. Derivation of grain-size distribution.
-

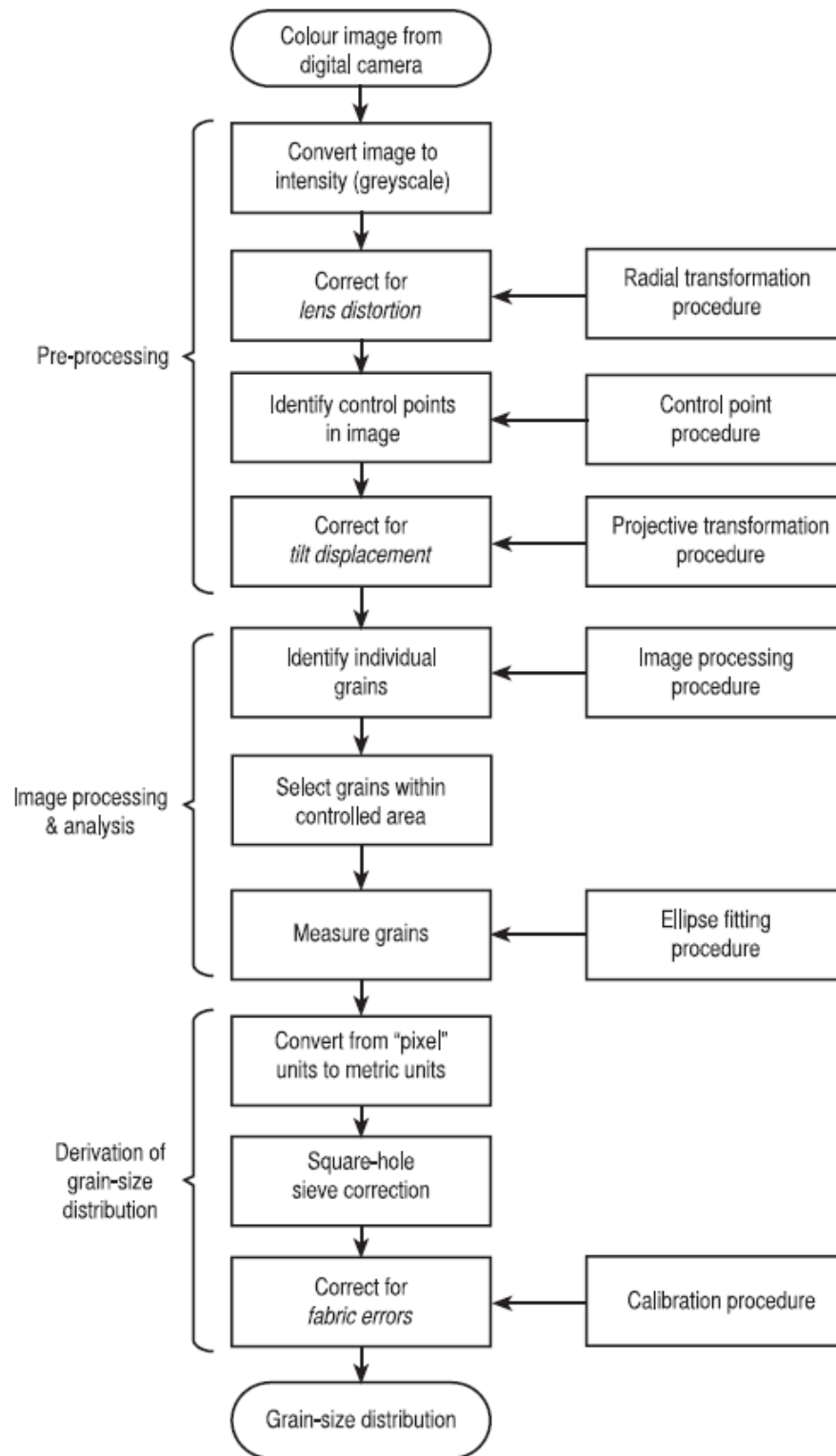


Figure 3.3: The stages followed by Graham, Stephan and Reid

### ***Image Collection***

In their report, Graham, Stephan and Reid, state that the scale of the image taken should be such that the smallest grain of interest has a  $b$ -axis larger than 23 pixels. Moreover, reference points were placed at each corner of the rectangular sample patch to provide a scale and define the boundary of the patch in the image (Graham 2005).

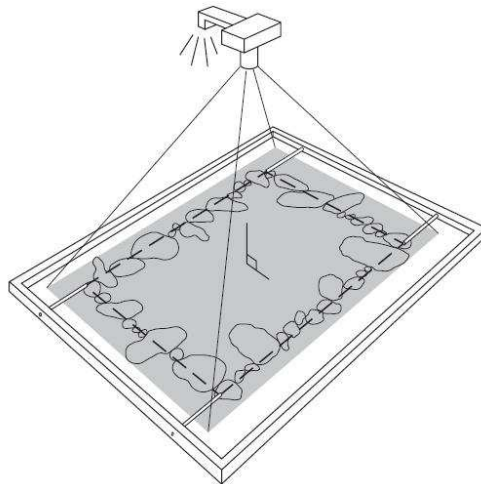


Figure 3.4: Photographic procedure followed by Graham et al. (Graham 2005)

### ***Image Pre-Processing***

The method converts the color image to gray scale and applies correction for radial distortions. Finally, the four corner reference points are identified (Graham 2005).

### ***Image Processing and Analysis***

The image is first manipulated by the application of a median filter which smooths markings on the grain surfaces while preserving edges. Interstices are then enhanced by the application of a morphological bottom-hat transform. A first segmentation is obtained using an adaptive double-threshold approach in which the threshold levels are defined in terms of percentiles in the image-intensity frequency distribution, and this is then refined using a watershed segmentation algorithm with minimum suppression. The selected objects are measured using an ellipse-fitting procedure to obtain the  $b$ -axis (Graham 2005).

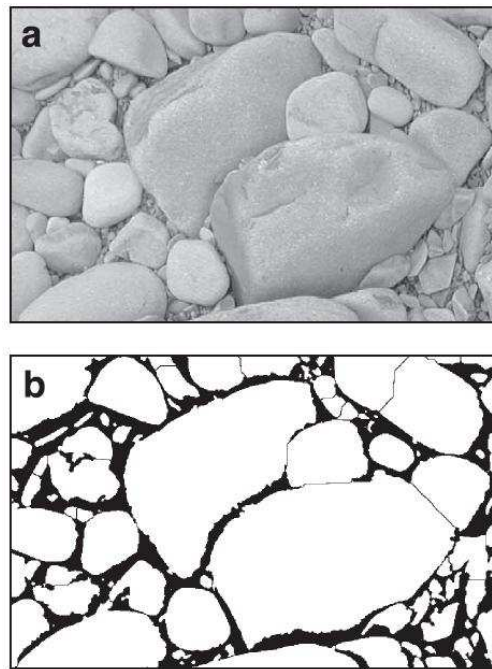


Figure 3.5: Image processing procedure (a) a color image, (b) the same image after image processing (Graham 2005)

### ***Derivation of Size Distribution***

At last a list of  $b$ -axis lengths of the objects identified in the image is obtained, measured in pixels. In the next step, the pixel units are converted to metric units. The measured  $b$ -axes are then used to generate a cumulative grain-size distribution (Graham 2005).

Both of the above mentioned methods utilize digital images and an algorithm that recognizes and measures individual grains to come up with the grain-size distribution of the sample. Even though some degree of success exists after using these methods with careful sampling and detaching of grains, as explained in Rubin (2004), it is difficult to achieve such success with natural field samples as these methods are subject to inaccuracies caused by overlapping of grains (which can cause partially covered grains to appear smaller than they are) or touching of grains of the same color or brightness (which can cause them to appear larger). Moreover, most of the methods so far developed including the methods mentioned above are designed to work only for coarse grains like gravel under completely dry conditions. All methods fail to address fine grains and samples under water.

## 4. Size Distribution from Digital Images Using Autocorrelation Analysis

This method, unlike the previously developed grain-size analyzers from digital image, considers the statistical properties of entire image to calculate the grain-size distribution. The method was developed by David M. Rubin, a geologist in the United States Geological Survey (USGS). Rubin (2004) developed two algorithms (explained in detail in sections 4.2 and 4.3) to come up with grain-size distribution of a sample from a digital image. The first algorithm calculates the autocorrelation values of an image at different offset distances within the image and is used in the calibration as well as in the grain-size analysis stages. The second algorithm, based on the calibration and observed autocorrelation data calculated using the first algorithm, determines the grain-size distribution and the mean grain-size of the sample image.

### 4.1. Spatial Autocorrelation Basics

Spatial autocorrelation is an assessment of the correlation or relationship of a variable in reference to the spatial location of the same variable (NRC 2008). In other words, it is the correlation of a variable with itself through space (Lembo 2000). Spatial autocorrelation measures the level of interdependence between the variables and the nature and strength of the interdependence. Spatial autocorrelation may be classified as either positive or negative. Positive spatial autocorrelation has all similar values appearing together, while negative spatial autocorrelation has dissimilar values appearing in close association (NRC 2008).

---

## 4.2. Spatial Autocorrelation Calculation - First Algorithm

The first algorithm, as mentioned in the previous sections, calculates the spatial autocorrelation of an image at different offset distances. The core of the algorithm is based on the idea that spatial autocorrelation in an image varies with grain-size (Rubin 2004). The idea with which the algorithm was developed is similar to the first law of geography which says:

*“Everything is related to everything else, but near things are more related than distant things”* (Tobler 1970).

Rubin (2004) used the intensity of the pixels in the image as a variable in calculating the autocorrelation values. He defines spatial autocorrelations as follows:

*“Spatial autocorrelation can be defined as the correlation between two rectangular regions (called plaquettes) in an image, measured by calculating the correlation between the intensity of each pixel in one plaquette with the pixel in the corresponding location in the second plaquette. To describe this operation in terms that are perhaps more familiar, the pixels in one plaquette might be considered ‘observed’ values, and the pixels in the second plaquette considered ‘predicted’ values; the spacial autocorrelation is the correlation between the two sets of values”.*

The spatial autocorrelation  $r$  between two plaquettes in an image is calculated using the following equation:

$$r = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2} \sqrt{\sum_i (y_i - \bar{y})^2}} \quad (4.1)$$

Where:

$r$  = the spatial autocorrelation between two plaquettes

$x_i$  and  $y_i$  = the intensities of corresponding pixels in the two plaquettes

---

$\bar{x}$  and  $\bar{y}$  = the mean intensities of pixels in the two plaquettes.

An autocorrelation curve, as shown in fig. 4.1, is determined from the output data of the algorithm which is autocorrelation value  $r$  as a function of the distance between the two plaquettes or the offset distance. As the distance between the two plaquettes is small (relative to the grain-size) the autocorrelation value approaches 1.0 and as the distance increases,  $x_i$  becomes increasingly unrelated to  $y_i$ , the two terms in the numerator become less likely to have the same sign, and the sum of their products decreases and approaches zero (Rubin 2004).

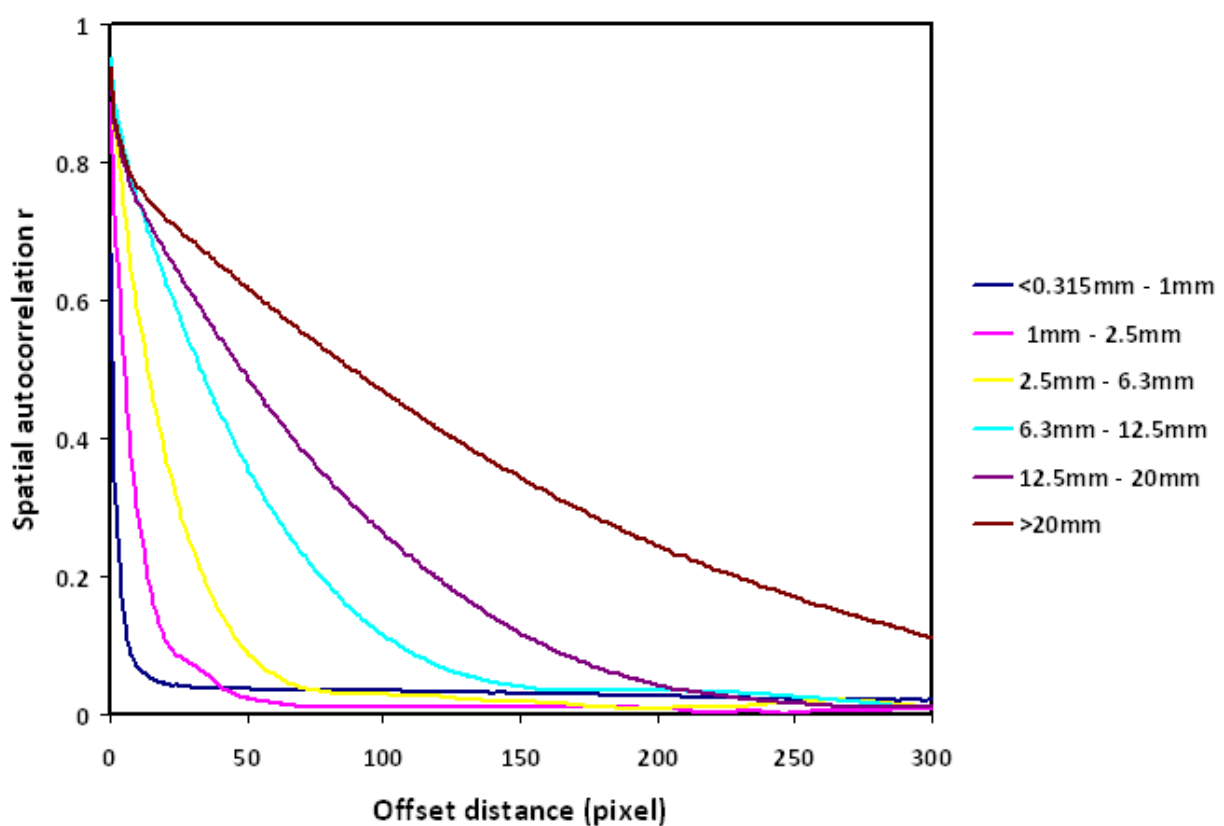


Figure 4.1: Autocorrelation curves of different grain-size fractions. Pixels in larger grains are more similar for a longer distance than pixels in smaller grains (Rubin 2004)

The algorithm works in three major steps:

1. Loads an image in color and displays it.
2. Converts the image to grayscale. A grayscale image is simply one in which the only colors

are shades of gray. The reason for changing the image to grayscale is that less information is needed to be provided for each pixel in the image. A 'gray' color is one which has the same intensity values for the 'red', 'green' and 'blue' components of an RGB (color) image, as a result, only a single intensity value is needed to describe each pixel in a grayscale image as opposed to the three values needed to describe a pixel in an RGB image. Normally, a grayscale intensity is stored as an 8-bit integer giving 256 possible different shades of gray ranging from black (with a value of zero) to white (with a value of 255) (Fisher 2004).

3. Calculates autocorrelation curve of the image for offset distances of one pixel to a user defined maximum-offset number of pixels.

Fig. 4.2 describes the third step of the algorithm. The maximum number of offsets is first defined by the user based on the largest grain-size in the sample. The width of the plaquette is then determined from the maximum offset - if a big number of offsets is specified, the size of the plaquette becomes smaller and vice versa. The first plaquette is always kept constant as a reference plaquette and all the other plaquettes are moved by 1 pixel horizontally each step and are compared to the reference plaquette from which the autocorrelation  $r$  is calculated. As an output the algorithm returns autocorrelation values for each offset distance of the plaquettes. As the offset distance becomes larger relative to the grain-size of the sample, the value of the autocorrelation approaches zero and if the offset distance is very small, it means that the two plaquettes are more or less resting on the same area and as a result the value of the autocorrelation nears 1.0.

### 4.3. Grain-Size Distribution Calculation - Second Algorithm

The second algorithm developed by Rubin calculates the grain-size distribution of a sample based on the autocorrelation curves of the calibration sample. The method involves three steps to determine a complete grain-size distribution of a sample from autocorrelation curves as follows (Rubin 2004):

1. calculating autocorrelation curves (calibration curves) for each grain-size in a sediment pop-
-

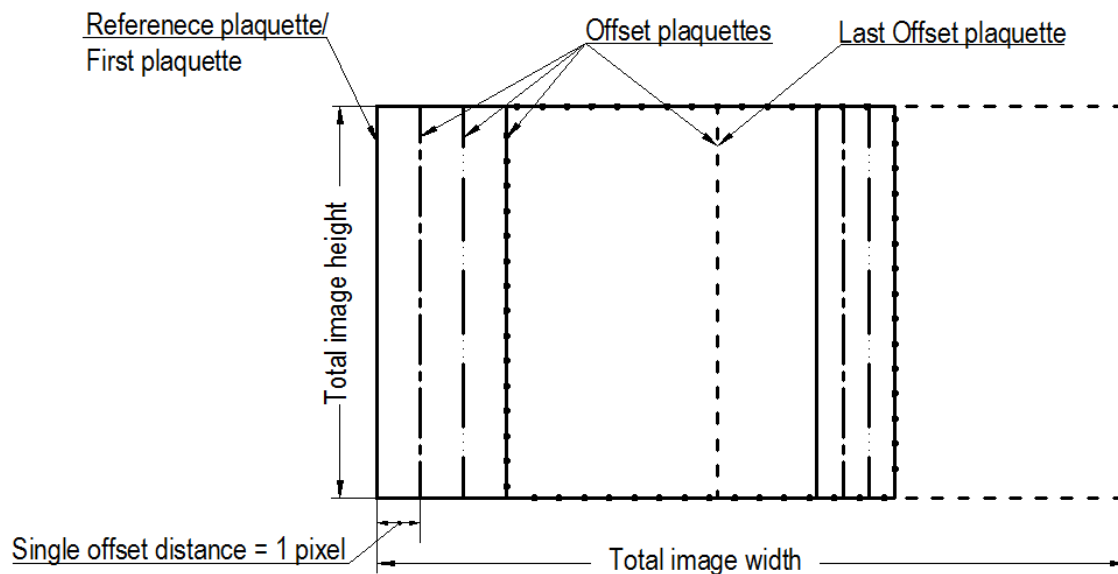


Figure 4.2: Diagram illustrating how the first algorithm works: offset plaquettes moving at single offset distances and each time being compared to the reference plaqueette

ulation (performed once for all samples in a specific region),

2. calculating the autocorrelation curve for a specific sample of sediment from that region, and then
3. solving for the proportions of calibrated sizes that collectively give the best fit to the sample's autocorrelation curve.

As explained above the algorithm determines the grain-size distribution of the sample by calculating the percentage proportion of each grain-size fraction in the calibration data so that a best fit to the autocorrelation curve of the sample can be achieved. This can be described using a set of

equations as follows (Rubin 2004):

$$\begin{aligned}
 a_{(1,1)}x_1 + a_{(1,2)}x_2 + \dots a_{(1,m)}x_m &= b_1 \\
 a_{(2,1)}x_1 + a_{(2,2)}x_2 + \dots a_{(2,m)}x_m &= b_2 \\
 &\vdots \\
 a_{(n,1)}x_1 + a_{(n,2)}x_2 + \dots a_{(n,m)}x_m &= b_n
 \end{aligned}
 \tag{4.2}$$

where:

$m$  = the number of calibration grain-size fractions

$n$  = the number of offsets in pixels at which the correlation coefficient value  $r$  is  $0 < r < 1$

$x_1 \dots x_m$  = the proportions of each of  $m$  grain-size fractions in the sample mixture

$a$  = coefficients representing the autocorrelation values observed for the calibration grain-size fractions

$b$  = coefficients representing the autocorrelation values observed for the sample being analyzed

Each equation in the above set of equations represents and relates the autocorrelation value of the sample  $b$  to the autocorrelation values of each calibration grain-size fractions  $a$  and their proportions in the sample mixture  $x_1 \dots x_m$  at a specific offset distance (Rubin 2004).

The minimum requirement for the above set of equations to be solved is that the number of unknowns ( $x_1 \dots x_m$ ) should be equal to the number of available equations ( $1 \dots n$ ). However, it is usually true that there are more equations than unknowns because the number of useful pixel offsets usually exceeds the number of calibration grain-size fractions ( $n > m$ ). Therefore, accurate solutions can be calculated using a least-squares fit method. Furthermore, as a negative value for a grain-size distribution is not physically possible, more accurate and meaningful solutions are the ones calculated using a non-negative least-squares fit method. The algorithm uses a non-negative least-squares fit method to calculate the proportions of each grain-size fractions and returns the values as an output (Rubin 2004).

To demonstrate how the system of equations is solved to determine the grain-size distribution of a soil sample, an example containing three calibration grain-size fractions and a validation soil sample is shown below.

Output from first algorithm (autocorrelation values) using 3 offset distances is shown in table 4.1:

Table 4.1: Autocorrelation values of calibration fractions

Calibration fraction 1	Calibration fraction 2	Calibration fraction 3	Validation sample
<b>Autocorrelation value (r)</b>			
0.9315	0.8342	0.8236	0.8122
0.8214	0.7023	0.5869	0.7523
0.7598	0.6125	0.4235	0.6255

The set of equations can be written in the following form:

$$\begin{aligned}
 0.9315x_1 + 0.8342x_2 + 0.8236x_3 &= 0.8122 \\
 0.8214x_1 + 0.7023x_2 + 0.5869x_3 &= 0.7523 \\
 0.7598x_1 + 0.6125x_2 + 0.4235x_3 &= 0.6255
 \end{aligned}
 \tag{4.3}$$

After solving the system of equations using a non-negative least squares method, we come up with the following values for the unknowns  $x_1$ ,  $x_2$ , and  $x_3$  which are percentage amounts of each calibration fraction present in the validation soil sample.

$$\begin{aligned}
 x_1 &= 0.45 \\
 x_2 &= 0.49 \\
 x_3 &= 0.00
 \end{aligned}
 \tag{4.4}$$

Fig. 4.3 shows the whole process of grain-size distribution determination of a soil sample using a flow chart.

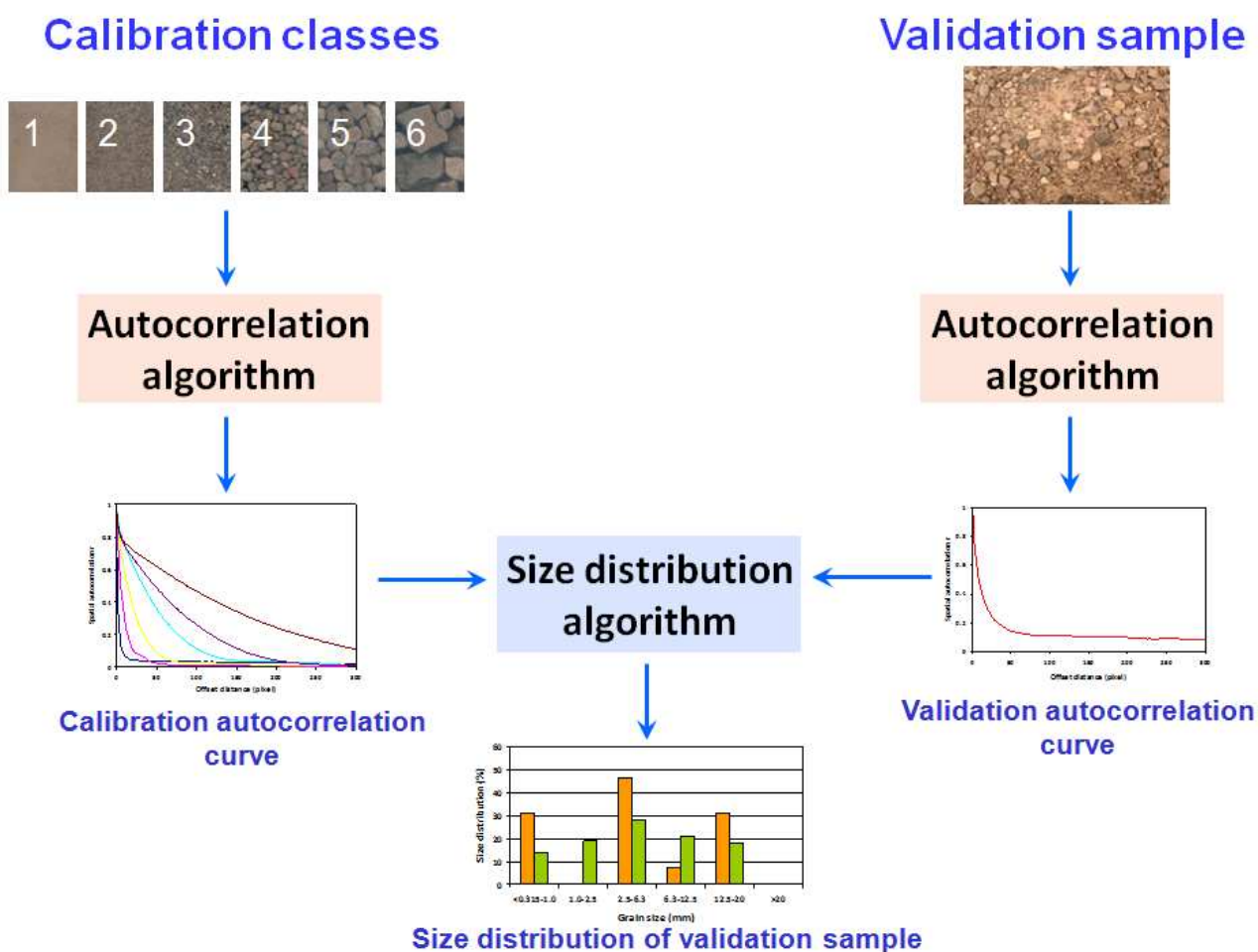


Figure 4.3: A schematic view of the stages of grain-size distribution determination of a soil sample from digital images using autocorrelation analysis

## 5. Experimental Procedures

The experiments were conducted in the laboratory of the Department of Hydraulic Engineering, University of Stuttgart. Furthermore, all the field images and soil samples used in this study were collected from the Schneitbach Creek, located in the Black Forest. To carry out the experiments and to collect the field images and soil samples, a variety of experimental equipment were used. This chapter explains the experimental equipment and procedures used during the study.

### 5.1. Experimental Equipment

#### *Scanner*

An HP-Scanjet 5530 Photosmart Scanner was used to scan the soil samples collected for both calibration and validation purposes. All the samples were scanned at resolutions of 100, 300 and 600 dpi.

#### *Wooden Frame*

A rectangular wooden frame with a size of 210 mm by 300 mm (size of an A4 sized paper), divided into four equal compartments inside, and with a 100% transparent thin plastic cover from the bottom was used to hold soil samples during scanning. The frame was painted completely matte black to avoid light reflection during scanning. Moreover, the transparent plastic cover was selected to be very thin (0.1 mm thick) to minimise refraction of light passing through it. A grayscale calibration strip was attached to the bottom of the box to control the intensity of the scanned images. Fig. 5.1 shows the wooden frame with its dimensions.

#### *Sieves and Shaker*

Standard sieves with opening sizes of 0.315 mm, 0.63 mm, 1.00 mm, 1.60 mm, 2.50 mm, 4.00

---

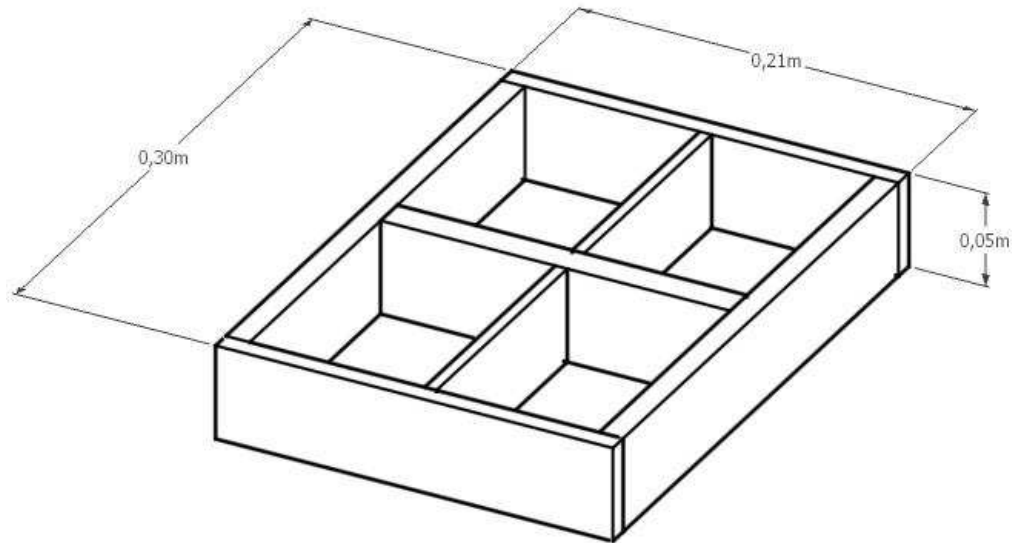


Figure 5.1: Wooden frame used for holding soil samples during scanning

mm, 6.30 mm, 10.00 mm, 12.50 mm, 16.00 mm, and 20.00 mm were used sequentially in a set to carry out all the sieve analysis of the soil samples. As shown in fig. 5.2 A FRITSCH made vibratory sieve shaker was used to shake the samples, each for a period of 10 minutes as recommended by the German standard (DIN-18123 1996).

### ***Drying Oven***

The samples were dried before sieving at a temperature of 105 °C for 24 hours as recommended by the German standard (DIN-18123 1996) using a MEMMERT UFB 500 drying oven. Fig. 5.2 shows the drying oven used for the study.

### ***Scale***

A SARTORIUS BA3100 digital bench scale was used to weigh the soil samples during all sieving experiments. The readings from the scale were recorded at an accuracy of two decimal points.

### ***Underwater Digital Camera***

Field and laboratory sample images were taken using a 7 Megapixel Canon PowerShot S70 underwater digital camera, shown in fig. 5.3. All images were taken using the highest resolution available in the camera which is 3072 x 2304 pixels.



Figure 5.2: Set of sieves mounted on a shaker (left), Drying oven (right)



Figure 5.3: Digital bench scale (left), Underwater digital camera (right)

### *Sampler*

An 18.5 cm by 18.5 cm metal box sampler open on one end was used to take soil samples of the top 3 cm (large rocks were removed by hand) immediately after taking images of the area so that the results from both sieve analysis and image analysis can be compared. Fig. 5.4 shows the sampler used during the field experiments.



Figure 5.4: Metal box sampler

### *Computer*

A 2.39 GHz, 1 GB RAM HP Compaq dc5750 Microtower PC was used for running the algorithms.

### *Software*

The algorithms were run using MATLAB version R2007b.

## 5.2. Experimental Procedures

The experiments carried out during the study were divided into three stages. The following sections explain the three experimental stages in detail.

### 5.2.1. First Stage

This stage is the simplest of the three stages. During this stage two samples of soil were collected from Schneitbach creek to be used for calibration and validation purposes. The samples were dried in an oven at a temperature of 105 °C for 24 hours. Both samples were mechanically sieved using a set of 8 sieves (0.315 mm, 0.63 mm, 1.00 mm, 1.60 mm, 2.50 mm, 4.00 mm, 6.30 mm, 10.00 mm). As the collected samples were very small in amount, the soil fractions were categorized into four classes as follows:

1. grains passing the 0.315 mm sieve (<0.315 mm),
2. mixture of grains passing the 0.315 mm sieve and retained in the 2.5 mm sieve (0.315 mm - 2.5 mm),
3. mixture of grains retained in the 2.5 mm sieve and those passing the 10 mm sieve (2.5 mm - 10 mm),
4. grains retained in the 10 mm sieve (>10 mm)

Each class from both samples was then scanned at resolutions of 100 dpi, 300 dpi, and 600 dpi consecutively. Moreover, two more classes were created by mixing classes 1 and 2 of the validation sample to form class-5 (<0.315 mm - 2.5 mm), and by mixing classes 3 and 4 of the validation sample to form class-6 (2.5 mm - >10 mm). Classes 5 and 6 were also scanned at the above mentioned three resolutions.

The autocorrelation curves for calibration were obtained by running the scanned images from the first sample and the images from the second sample were used to generate the autocorrelation curves for validation. At last grain-size distribution of each class in the validation sample was calculated and was compared to the results obtained by mechanical sieving.

### 5.2.2. Second Stage

At the second stage of the experiments a larger soil sample was collected from Schneitbach creek to be used for calibration purpose. The sample was mechanically sieved and categorized into six

---

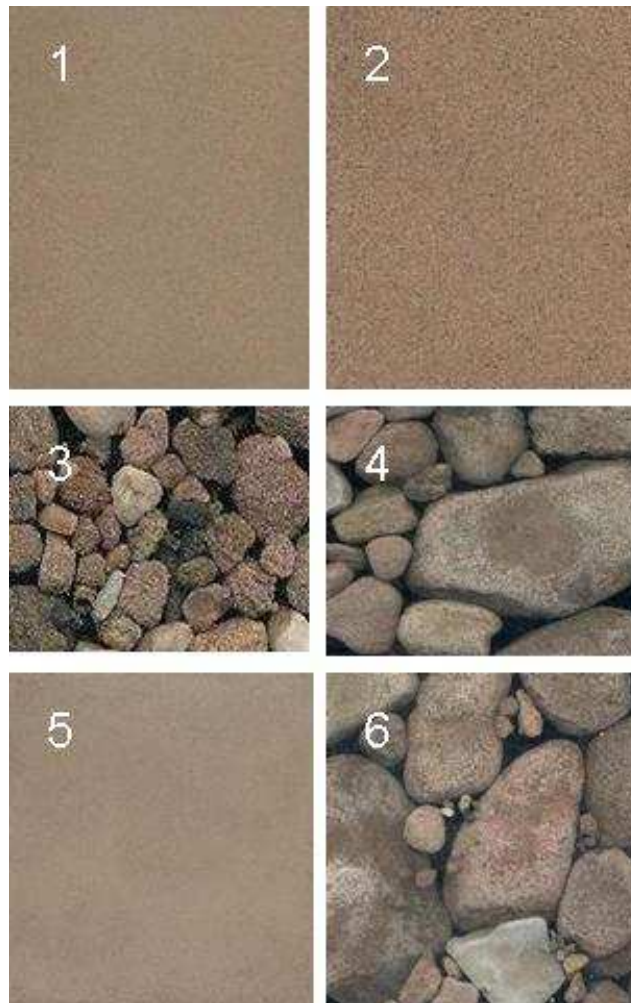


Figure 5.5: Images from first stage scans, (1) class-1 ( $<0.315$  mm), (2) class-2 (0.315 mm - 2.5 mm), (3) class-3 (2.5 mm - 10 mm), (4) class-4 ( $>10$  mm), (5) class-5 ( $<0.315$  mm - 2.5 mm), (6) class-6 (2.5 mm -  $>10$  mm)

classes as shown in fig. 5.6. The classing of the grain-size fractions was performed eventually to come up with soil ranges which are typically used for river habitat modeling. Grain-size fractions with very close grain-sizes and resulting in very close autocorrelation curves were mixed to form one class. The following list explains the created classes:

1. mixture of grains passing the 0.315 mm sieve and passing the 1 mm sieve ( $<0.315$  mm - 1 mm),
2. mixture of grains retained in the 1 mm sieve and passing the 2.5 mm sieve (1.0 mm - 2.5 mm),

3. mixture of grains retained in the 2.5 mm sieve and passing the 6.3 mm sieve (2.5 mm - 6.3 mm),
4. mixture of grains retained in the 6.3 mm sieve and passing the 12.5 mm sieve (6.3 mm - 12.5 mm),
5. mixture of grains retained in the 12.5 mm sieve and passing the 20 mm sieve (12.5 mm - 20 mm),
6. grains passing the 20 mm sieve (>20 mm)

Each class was then scanned at 300 dpi and 600 dpi. The scanned images were used to calculate the calibration autocorrelation curves.

As a second step during this stage underwater river bed images were obtained at four locations from Schneitbach creek using a camera placed at a height of 40 cm from the bed (150 dpi) for field image 1 and 30 cm (200 dpi) for image 2, 3 and 4. Fig. 5.7 shows the field images taken at Schneitbach creek. At the same time the surficial sediment material (approximately top 3 cm) included in the images at the four locations was collected using the sampler shown in section 5.1. The grain-size distribution of the field images was then calculated using the calibration curves from the scanned images. Finally, the surficial sediment material collected at the four locations were mechanically sieved and the size distribution was compared to that obtained from image analysis.

### 5.2.3. Third Stage

In the first two stages of the experiment only scanned images were used for calibrating the models, at the third stage however, it was decided to calibrate the model using images of sediment classes taken using a camera. Images of each calibration sediment class was taken under the following five conditions separately:

- indoors, camera placed at a height of 30 cm (200 dpi resolution), with flash (30FI),
  - indoors, camera placed at a height of 30 cm (200 dpi resolution), without flash (30NFI),
-

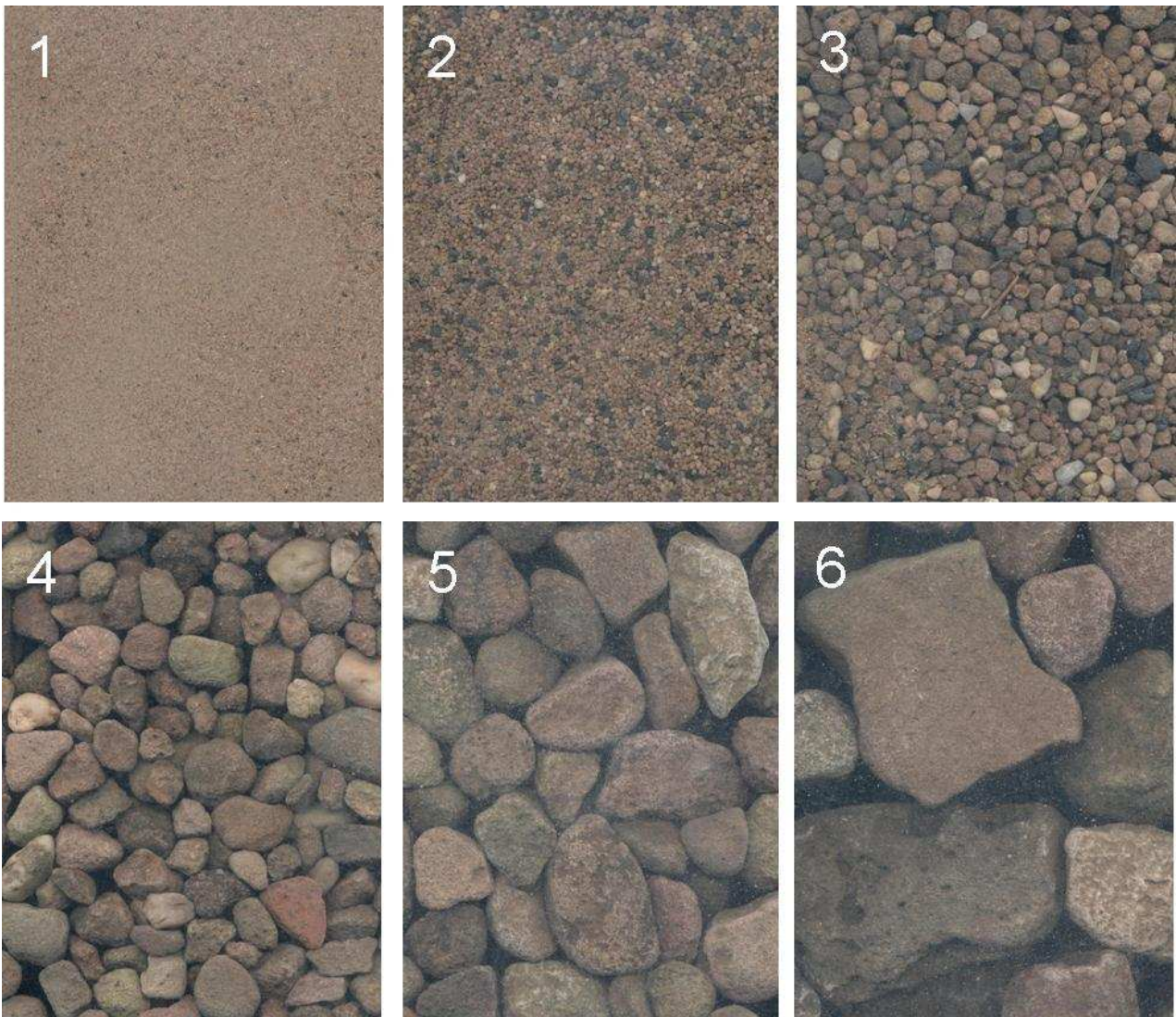


Figure 5.6: Images from second stage scans, (1) class-1 ( $<0.315\text{ mm} - 1\text{ mm}$ ), (2) class-2 (1.0 mm - 2.5 mm), (3) class-3 (2.5 mm - 6.3 mm), (4) class-4 (6.3 mm - 12.5 mm), (5) class-5 (12.5 mm - 20 mm), (6) class-6 ( $>20\text{ mm}$ )

- indoors, camera placed at a height of 40 cm (150 dpi resolution), without flash (40NFI),
- outdoors, camera placed at a height of 30 cm (200 dpi resolution), without flash (30NFO),
- outdoors, camera placed at a height of 40 cm (150 dpi resolution), without flash (40NFO)

After running the first algorithm for each image, five calibration data sets (autocorrelation curve sets) were obtained, each for one of the above conditions. The calibration curves were then used to calculate the grain-size distribution of the four underwater field images obtained in the second

---

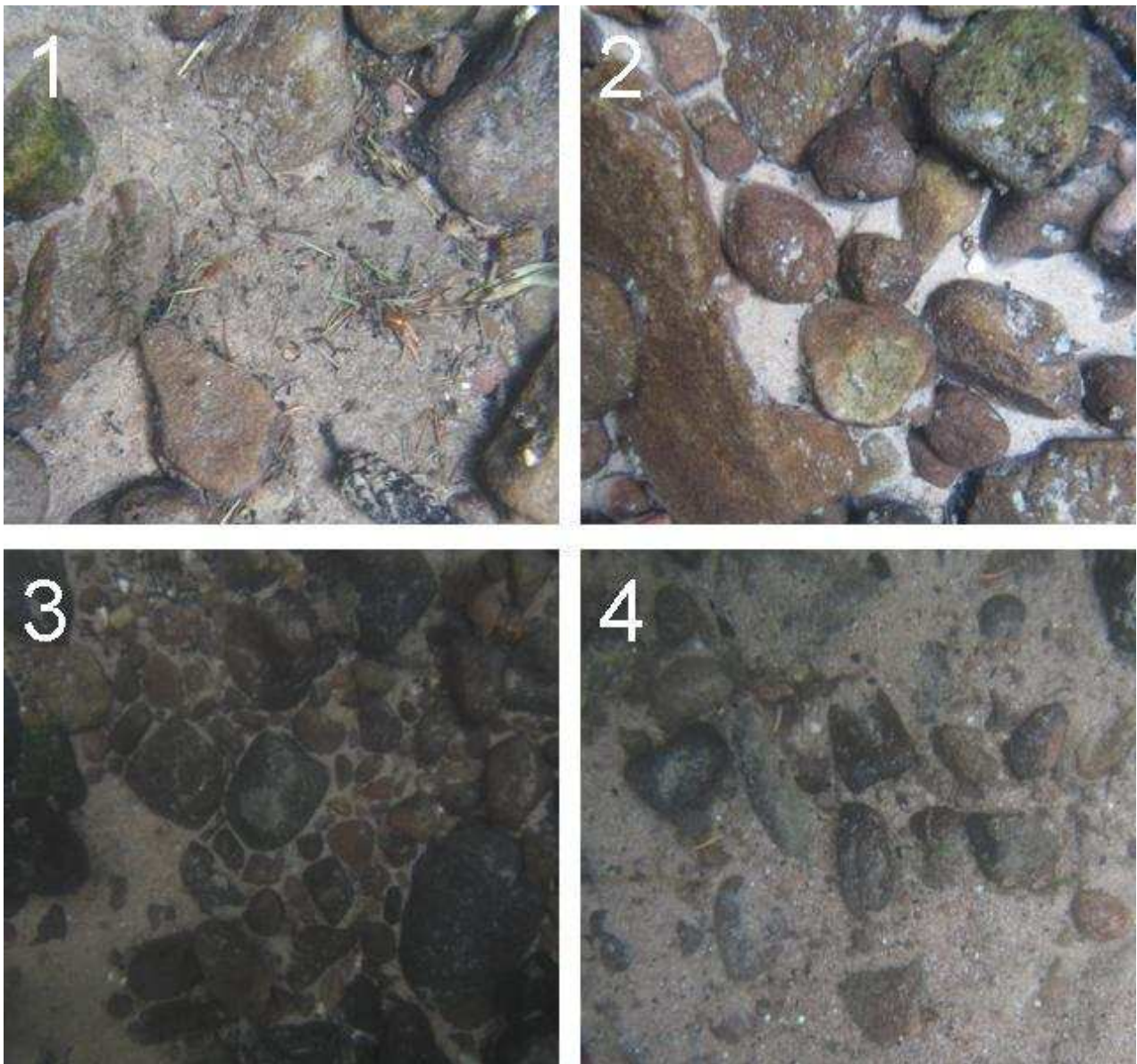


Figure 5.7: underwater river bed images, (1) field image 1 (downstream of Scheitbach creek), (2) field image 2 (downstream of Scheitbach creek), (3) field image 3 (upstream of Scheitbach river), (4) field image 4 (upstream of Scheitbach creek)

stage.

At last, three random mixture of sediments (classes 7, 8 and 9 in fig. 5.8) of known amount (known size distribution) were made from the calibration sediment classes and their images were taken under all the above five conditions. Their size distribution were calculated for every condition and were compared to the real grain-size distribution obtained by sieving. One of the samples was made to include more coarse grains ( $> 20$  mm) to check the algorithm's sensitivity in sensing

---

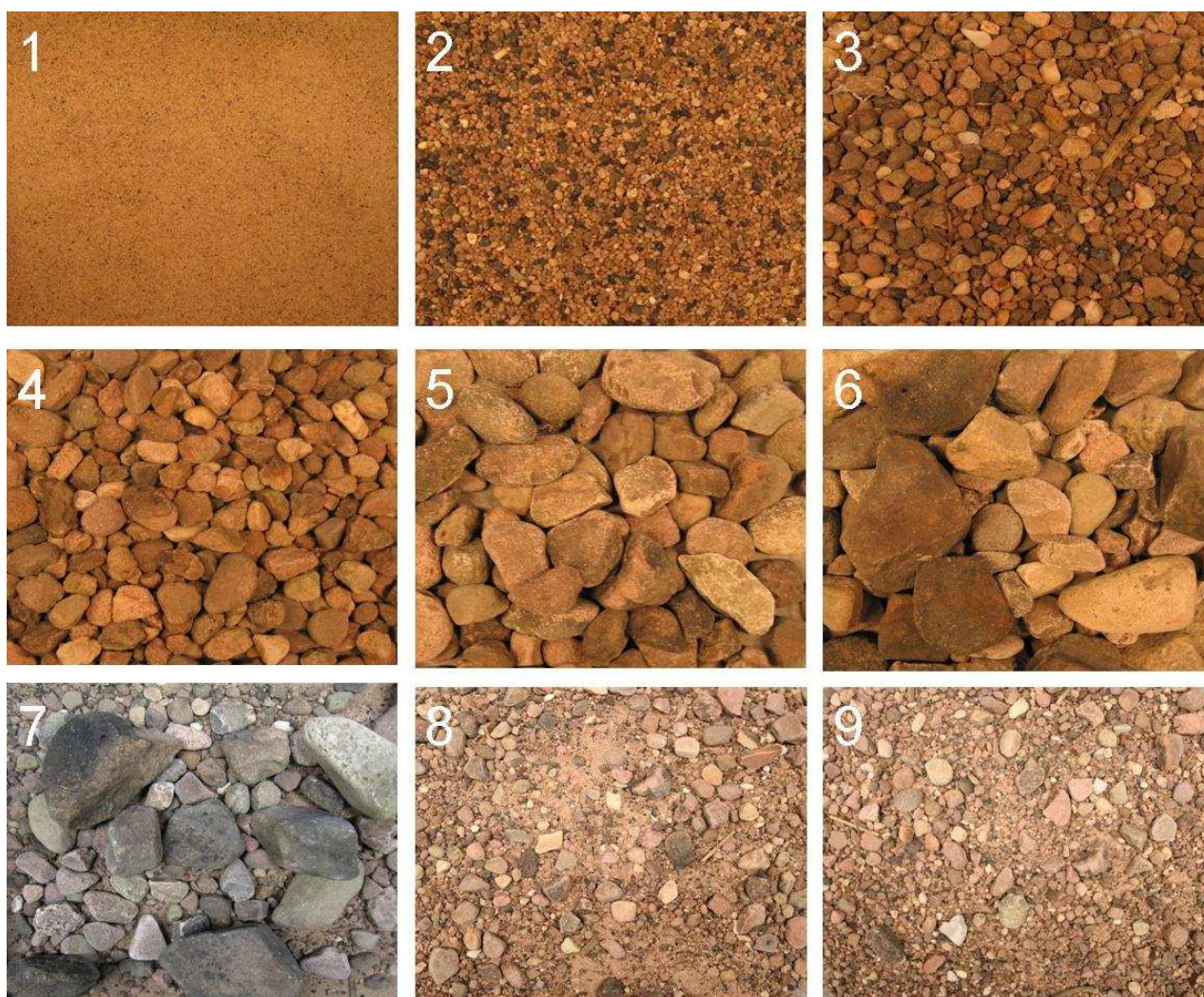


Figure 5.8: Images from third stage taken at 30Fl, (1) class-1 ( $<0.315$  mm - 1 mm), (2) class-2 (1.0 mm - 2.5 mm), (3) class-3 (2.5 mm - 6.3 mm), (4) class-4 (6.3 mm - 12.5 mm), (5) class-5 (12.5 mm - 20 mm), (6) class-6 ( $>20$  mm), (7) class-7, (8) class-8, (9) class-9

coarse grained sample. In the other two samples grains greater than 20 mm were not included and the distribution was made uniform (equal amount from each grain-size fraction) to check the algorithm's ability to recognize all the existing grain fraction with out being biased by the presence of coarse grains.

Tables 5.1 and 5.2 summarize the grain-size classes, soil samples, image resolutions and lighting conditions used for calibration and validation purposes respectively during the three stages of the experiments carried out during this thesis.

Table 5.1: Summary of grain-size classes, image resolutions, and lighting conditions used for calibration during the three stages of the study

Calibration				
Stages	Image type	Calibration grain-size classes	Resolutions	Lighting conditions
1	Scanned	class-1 (<0.315 mm), class-2 (0.315 mm - 2.5 mm), class-3 (2.5 mm - 10 mm), class-4 (>10 mm)	100, 300, 600 dpi	Scanner lighting
2	Scanned	class-1 (<0.315 mm - 1 mm), class-2 (1.0 mm - 2.5 mm), class-3 (2.5 mm - 6.3 mm), class-4 (6.3 mm - 12.5 mm), class-5 (12.5 mm - 20 mm), class-6 (>20 mm)	300, 600 dpi	Scanner lighting
3	Camera images in the laboratory	The same as stage 2	Camera placed at 30 cm, 40 cm from the ground	Indoor with flash (FI), indoor without flash (NFI), outdoor without flash (NFO)

Table 5.2: Summary of soil samples, image resolutions, and lighting conditions used for validation during the three stages of the study

Validation				
Stages	Image type	Validation soil samples	Resolutions	Lighting conditions
1	Scanned	class-1 (<0.315 mm), class-2 (0.315 mm - 2.5 mm), class-3 (2.5 mm - 10 mm), class-4 (>10 mm), class-5 ((<0.315 mm -2.5 mm), class-6 (2.5 mm - >10 mm)	100, 300, 600 dpi	Scanner lighting
2	Field photos	field sample-1, field sample-2, field sample-3, field sample-4,	Camera placed at 30 cm, 40 cm from the ground	Natural lighting
3	Camera images in the laboratory and field photos	field sample-1, field sample-2, field sample-3 field sample-4 lab sample class-7 lab sample class-8 lab sample class-9	Camera placed at 30 cm, 40 cm from the ground from the ground	Indoor with flash (FI), indoor without indoor without flash (NFI), outdoor without flash (NFO)

## 6. Results

Some of the results of stages 1, 2 and 3 are presented in this chapter. All the results show, using bar graphs, comparison of the sample's grain-size distribution obtained by digital image analysis and that obtained by mechanical sieving. Complete results from all the stages can be found in the Appendix of this report.

Table 6.1 shows a summary of the scans, camera imaging, and sieving performed during the study period. Two to five images per sample were taken and autocorrelation results were averaged.

Table 6.1: Summary of total experiments performed for this study

Stage	Total scans	Total camera images	Total autocorrelation analysis	Total image GSD analysis	Total sieve analysis
1	72	0	72	18	2
2	48	16	64	28	7
3	from stage 2	41	41	11	from stage 2
<b>Total</b>	<b>120</b>	<b>57</b>	<b>177</b>	<b>57</b>	<b>9</b>

### 6.1. First Stage

#### 6.1.1. 100 dpi Resolution

Fig. 6.1 shows the grain-size distribution of a sample containing grains less than 0.315 mm (sample class-1). It can be shown from fig. 6.1 that the algorithm was not able to detect all the grains in the sample. This is because at a resolution of 100 dpi each grain was not represented by more

than one pixel in the image. As mentioned in the previous chapters, for the algorithm to give good results, each grain in the sample must be represented by more than one pixel. If calculated accordingly, a sample of grains of size 0.315 mm should be imaged at a minimum resolution of 162 dpi for the algorithm to recognize the grains.

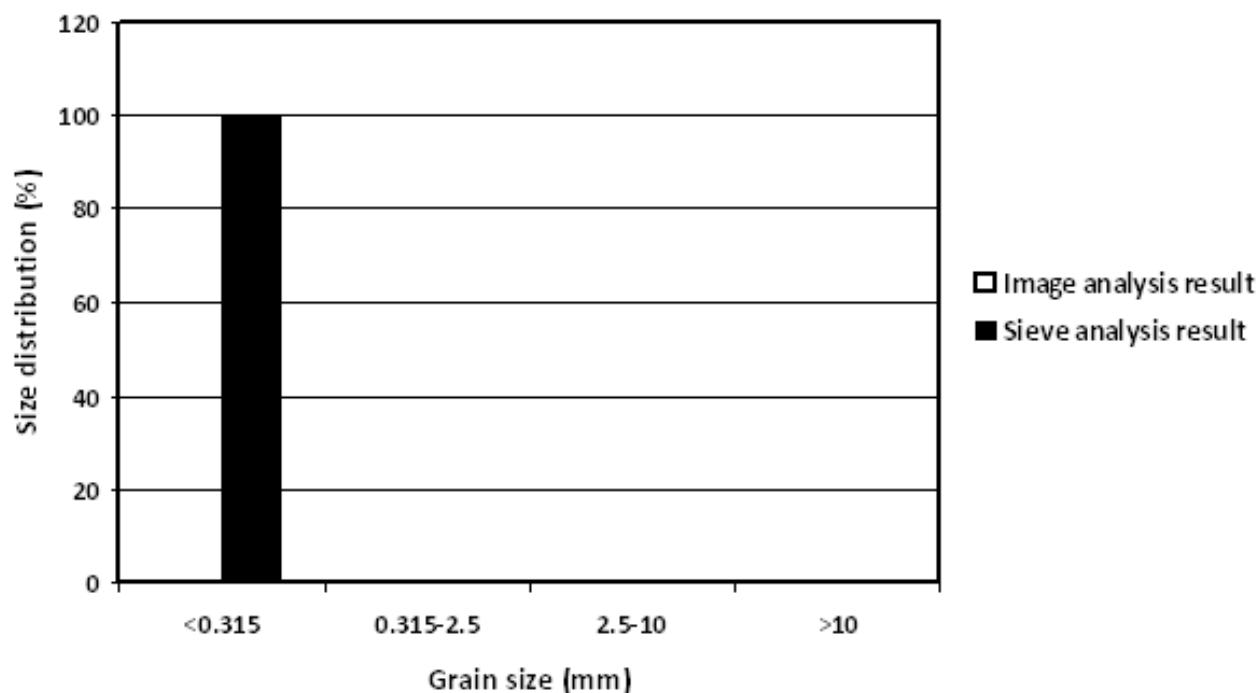


Figure 6.1: Comparison of actual and computed GSD for grains <0.315 mm (class-1), images taken at a resolution of 100 dpi. The algorithm was unable to detect the grains of the sample due to reasons explained in section 6.1.1.

Grain-size distribution of class-2 (grains between 0.315 - 2.5 mm) is shown in fig. 6.2. For the same reason mentioned above, the algorithm was unable to detect the grains in this sample.

The algorithm returned very good results for sample class-4 (grains greater than 10 mm) as shown in fig. 6.3. The absolute mean error calculated for this result was 3.15%, which is very small and completely acceptable. The calculated correlation between the image GSD (grain-size distribution) and the sieved GSD was 1, showing that the distribution predicted by the algorithm followed the same trend as the distribution determined by sieving.

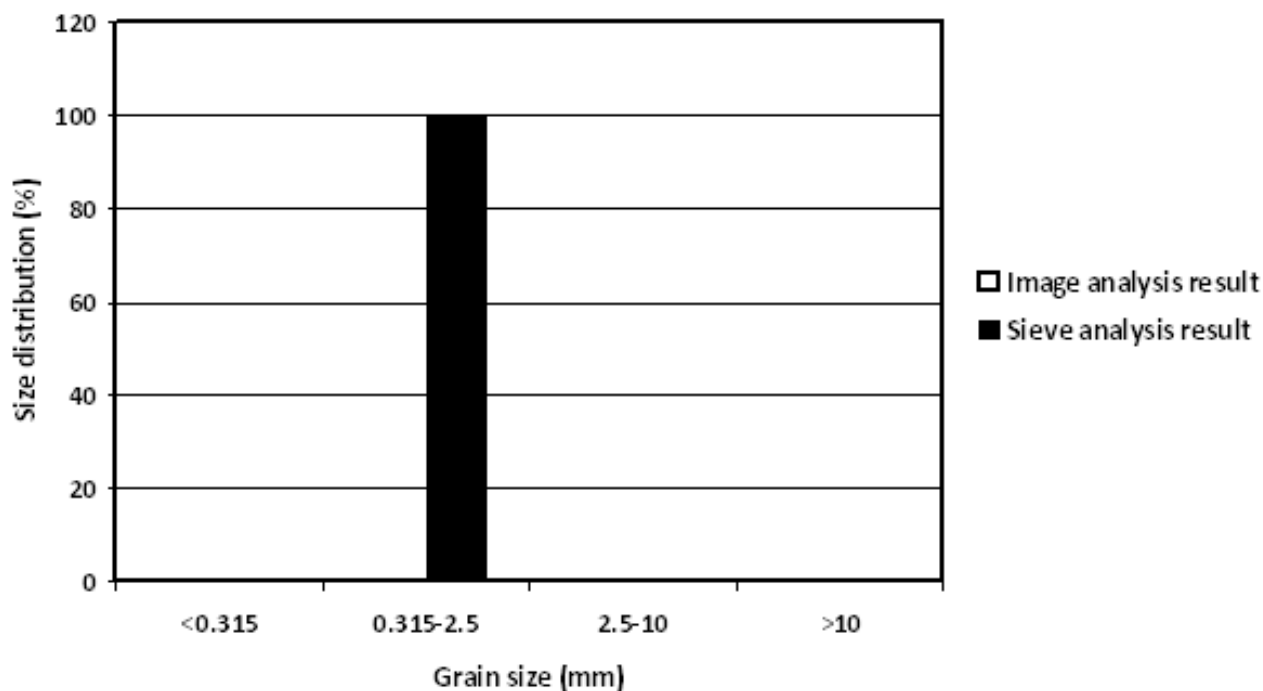


Figure 6.2: Comparison of actual and computed GSD for grains between 0.315 - 2.5 mm (class-2), images taken at a resolution of 100 dpi. The algorithm was unable to detect the grains of the sample due to reasons explained in section 6.1.1.

### 6.1.2. 300 dpi Resolution

The algorithm was able to detect the grains in class-1 (<0.315 mm) and class-2 (0.315-2.5 mm) for images taken at a resolution of 300 dpi. This results show that at 300 dpi, all the grains were represented by more than one pixel. The results are shown in figs. 6.4 and 6.5.

A relatively large error of 22.17% was detected for sample class-5 (grains <0.315-2.5 mm) as shown in fig. 6.6. This is because the sample was relatively big and formed a thickness of soil when put in the scanning box (explained in section 5.1). The formation of the thickness protected part of the sample from being included in the scan.

### 6.1.3. 600 dpi Resolution

For 600 dpi resolution images the algorithm predicted the grain distributions for all size classes with smaller errors. This shows that as the resolution of the images is higher the higher the number of pixels per grain and thus better GSD results. The results of sample class-2 (0.315-2.5 mm) and

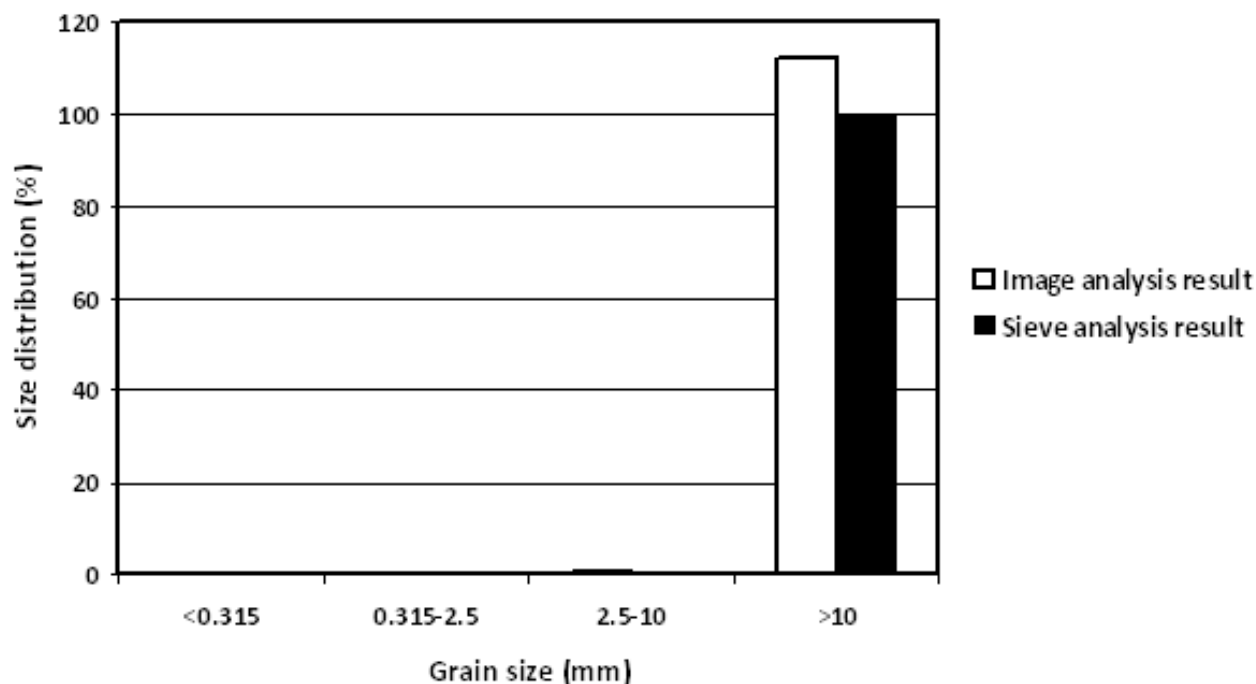


Figure 6.3: Comparison of actual and computed GSD for grains >10 mm (class-4), images taken at a resolution of 100 dpi

sample class-6 (2.5->10 mm) are shown in fig. 6.7 and in fig. 6.8 respectively.

## 6.2. Second Stage

In this stage very large errors and small correlation values between image and sieve analysis results were observed. This was due to two reasons; first, since the calibration images were scanned at resolutions of 300 and 600 dpi and the field images were taken using a camera at a resolution of 200 dpi, the grains were represented with different number of pixels in both cases. Moreover, the lighting conditions in both cases were completely different, one being bright lighting from a scanner and the other at natural lighting conditions. The second reason is that field images 3 and 4, shown in fig. 5.7, were taken in a completely shadowed environment and resulted in a very low brightness and thus bad autocorrelation values.

Figures 6.9, 6.10, 6.11, and 6.12 show the results of field images 1, 2, 3, and 4 respectively for a calibration using scanned images at a resolution of 300 dpi.

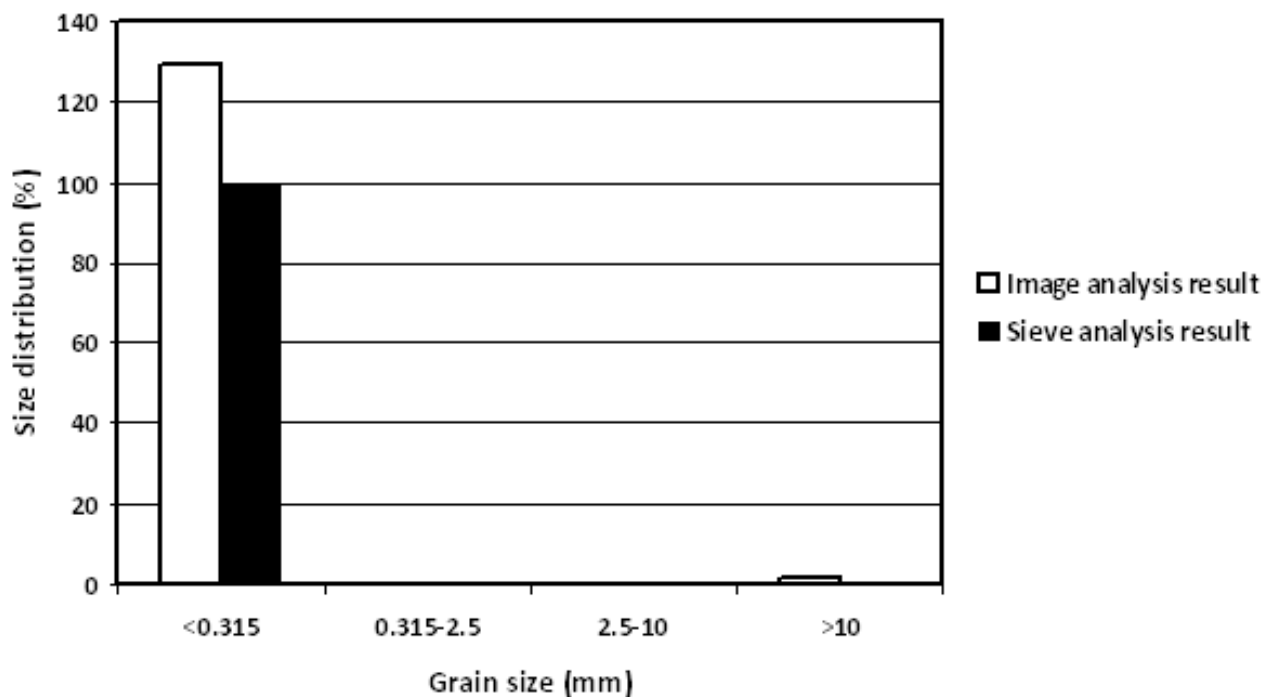


Figure 6.4: Comparison of actual and computed GSD for grains <0.315 mm (class-1), images taken at a resolution of 300 dpi

### 6.3. Third Stage

In the first part of this stage the grain-size distribution of the four underwater field images was calculated using laboratory photos as calibration. The algorithm predicted size distribution for field images 1 and 2 accurately with small errors and strong correlation between the computed and actual distributions. Particularly, calibration using images taken indoors, camera placed at 40 cm, without flash (40NFI) returned the best result for field image 1; and calibration using 30NFI returned the best result for field image 2 as shown in figs. 6.13 and 6.14 respectively. This is because field image 1 was taken at a height of 40 cm and field image 2 at 30 cm making their resolution similar to the calibration images. Moreover, the lighting conditions in the field was very similar to the indoor lighting conditions used for calibration images. It can be concluded from this fact that using similar resolution and lighting conditions for the calibrations and validation images returns the most accurate results.

The results for field images 3 and 4 in this stage showed relatively large errors and small correlation

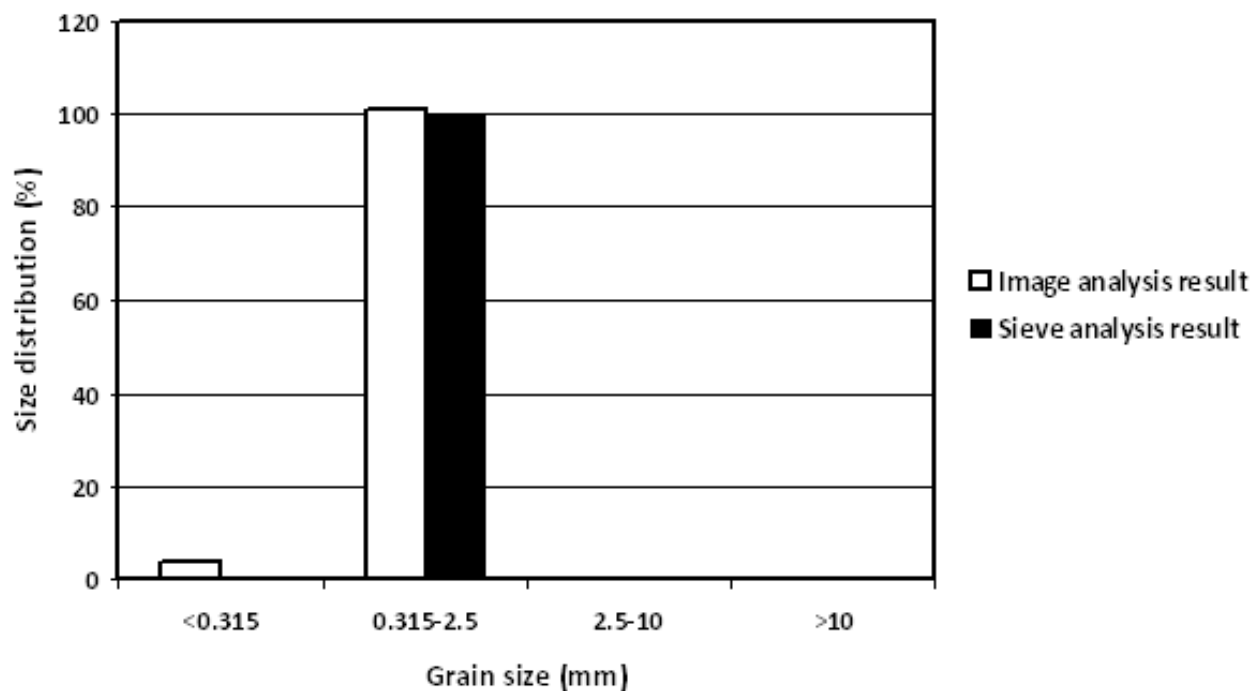


Figure 6.5: Comparison of actual and computed GSD for grains between 0.315 - 2.5 mm (class-2), images taken at a resolution of 300 dpi

values. This is, as explained in section 6.2, due to the fact that field images 3 and 4, shown in fig. 5.7, were taken in a completely shadowed environment and resulted in a very low brightness (underexposed) and thus bad autocorrelation values. The results are shown in figs. 6.15 and 6.16. During the second part of this stage the grain-size distributions of the three random samples (classes 7, 8 and 9), prepared by mixing grain fractions of known quantity and shown in fig. 5.8, were calculated. Sample class-7 contained more coarse grains ( $>20$  mm) and the algorithm returned accurate results with acceptable errors for all the five calibration lighting conditions. Fig. 6.17 shows the result of sample class-7 calculated using 30NFI calibration images.

Only finer grains and no coarse grains ( $>20$ mm) were included in sample classes 8 and 9. They were both analyzed using 30FI, 30NFI and 40NFI calibration conditions. It was observed from the results, as shown in fig. 6.18, that images taken with flash seem to exaggerate the amount of very fine grains ( $<0.315-1$  mm). The 30NFI and 40NFI returned relatively accurate results. Fig. 6.19 shows the result of sample class-9 using 30NFI calibration images. It was concluded from this second part of the third stage experiments that the algorithm is able to predict reliable results for

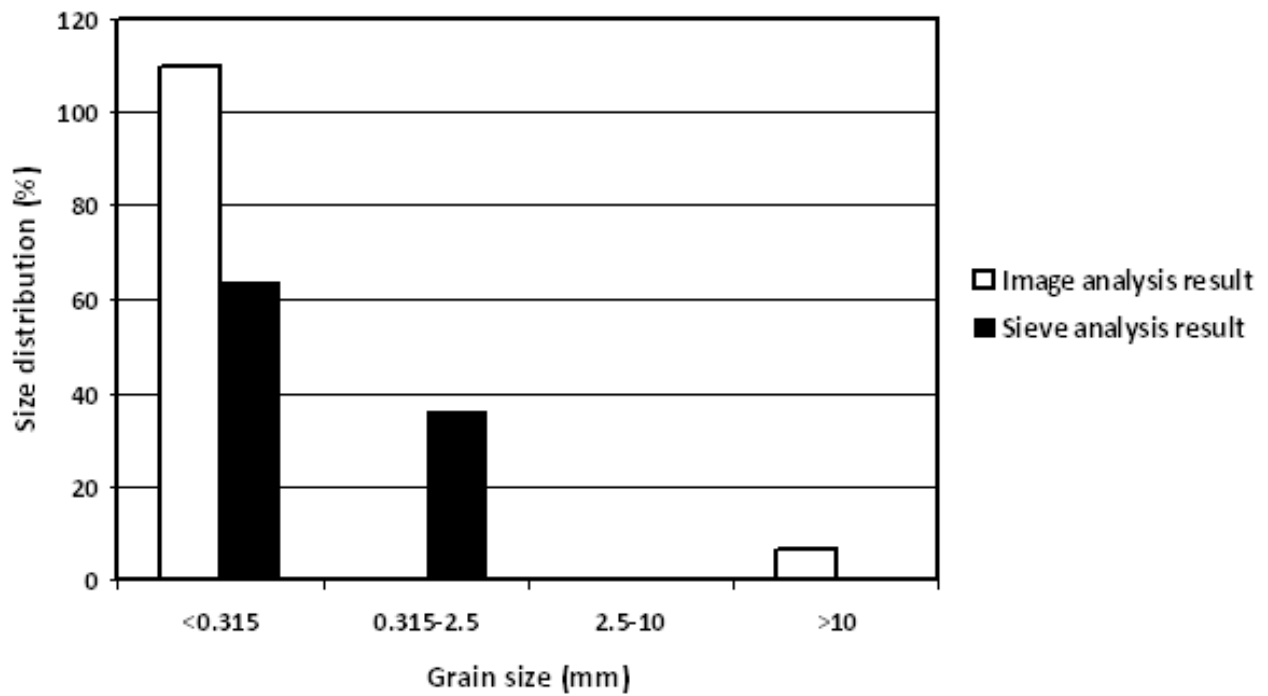


Figure 6.6: Comparison of actual and computed GSD for grains <0.315 - 2.5 mm (class-5), images taken at a resolution of 300 dpi

samples containing fine grains as well as samples with a large percentage of coarse grains.

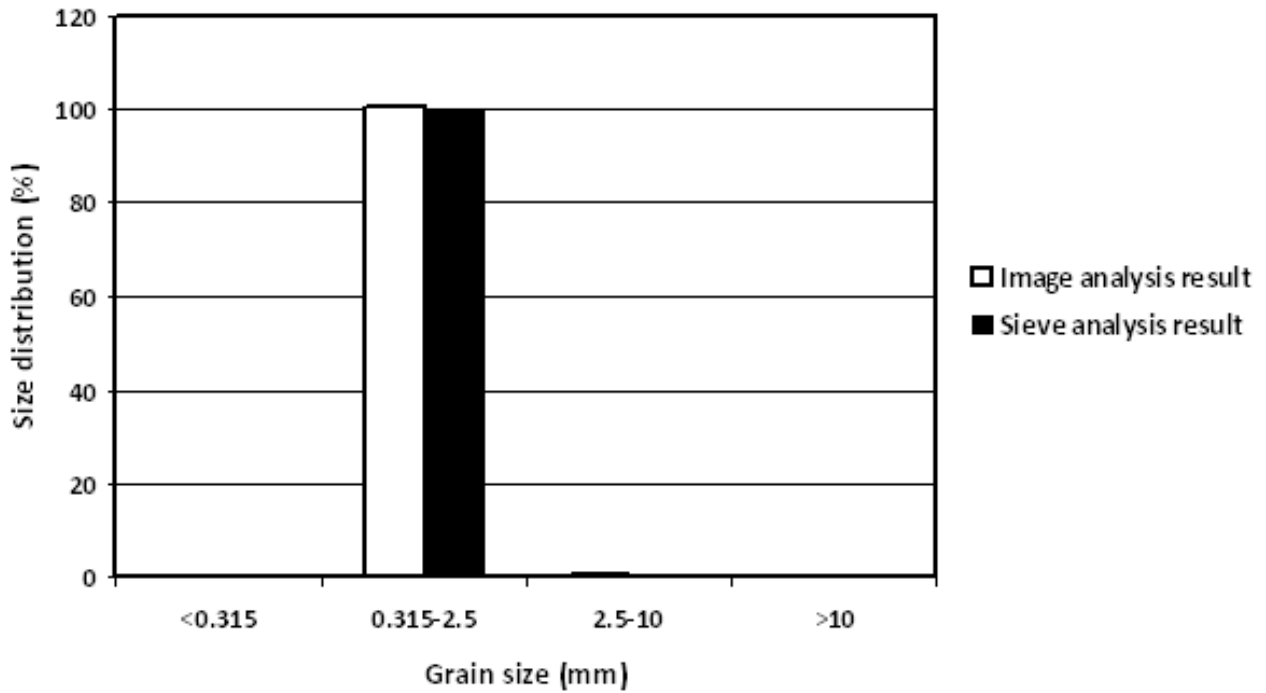


Figure 6.7: Comparison of actual and computed GSD for grains between 0.315 - 2.5 mm (class-2), images taken at a resolution of 600 dpi

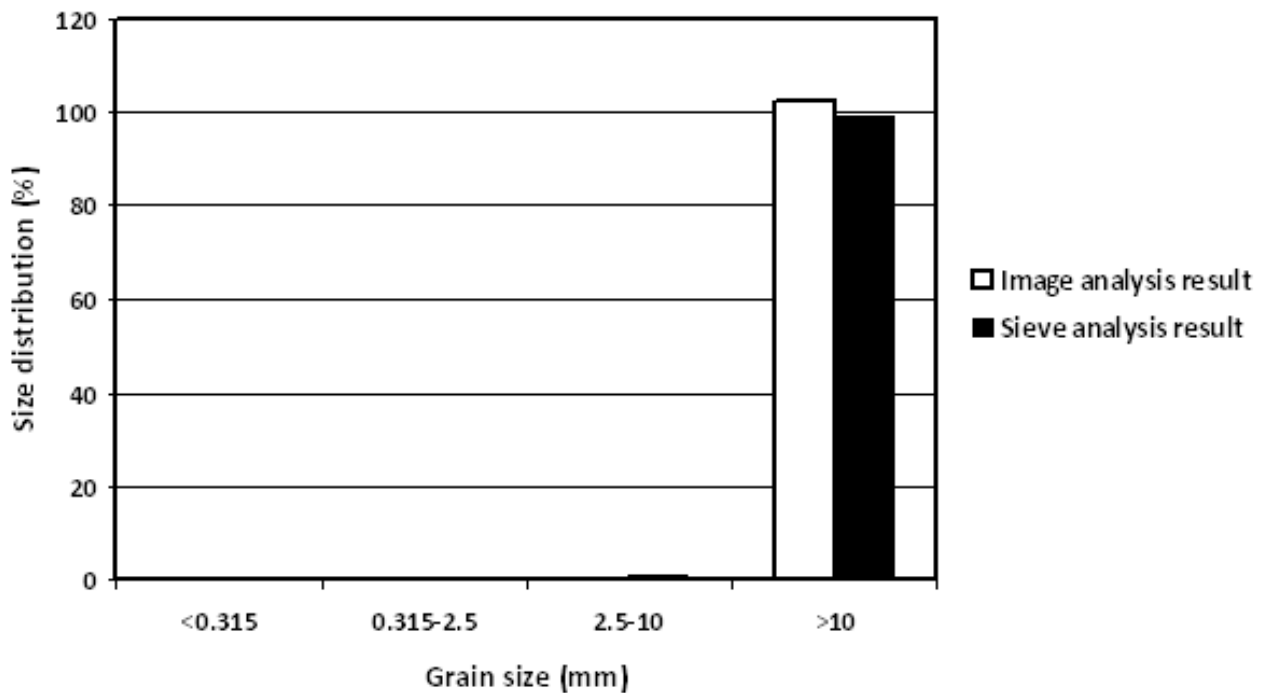


Figure 6.8: Comparison of actual and computed GSD for grains between 2.5 - >10 mm (class-6), images taken at a resolution of 600 dpi

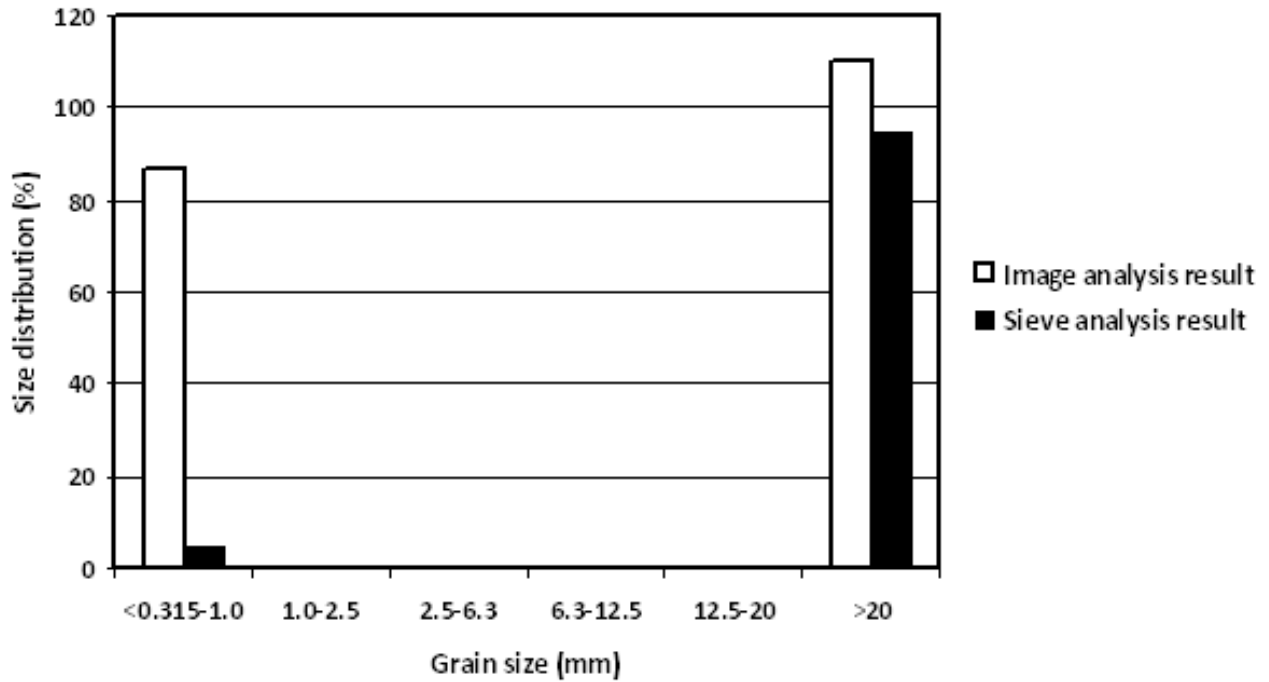


Figure 6.9: Comparison of actual and computed GSD for field image 1 in stage two (scanned calibration images). Very large error and very small correlation between the actual and computed results observed due to reasons explained in section 6.2

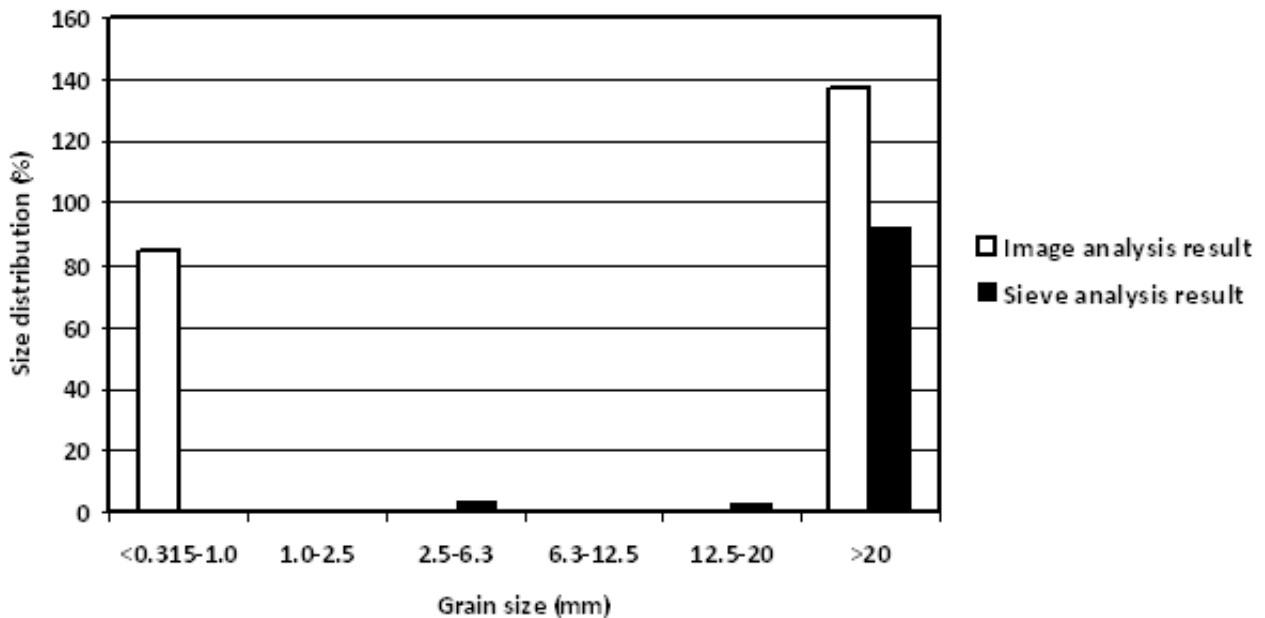


Figure 6.10: Comparison of actual and computed GSD for field image 2 in stage two (scanned calibration images). Very large error and very small correlation between the actual and computed results observed due to reasons explained in section 6.2

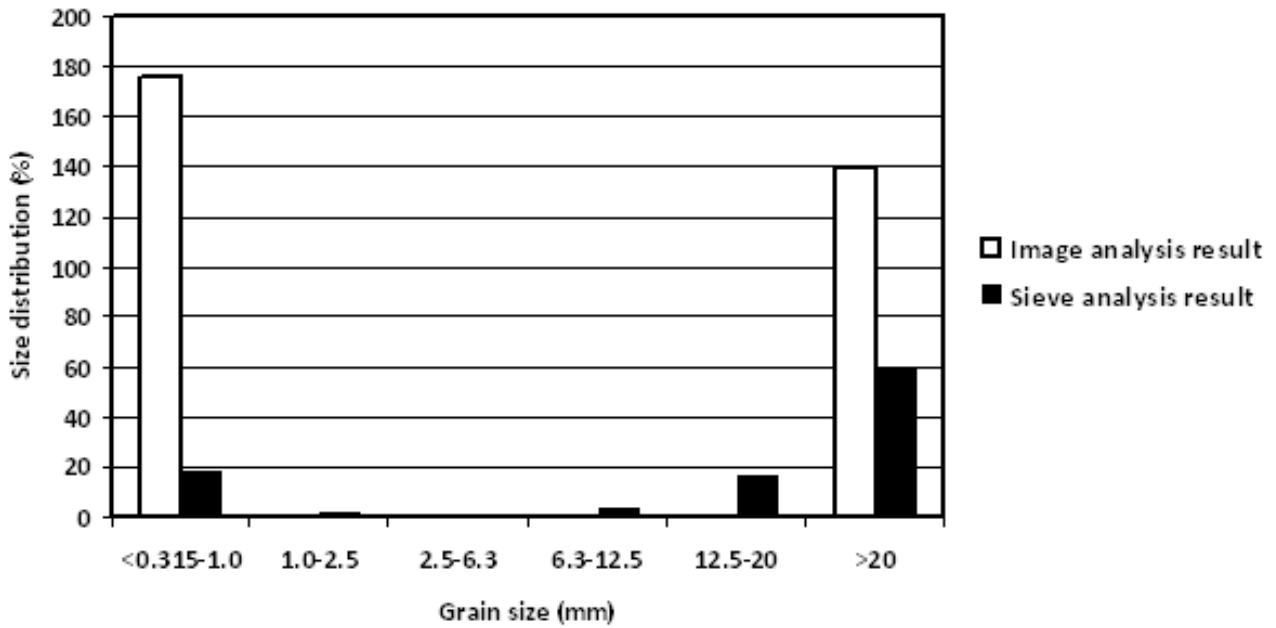


Figure 6.11: Comparison of actual and computed GSD for field image 3 in stage two (scanned calibration images). Very large error and very small correlation between the actual and computed results observed due to reasons explained in section 6.2

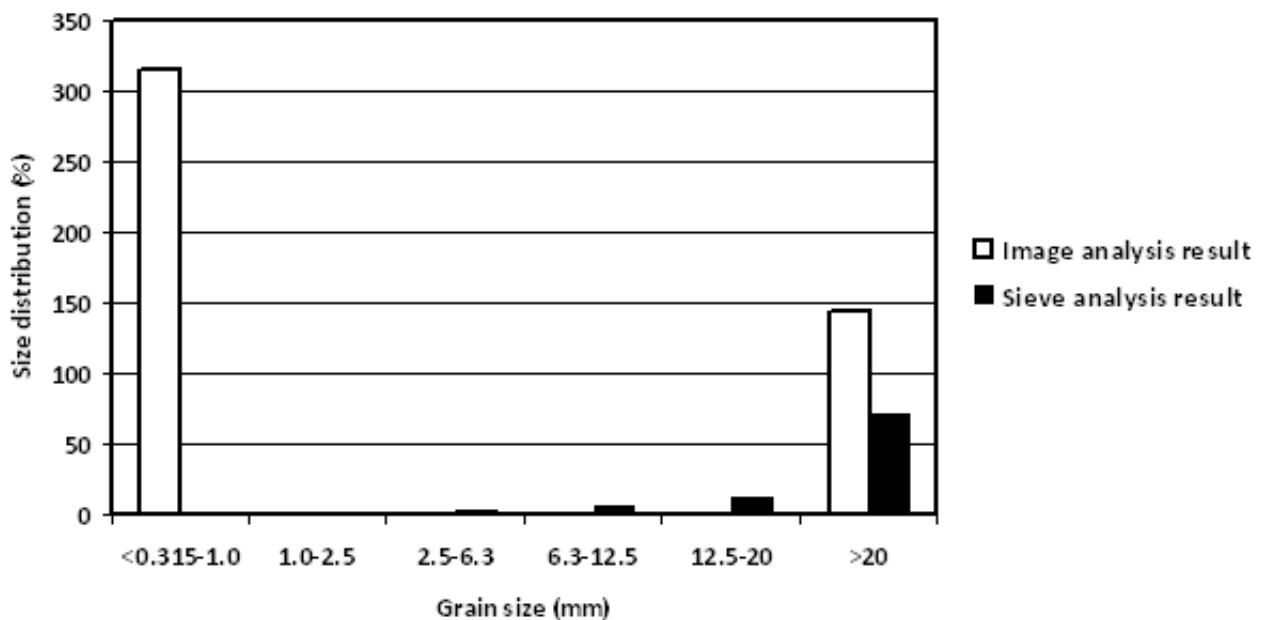


Figure 6.12: Comparison of actual and computed GSD for field image 4 in stage two (scanned calibration images). Very large error and very small correlation between the actual and computed results observed due to reasons explained in section 6.2

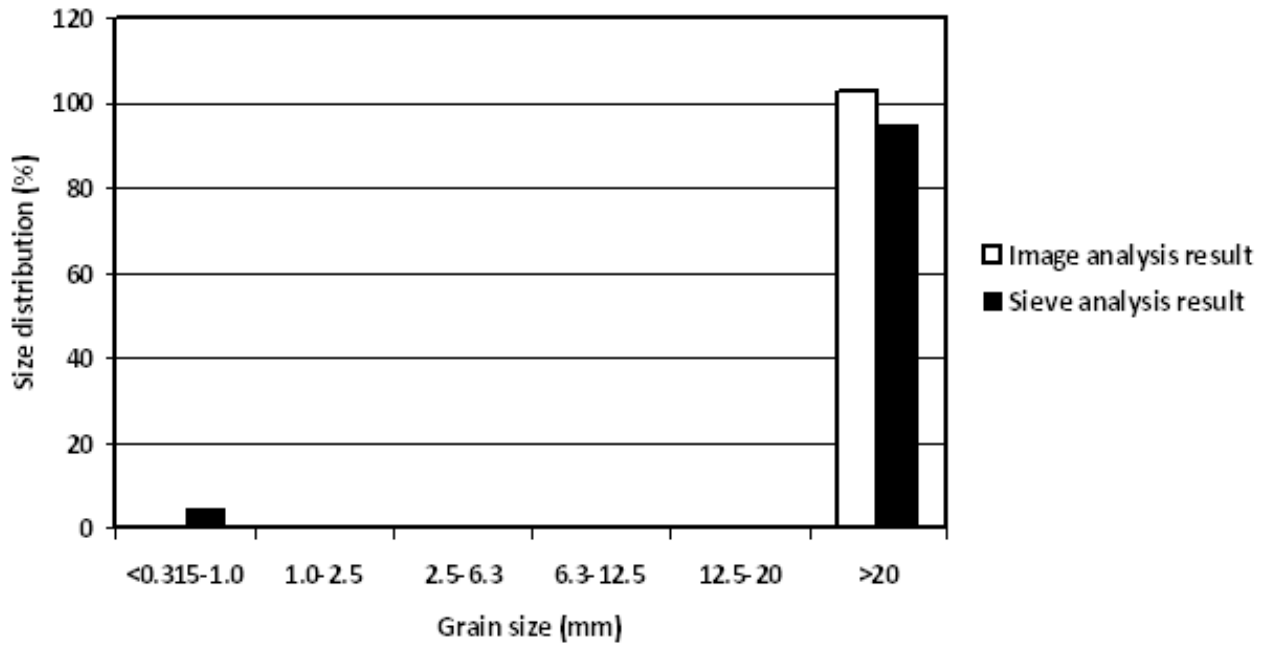


Figure 6.13: Comparison of actual and computed GSD for field image 1 using 40NFI calibration images

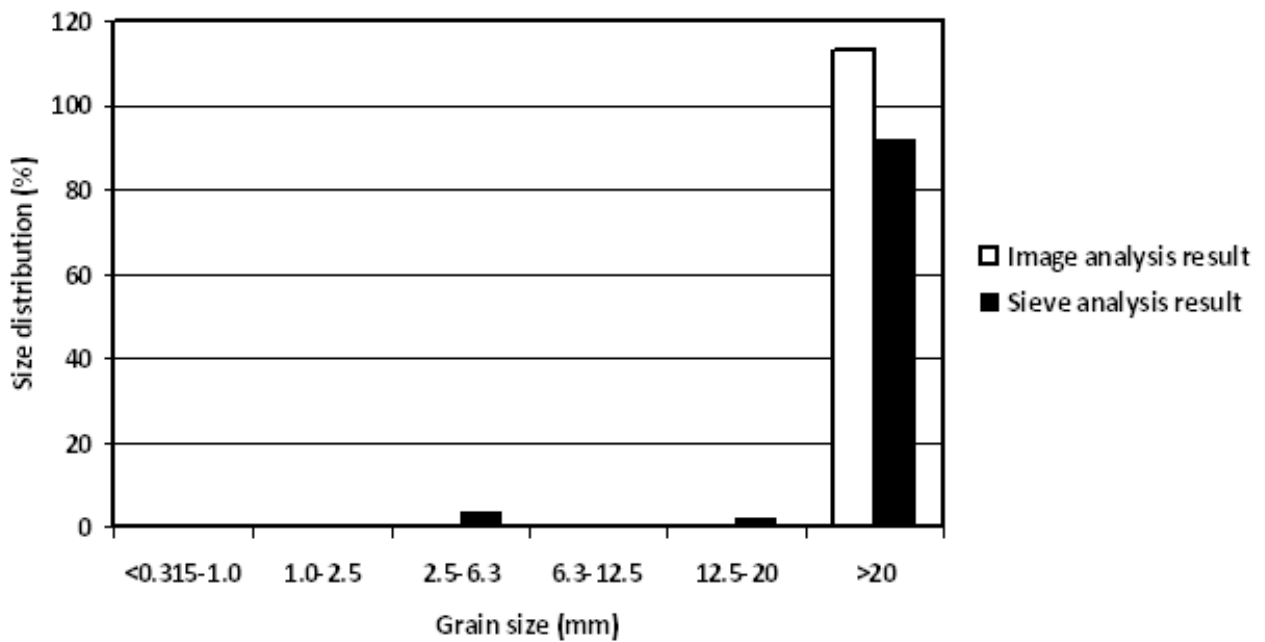


Figure 6.14: Comparison of actual and computed GSD for field image 2 using 30NFI calibration images

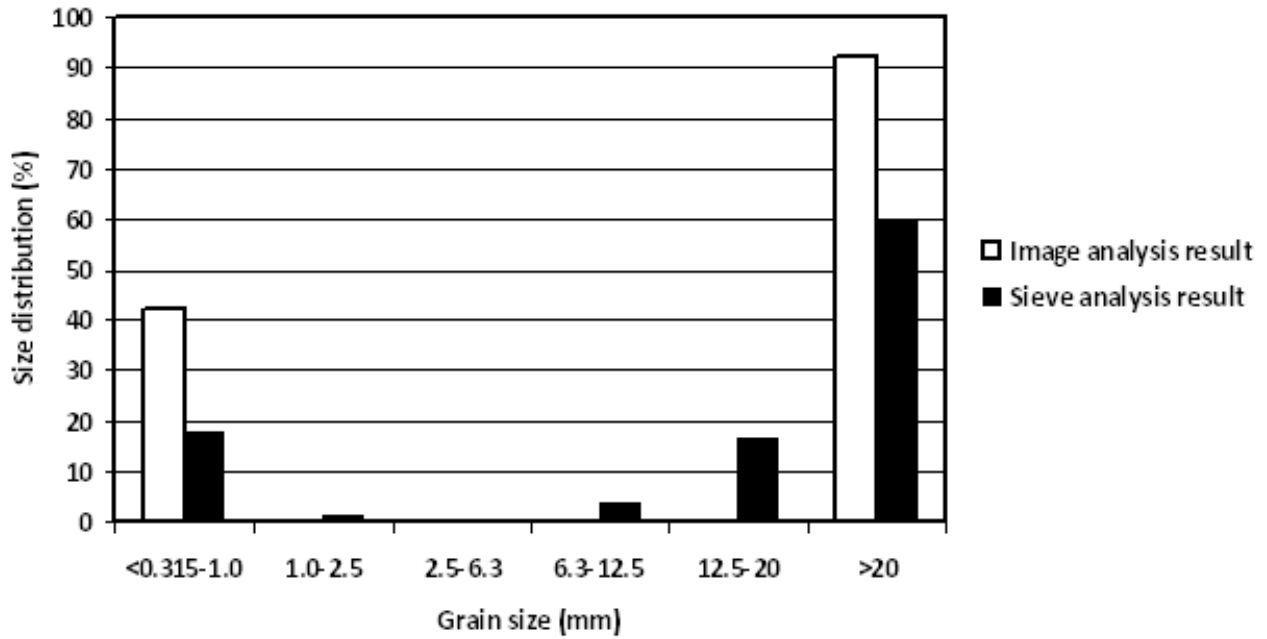


Figure 6.15: Comparison of actual and computed GSD for field image 3 using 30NFI calibration images. Very large error and very small correlation between the actual and computed results observed due to shadowed (underexposed) lighting conditions.

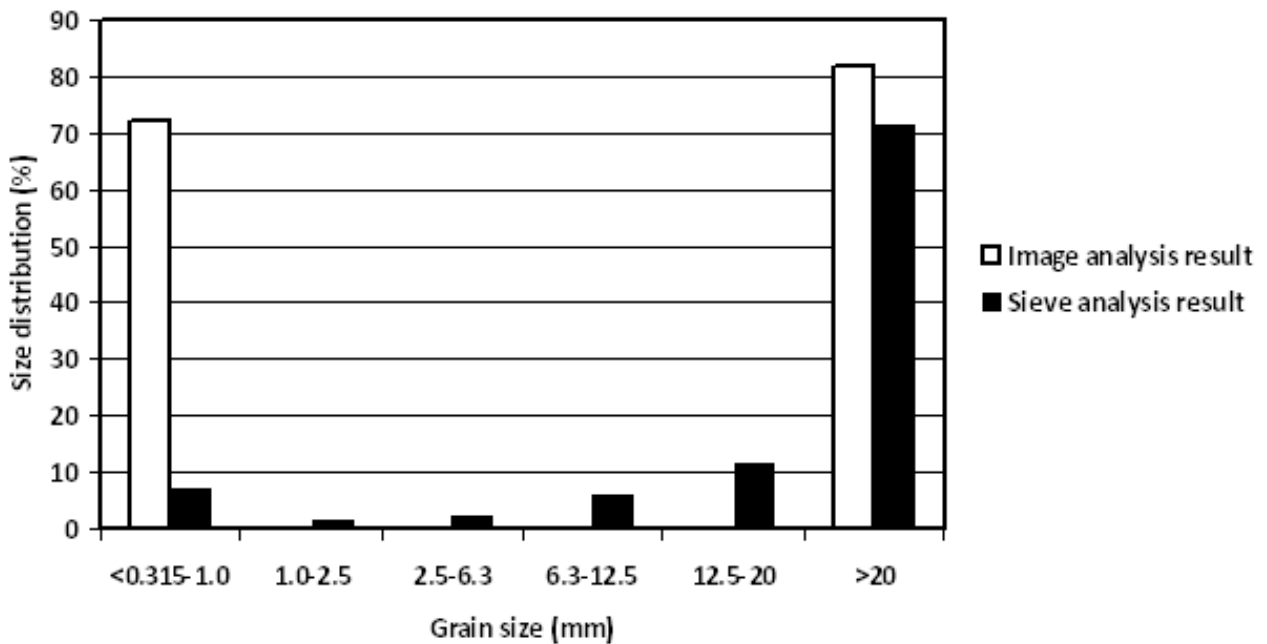


Figure 6.16: Comparison of actual and computed GSD for field image 4 using 30NFI calibration images. Very large error and very small correlation between the actual and computed results observed due to shadowed (underexposed) lighting conditions.

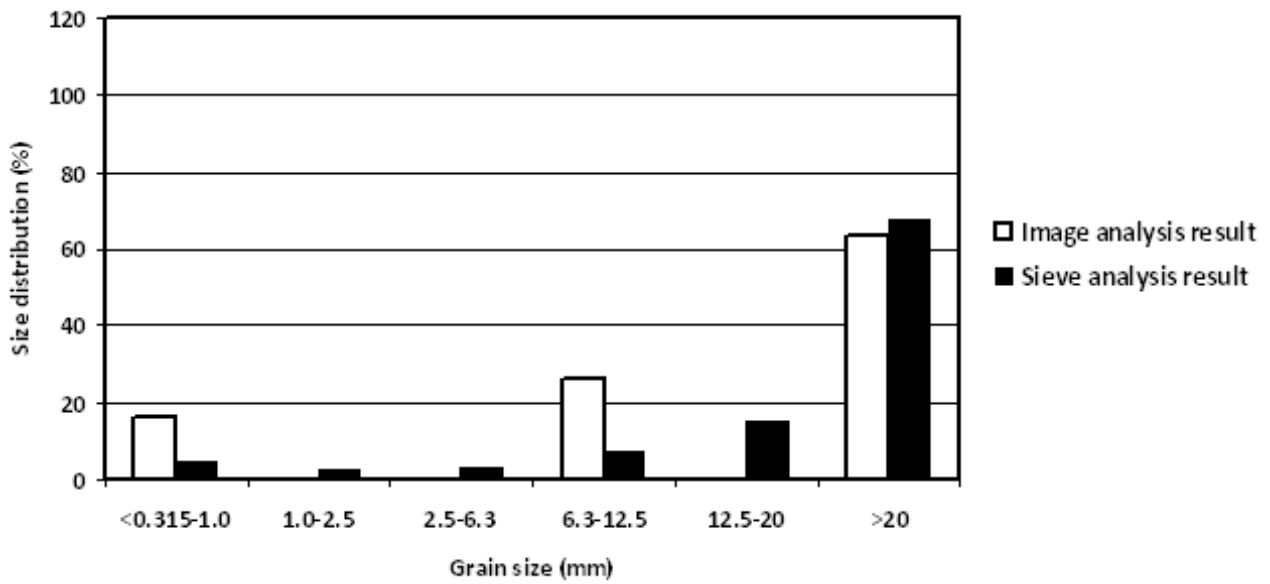


Figure 6.17: Comparison of actual and computed GSD for sample class-7 using 30NFI calibration images

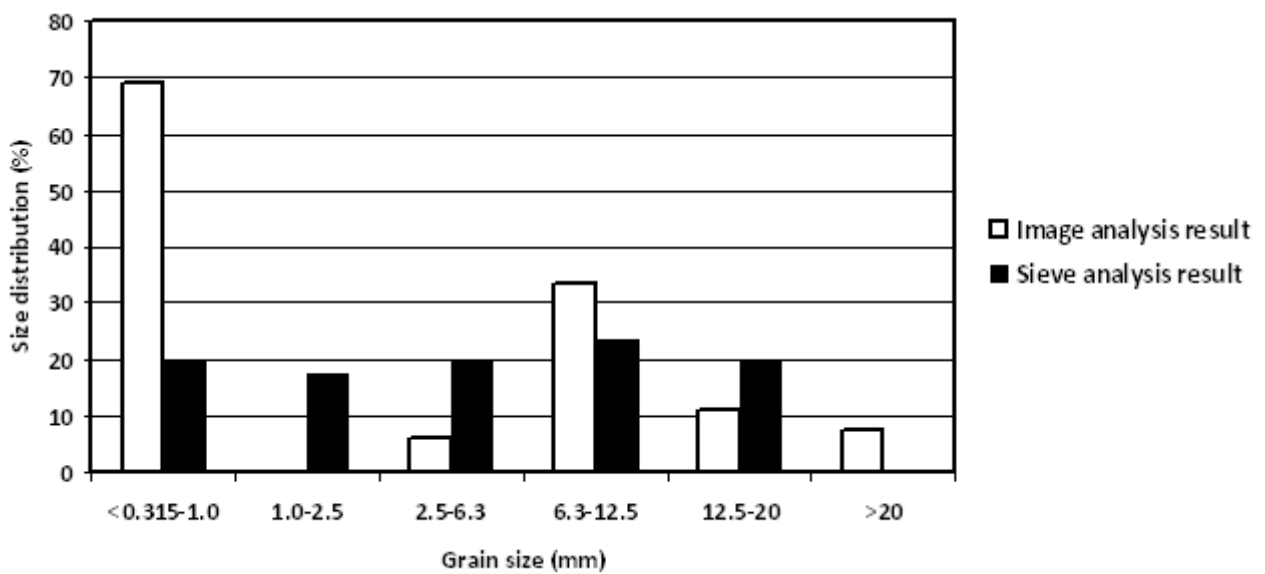


Figure 6.18: Comparison of actual and computed GSD for sample class-8 using 30FI calibration images

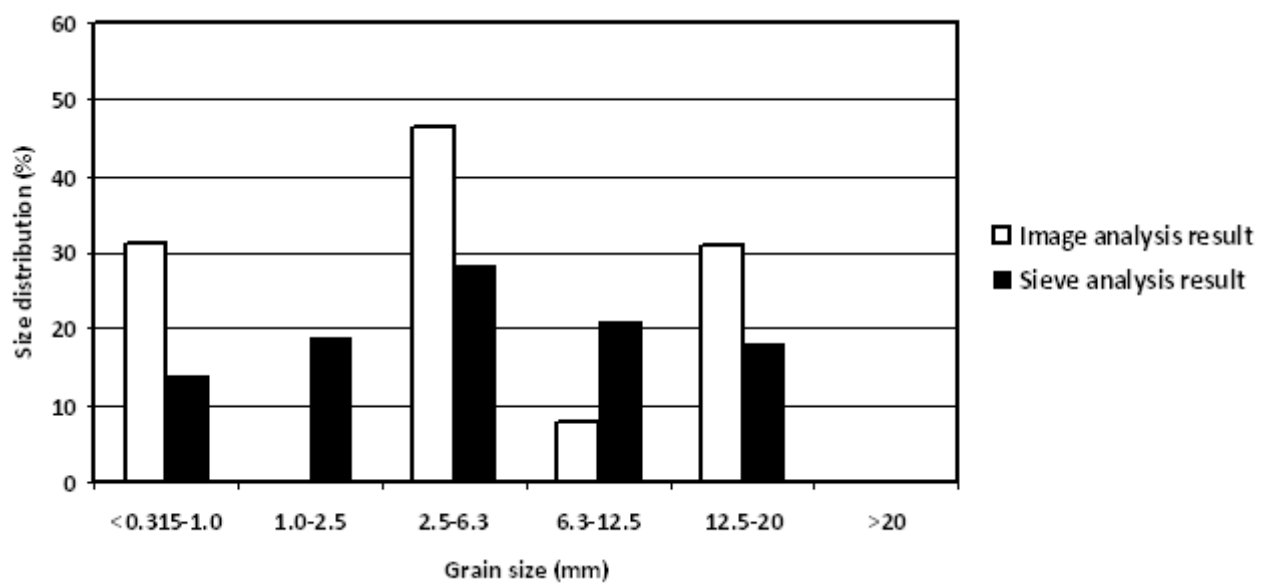


Figure 6.19: Comparison of actual and computed GSD for sample class-9 using 30NFI calibration images

## 7. Analysis and Discussion

### 7.1. Effect of Image Resolution on Calculated Grain-Size Distribution

It was observed during this study that images taken at higher resolutions give better grain-size distribution results than images taken at lower resolutions. Fig. 7.1 compares the results of soil samples investigated during the first stage of the experiments. It can be seen for all the six soil samples investigated, images taken at a resolution of 600 dpi predicted the best size distribution results, images at 300 dpi being the second best. The same can be observed from the results of the second stage experiments as shown in fig. 7.2. Calibration images taken at 600 dpi showed smaller errors than the 300 dpi images. This is because higher image resolution means higher number of pixel representation per grain which allows the algorithm to return accurate autocorrelation curves thereby predicting more accurate results.

### 7.2. Effect of Lighting Conditions and Camera Height on Calculated Grain-Size Distribution

Fig. 7.3 shows comparison of grain-size distribution of field samples predicted for different lighting conditions and camera heights from the ground. Calibration images taken at a camera height of 40 cm, with out flash and indoors gave the best results for field sample 1, while those images taken at a height of 30 cm, without flash and indoors gave the best results for samples 2, 3 and 4. This is because field sample 1 was taken at a camera height of 40 cm and samples 2, 3 and 4 at 30 cm.

---

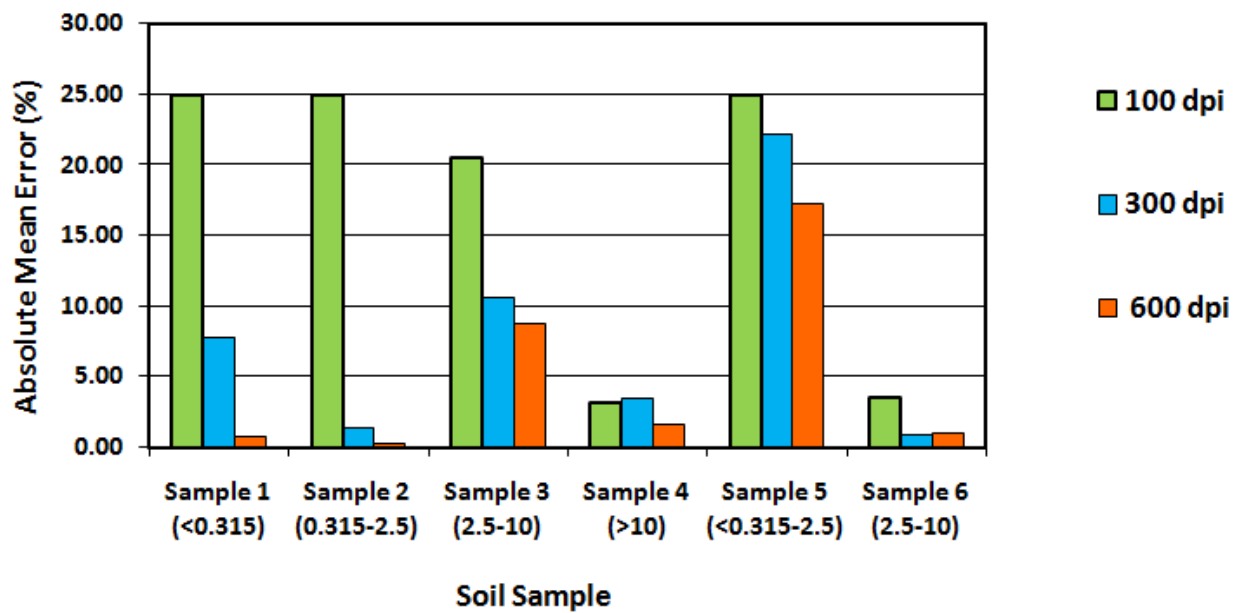


Figure 7.1: Comparison of scanned soil samples using errors for different image resolutions

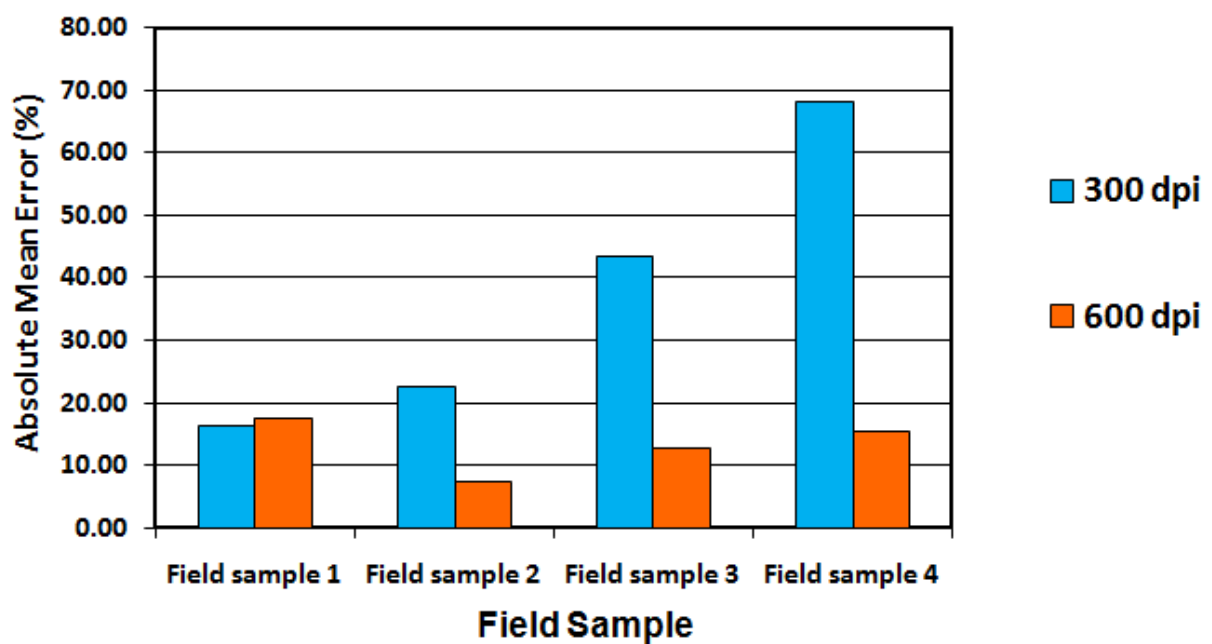


Figure 7.2: Comparison of field soil samples using errors for different correlation image resolutions

Field lighting condition was similar to indoors lighting condition on a typical summer day. It can be concluded from this finding that using similar lighting conditions and camera heights for both calibration and validation images give better size distribution predictions.

Furthermore, using flash when taking images seems to exaggerate the amount of fine grains. This might be due to small patches on coarse grains being interpreted as fine grains. Taking images of samples on a very sunny day might also have the same effect. It should, therefore, be avoided using flashes except in dark or highly shadowed lighting conditions.

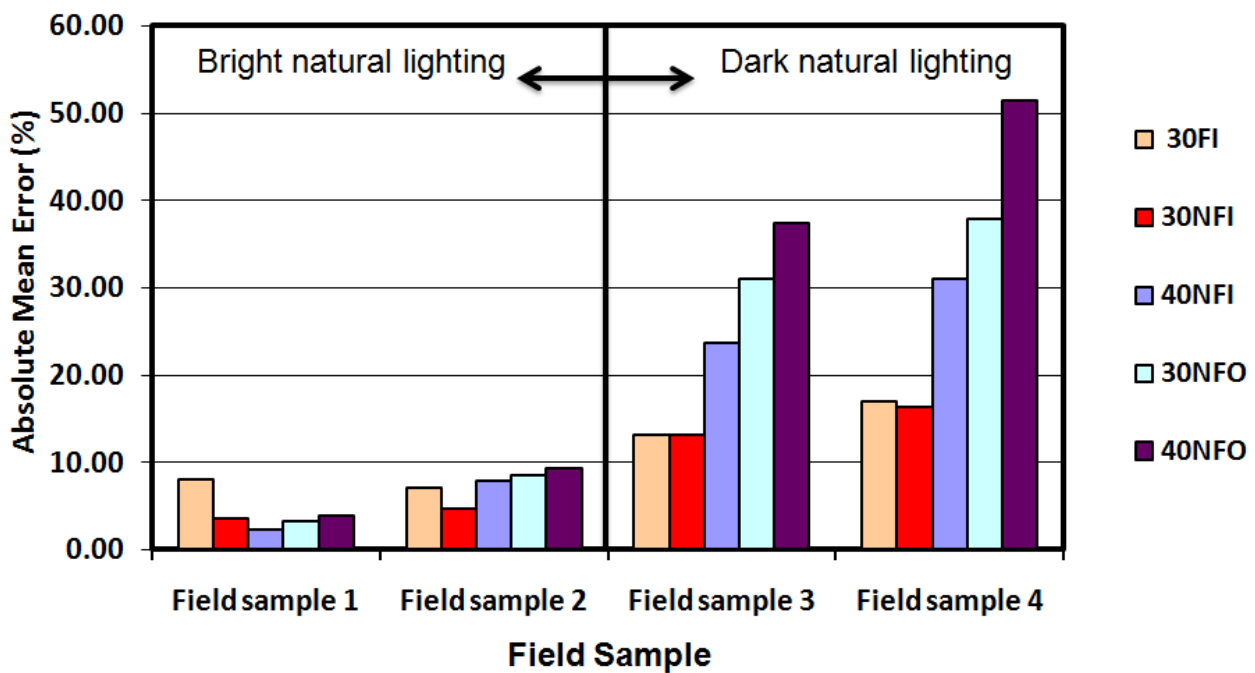


Figure 7.3: Comparison of field soil samples using errors for different lighting conditions and camera heights

### 7.3. Effect of Calibration Method on Calculated Grain-Size Distribution

Field samples taken during the second stage of the experiments were analyzed twice, first using scanned images as calibration and second using images taken in the laboratory using a digital camera. This was done to compare the effect of changing a calibration method on the size

distribution prediction. It can be shown from fig. 7.4 that results obtained using camera images for calibration showed smaller errors for samples 1 and 2, and the same errors for samples 3 and 4 when compared to the results obtained using scanned images for calibration. As explained in sections 7.2 and 6.2, similarity in the lighting conditions of the field and camera images resulted in better predictions of samples 1 and 2. On the other hand, relatively higher errors observed in samples 3 and 4 were due to the fact that both images were taken in a completely shadowed environment and resulted in a very low brightness and thus bad autocorrelation values.

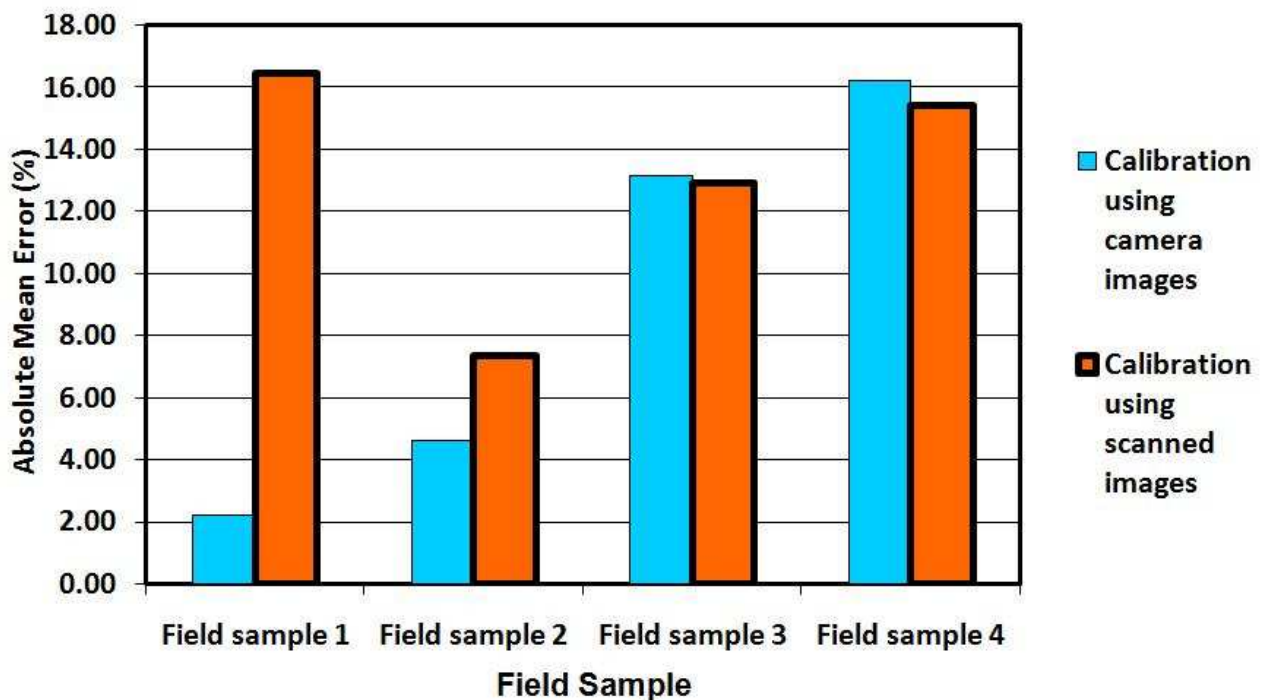


Figure 7.4: Comparison of field soil samples using errors for different Calibration Methods

## 7.4. Correlation Strength of Image and Sieve Analysis

### Results

Around 75% of the samples analyzed using image analysis showed strong correlation to the actual size distribution results obtained from sieve analysis. This means that 75% of the samples analyzed showed the same trend to their actual size distribution. It can be concluded from this observation

that the algorithms were able to recognize and predict the distributions of all size fractions, from fine to coarse grains, in most of the samples. Fig. 7.5 shows the results.

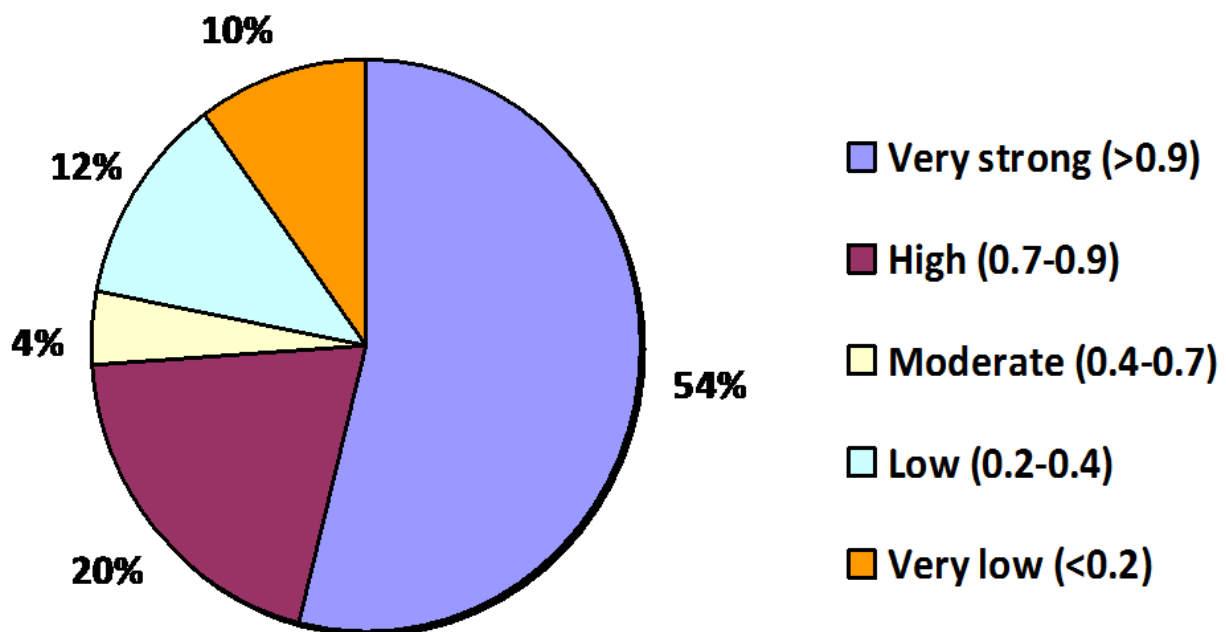


Figure 7.5: Overall correlation strength of image and sieve analysis results

## 7.5. Sources of Error

### 7.5.1. Errors Associated with Projected Area and Volume / Weight Comparison

The quantities of grain-size fractions in mechanical sieve analysis are determined by weighing and are expressed in terms of percentages by weight of the sample. In size-distribution from images on the other hand, a two-dimensional projected area of the grains is measured and not their weights. Image analysis techniques do not even consider the volume of the grains, while volume has an effect on the results of mechanical sieve analysis. Moreover, the sizes of grains in mechanical sieving are measured using the b-axis of the grains and whether they are passing through a square shaped sieve opening or not. The grain-size determined by image analysis might not be the same as that

determined by mechanical sieving. It is, therefore, difficult to compare the results of both methods and most of the errors encountered during this study are attributed to this effect.

### **7.5.2. Errors Associated with Sampling**

As explained in section 5.1 an 18.5 cm by 18.5 cm metal box, open on one end, was used to take soil samples in the field. As the sampler was small, it was difficult to collect the coarse grains and hand picking of grains had to be used. At the same time, because of its inability to penetrate and the firmness of the ground, the sampler was not able to collect all the fine particles. These limitations affected the results of the mechanical sieving and contributed to the errors observed during the second and third stages of the experiments.

### **7.5.3. Errors Associated with Calibration and Validation**

#### **Inconsistency**

In most of the experiments carried out during the second and third stages, image resolutions and lighting conditions used for both calibration and validation were different. Some times the calibration images being scanned and the validation being images taken with a camera, and at other times the calibration being images taken in the laboratory and the validation being images taken in the field. This inconsistency between the calibration and validation images might have influenced the results.

## **7.6. Image Collection**

Rubin (2004) recommends in his article that images should be taken by aiming the camera perpendicular to the ground and by keeping it at the same distance from the surface for both calibration and validation purposes. The camera should be configured (zoom, focus, and file format) in the same way for all the images taken. If artificial lighting, such as camera flash, is used, it should always be from a fixed angle. Images which are underexposed or overexposed to lighting should be avoided as much as possible as the grains are interpreted as large individual grains by the algorithm.

---

Moreover, it has to be made sure to include enough representative grains in terms of shape and texture in the calibration size fractions.

As already explained in previous chapters, image resolutions should be so that the smallest grain to be measured in the sample is represented by more than one pixels. Based on this, an equation relating the smallest grain-size in the sample, height of the camera from the ground, and camera resolution was developed for the particular camera used during this study. It is believed that such a simple relation could be developed for any digital camera used. The equation was developed such that the smallest grain-size in the sample is represented by a minimum of three pixels to achieve good results.

$$h = \frac{d\sqrt{C}}{2.46} \quad (7.1)$$

Where:

$h$  = camera height from the ground [mm]

$C$  = camera resolution [pixels]

$d$  = grain-size of the smallest grain of interest [mm]

Equation 7.1 was derived as follows:

It was first determined from experiments that at a height of 300 mm from the ground, the camera could fully capture an area of 150 x 200 mm on the ground. This fact was then used to calculate the angles of inclination of the rays of the camera ( $\theta_1$  and  $\theta_2$ ) on both sides of the rectangular area as shown in figures 7.6 and 7.7.

$$\tan \theta_1 = \frac{x}{h} = \frac{150mm}{300mm} = 0.5 \rightarrow x = 0.5h \quad (7.2)$$

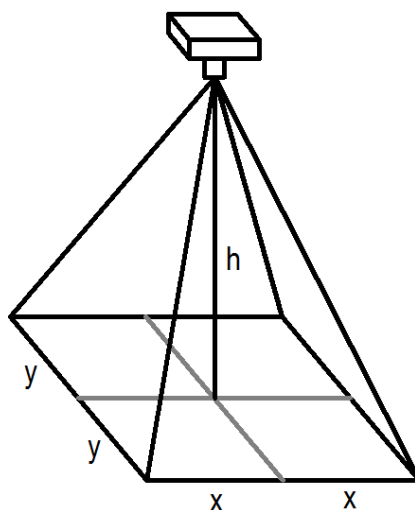


Figure 7.6: A  $2x$  by  $2y$  ground area captured by placing the camera at a height of  $h$  mm from the ground

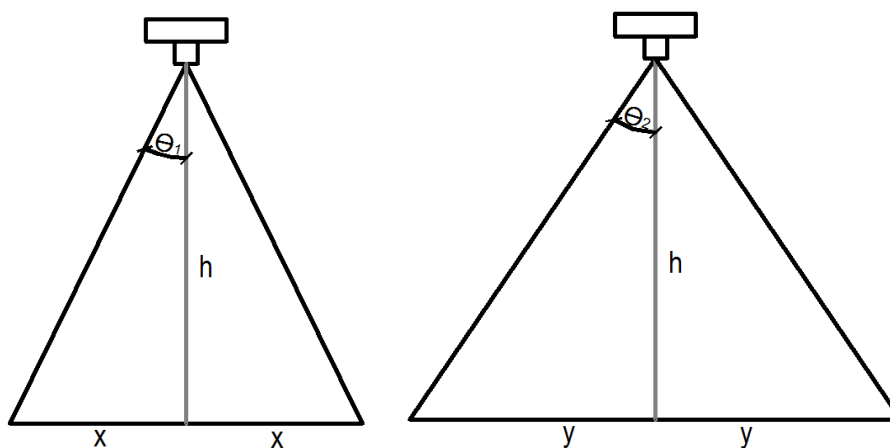


Figure 7.7: Side views of the captured ground area

$$\tan \theta_2 = \frac{y}{h} = \frac{200\text{mm}}{300\text{mm}} = 0.67 \rightarrow y = 0.67h \quad (7.3)$$

Area of the ground  $A$ :

$$A = 2x \cdot 2y = 2(0.5h \cdot 0.67h) = 0.67h^2 \quad (7.4)$$

The minimum number of pixels needed to represent the smallest grain-size in the sample was

decided to be 3 pixels. From this, the following relation could be developed:

$$\frac{C}{A} = \left(\frac{3}{d}\right)^2 \rightarrow \frac{C}{0.67h^2} = \left(\frac{3}{d}\right)^2 \quad (7.5)$$

Solving for  $h$  in equation 7.5 results in the final equation:

$$h = \frac{d\sqrt{C}}{2.46} \quad (7.6)$$

Finally, based on equation 7.1 a graph (fig. 7.8) was developed showing camera height versus smallest grain-size measurable for a range of camera resolutions.

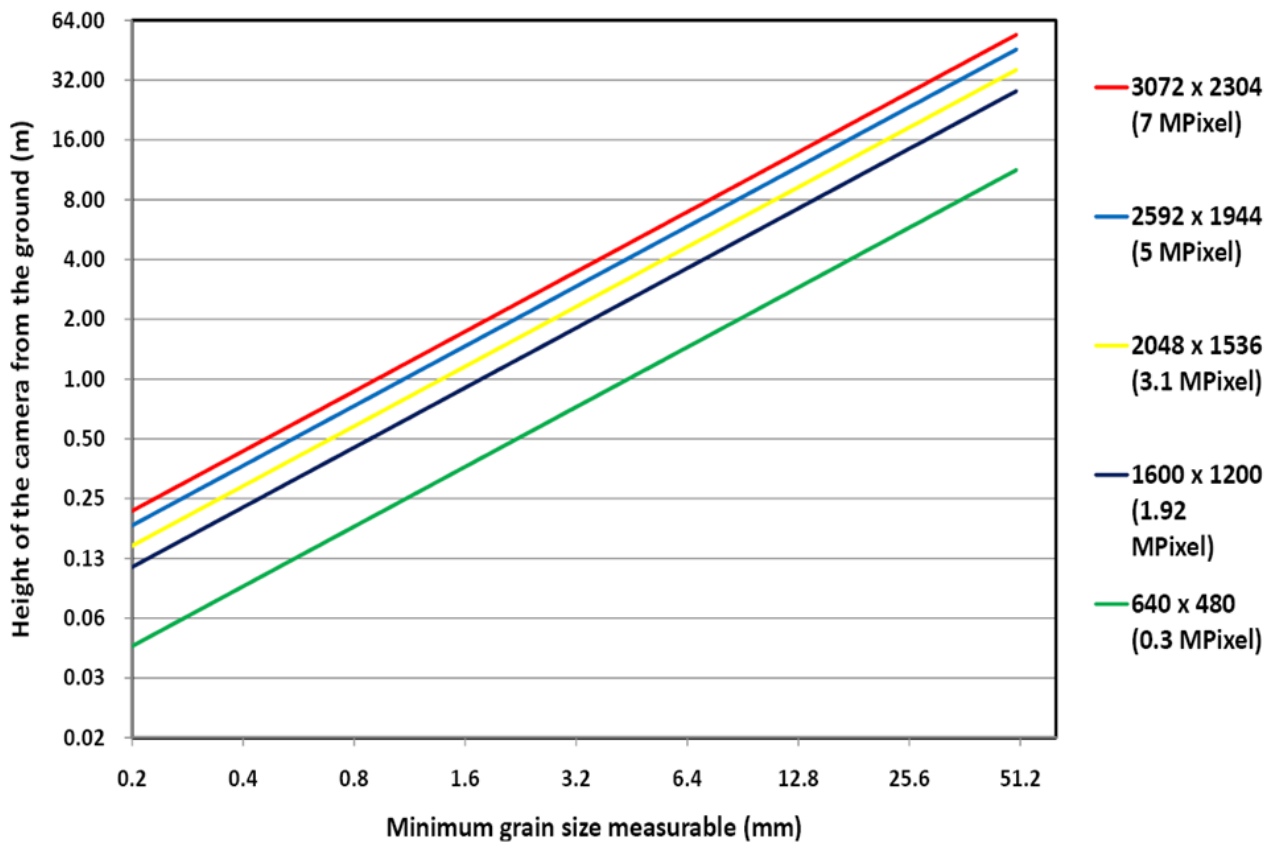


Figure 7.8: Graph showing the relation between camera height and smallest grain-size of interest for different camera resolutions

## 8. Conclusions and Recommendations

### 8.0.1. Conclusions

Based on the objectives of this thesis and the works carried out during the study period, the following conclusions can be drawn:

- Previously developed methods of grain-size distribution determination from digital images are mostly based on edge detection of individual grains. These methods suffered from a limitation that overlapping and touching grains are interpreted as one big grain.
  - The method identified and used during this thesis, autocorrelation analysis of digital images, is a statistical method and overcomes the problem of touching and overlapping grains.
  - The method of size distribution from autocorrelation analysis of digital images was successful in predicting size distribution of a diverse range of sediment types: from sand ( $<0.315$  mm) to coarse gravel ( $>20$  mm).
  - It was also shown that the method of autocorrelation analysis is very simple to use with out specialist equipment or technical knowledge.
  - Run time of the algorithms of the autocorrelation analysis method range form  $<1$  min to 5 min depending on grain size, image resolution and maximum offset distance. This means that the method is faster than mechanical sieving by more than 100%.
  - An overall error, from the results of mechanical sieving, of around 10% was observed. This error is acceptable for the purposes of preparing fluvial substrate maps for habitat modeling. Furthermore, taking into account that the compared values of size distribution were based
-

on a projected 2D area for the image analysis and on weight and 3D shape for the mechanical sieve analysis, the observed error is acceptable.

- 75% of the samples analyzed using the method of autocorrelation analysis showed strong correlation to mechanical sieve analysis results.
- A grain should be represented by 2 or more pixels to have good autocorrelation results.
- The higher the resolution of the images, the better are the size distribution results.
- It was observed that similar calibration and validation conditions (lighting and image resolution) give better results.
- Use of flash seems to exaggerate the amount of fine grains.
- Images underexposed or overexposed to lighting return bad autocorrelation curves and as a result bad size distribution predictions.

### 8.0.2. Recommendations and Future Work

The most important fact identified during this thesis is that **the accuracy of the results of size distribution from autocorrelation analysis of digital images highly depends on the consistency of calibration and validation images with regards to their resolution and lighting conditions**. Therefore, it is recommended that all the images be taken at similar resolution and lighting conditions. Use of flash should be avoided when taking images except in highly shadowed and dark conditions. Furthermore, it is recommended that images which are underexposed or overexposed to lighting should not be used for autocorrelation analysis.

Future work on this method should focus on finding a correction factor or a method for converting the results from this method to mechanical sieve analysis results. Future research should also focus on a technique of using image analysis results in preparing fluvial substrate maps.

---

## A. Appendix A

The MATLAB algorithms used in this study are presented here.

---

## B. Appendix B

The results of all the experiments carried out during the study period of this master's thesis are presented here.

The headings used on each page of the following pages can be interpreted as follows:

Example: Stage1 - scanned calibration/validation images - sample class 1 (0.315 - 2.5 mm) - 100 dpi resolution

This means that the experiment was carried out during stage1, scanned images were used for both calibration and validation purposes, name of the sample is class1 and includes grains between 0.315 mm and 2.5 mm, and 100 dpi resolution images were used.

---

# Bibliography

## Berg 1939

BERG, Soren: *Studies on Particle Size Distribution*. Academy of Technical Sciences, Denmark, 1939

## Bridge 1981

BRIDGE, John S.: Hydraulic Interpretation of Grain Size Distribution Using a Physical Model for Bedload Transport. In: *Journal of Sedimentary Petrology* 51 (1981), S. 1109–1124

## Bunte 2001

BUNTE, Kristin: Sampling Surface and Sub-surface Particle Size Distribution in Wadable Gravel- and Cobble-Bed Streams for Analyses in Sediment Transport, and Streambed Monitoring / United States Department of Agriculture - ForestService. 2001. – Forschungsbericht

## Carbonneau 2004

CARBONNEAU, Patrice E.: Catchment-scale mapping of surface grain size in gravel bed rivers using airborne digital imagery. In: *Water Resources Research* W07202 (2004), S. W07202

## Cheel 2005

CHEEL, R.J.: Standard procedure for sieve analysis of sand. Version: 2005. <http://spartan.ac.brocku.ca/~rcheel/teaching/sedimentology/SedNotes/>. In: *Introduction to Clastic Sedimentology*. Department of Earth Sciences, Brock University, Kapitel Appendix 1. – Online-Ressource

## Collins 2007

---

COLLINS, M.: Stream barrier removal monitoring guide. Gulf of Maine Council on the Marine Environment. Version: 2007. [www.gulfofmaine.org/streambarrierremoval](http://www.gulfofmaine.org/streambarrierremoval). Gulf of Maine Council on the Marine Environment. – Online-Ressource

#### **DIN-18123 1996**

DIN-18123: Bestimmung der Korngrößenverteilung - DIN 18123 / Deutsches Institut für Normung. 1996. – Forschungsbericht

#### **Eberstaller 2006**

EBERSTALLER, Jürgen: Sustainable Sediment Management in Alpine Reservoirs Considering Ecological and Economical Aspects / Institut für Wasserwesen - Universität der Bundeswehr München. 2006. – Forschungsbericht

#### **Fisher 2004**

FISHER, Robert: Hypermedia Image Processing Reference. Version: 2004. [http://homepages.inf.ed.ac.uk/rbf/HIPR2/hipr\\_top.htm](http://homepages.inf.ed.ac.uk/rbf/HIPR2/hipr_top.htm). John Wiley & Sons. – Online-Ressource

#### **Ghalib 1999**

GHALIB, Ali: Soil Particle Size Distribution by Mosaic Imaging and Watershed Analysis. In: *Journal of Computing in Civil Engineering* 13 (1999), S. 80 – 87

#### **Graham 2005**

GRAHAM, D. J.: A Transferable Method for The Automated Grain Sizing of River Gravels. In: *Water Resources Research* 41 (2005), S. 1 – 12

#### **Kondolf 2003**

KONDOLF, Mathias: *Tools in Fluvial Geomorphology*. John Wiley and Sons, 2003

#### **Latulippe 2001**

LATULIPPE, Christian: Visual Characterization Technique for Gravel-Cobble River Bed Surface Sediments; Validation and Environmental Applications. In: *Earth Surface Processes and Landforms* 26 (2001), S. 307–318

---

**Lembo 2000**

LEMBO, Arthur J.: Spatial Autocorrelation. Version: 2000. [www.css.cornell.edu/courses/620/lecture9.pp](http://www.css.cornell.edu/courses/620/lecture9.pp). Cornell University. – Online-Ressource

**McLaren 1981**

MCLAREN, Patrick: An Interpretation of Trends in Grain Size Measures. In: *Journal of Sedimentary Petrology* 51 (1981), S. 0611–0624

**NRC 2008**

NRC, Natural Resources C.: Multivariate Statistics - Spatial Autocorrelation. Version: 2008. [http://www.pfc.forestry.ca/profiles/wulder/mvstats/spatial\\_e.html](http://www.pfc.forestry.ca/profiles/wulder/mvstats/spatial_e.html). Natural Resources Canada. – Online-Ressource

**Platts 1983**

PLATTS, W.S.: Methods for Evaluating Stream, Riparian, and Biotic Conditions / USDA, Forest Service, Intermountain Forest and Range Experiment Station, Ogden UT. 1983. – Forschungsbericht

**Poppe 2001**

POPPE, Lawrence J.: Map Showing the Distribution of Surficial Sediments in Long Island Sound. Version: 2001. <http://pubs.usgs.gov/of/2000/of00-304/htmldocs/chap04/index.htm>. U.S. Geological Survey (USGS). – Online-Ressource

**Retsch 2008**

RETSCH, Solutions in Milling & S.: Fine Tuning Sieve Analysis for Accurate Particle Size. Version: 2008. <http://www.retsch.com/uploads/s1/c6b16ea74b1d95baa522f6d5754a9fc0.jpg>. – Online-Ressource

**Rubin 2004**

RUBIN, David M.: A Simple Autocorrelation Algorithm for Determining Grain Size from Digital Images of Sediment. In: *Journal of Sedimentary Research* 74 (2004), S. 160 – 165

---

**Sahu 1965**

SAHU, Basanta K.: Theory of Sieving. In: *Journal of Sedimentary Research* 35 (1965), S. p. 750–753

**Shin 2004**

SHIN, Seungcheol: Wavelet Analysis of Soil Mass Images for Particle Size Determination. In: *Journal of Computing in Civil Engineering* (2004), S. 19–27

**Sime 2003**

SIME, L. C.: Information on Grain Size in Gravel-Bed Rivers by Automated Image Analysis. In: *Journal of Sedimentary Research* 73 (2003), S. 630 – 636

**Syvitski 1991**

SYVITSKI, James P.: *Principles, Methods, and Application of Particle Size Analysis*. Cambridge University Press, 1991

**Tobler 1970**

TOBLER, W. R.: A computer movie simulating urban growth in the Detroit region. In: *Economic Geography* 46 (1970), S. 234 – 240

**USGS 2001**

USGS: *USGS EAST-COAST SEDIMENT ANALYSIS: PROCEDURES, DATABASE, AND GEO-REFERENCED DISPLAYS*. U.S. GEOLOGICAL SURVEY (USGS), 2001

---