

Nataša Sladoje* and Joakim Lindblad*

THE COVERAGE MODEL AND ITS USE IN IMAGE PROCESSING

Abstract. The coverage model provides a framework for representing continuous objects present in digital images as spatial fuzzy subsets. Assigned membership values indicate to what extent image elements are covered by the imaged objects. We present the basic definitions and properties of this model and show how it can be used to improve information extraction from digital images and to reduce problems originating from limited spatial resolution. We describe a number of image segmentation methods that result in coverage representations. We present methods for estimating geometric moments and object perimeter from coverage representations and derive the corresponding maximal estimation errors as functions of sampling density and number of quantization levels. Compared to a classic binary approach the coverage model provides greatly increased precision. We show how to generate an appropriate binary representation from a coverage one, and also how to use the information rich coverage representation to reconstruct a binary representation at an increased resolution. Empirical studies as well as presented image analysis applications demonstrate the practical advantages of the coverage model and the superior performance of the described methods.

Mathematics Subject Classification (2010): 68U10, 65D18

Keywords: digital image analysis, object representation, image segmentation, feature extraction, error estimation, fuzzy sets

**Faculty of Technical Sciences, University of Novi Sad, Serbia*

CONTENTS

1. Introduction	40
2. Background and related work	45
2.1. Grey-scale information	46
2.2. Remote sensing–fractional pixels	47
2.3. Tomographic images–partial volume effect	49
2.4. Fuzzy set theory in image processing	50
3. The coverage model	52
3.1. Basic definitions	53
3.2. Properties of coverage digitizations	55
4. Coverage segmentation methods	57
4.1. Coverage segmentation based on double thresholding	58
4.2. Coverage segmentation by local unmixing	59
4.3. Graph based coverage segmentation	63
5. Feature extraction	68
5.1. Shape analysis	69
5.2. Geometric moments	70
5.3. Perimeter	78
6. Defuzzification and high resolution reconstruction	87
7. Performance of the presented approach	98
7.1. Comparison of different perimeter estimation methods	99
7.2. Coverage segmentation followed by feature estimation for noisy data	99
7.3. Estimation of affine deformations of shapes	104
7.4. High resolution reconstruction–two examples	108
8. Conclusions and future work	112
References	114

1. Introduction

The interdisciplinary field of imaging science, including image processing, image analysis, image understanding and visualization, is undergoing a very rapid development. Closely tied to advancements in technology, digital imaging and digital image processing, have grown to, not only become very important parts of the scientific world of information processing, but also to become important parts of our everyday lives, in the sense that imaging and image processing are integrated in both society and science.

Together with the development of science and technology, imaging has become a rather general term, related not only to, as traditionally understood, capturing of light reflected of a two-dimensional (2D) surface, but also the measuring many other physical properties of objects of interest, in two or more dimensions. Such different types of imaging, naturally create rather different types of images. These images have in common the acquisition of, in some sense, spatially distributed measurements and they, in general, constitute a valuable source of information about the observed objects. “Visible” is not always in focus any longer; the challenge became to capture images of objects we cannot see (distant stars, atomic-size objects, parts of a living human body, unborn babies, blood flow), or images of unconventional properties of objects (heat, density, water content, etc). To be able to understand and interpret such images, the observer has to know what physical property is expressed and how that property relates to the intensity levels expressed in the image.

Essentially all imaging techniques provide some kind of geometric information about the object: some provide information about anatomy and/or function (e.g., magnetic resonance angiography–MRA, positron emission tomography–PET), some show topographic properties of an object (radar, ultrasound), others may provide very detailed spectral or temporal information (hyperspectral or high speed cameras). The acquired data are, in general, organized in a way that preserves some spatial structure of the object of interest, even if that is not necessarily the main observed property; this analogy with the traditional concept of an image is why the process of creating such data structures is called imaging, and the data themselves–images.

Digital images. In most application areas of imaging sciences, information about some objects of interest, captured in images, needs to be extracted, visualized, manipulated and analysed. When addressing such tasks, we more and more rely on the power of computers. Computers can handle huge amounts of data and accomplish many tasks, primarily those defined in terms of processing large sets of numerical values, much faster, and more reliably than humans can. Connected with digital computers inability to represent continuous information, the imaging process is generally assumed to, instead of capturing a continuum of an observed piece of space containing objects of interest, only observe a sample of points. In such a way, the image domain is discretized and mapped onto a discrete set of points.

The image sample points are often regularly distributed in a grid, and are, in many cases, addressed by integer coordinates. Each such grid point, in some sense, represents a portion of the observed continuous space (often a Voronoi region of the point, i.e., the part of the image space which is closer to the observed grid point than to any other grid point). Every image element (called pixel in 2D, voxel in 3D, or spel–spatial element–in the general case) is, in the imaging process, assigned a value corresponding to the intensity of the physical property observed in that piece of continuous space.

Despite us referring to the “value” or “intensity” of an image element, the image function may not be a scalar function and these words may well refer to a vector

of values instead. For example, in a typical colour image, every image point is assigned, not one, but three intensities: red, green, and blue, or RGB for short.

The observed physical property can essentially never be measured perfectly and without error. In addition, the obtained values are stored in limited memory space in a computer. Therefore, the range of the image function is usually restricted to a set of integers (or fractions with common denominator). This process is called quantization. Discretization (sampling) and quantization, applied together to an, initially continuous, image function (theoretically continuous), lead to what is called a digital image, where both the domain and the range are discrete and limited. The number of sample points per unit (density of sampling) is often referred to as image resolution (spatial, spectral, time), where higher image resolution in general provides more information about the imaged objects, and most often better subsequent analysis results. (It should be noted that this is not really a strict usage of the notion of *resolution*, since it does not say what we actually can resolve in the image; the latter is dependent on the physics of the imaging device and not on the number of pixels in an image.) Unfortunately, to increase image resolution deliberately is seldom possible; resolution is imposed by the imaging conditions. Therefore, a challenge within the field of image analysis is to develop creative methods that are capable to utilize and extract as much as possible from the data that is available. Our work summarized in this paper, is in line with this challenge of overcoming limitations of a given spatial resolution and to increase the quality of image analysis results by utilizing the available information as well as possible.

Segmentation. No matter what physical property is imaged, it is practically never exhibited so that it creates well defined and homogeneous regions. Imprecision is a result of imaging conditions, like noise or limited resolution, but also of the properties of the imaged objects. This makes it difficult to clearly separate and outline different objects appearing in the image. Image segmentation aims at defining the extents of the different objects in the image by partitioning the image into a number of regions characterized by a certain intra-component homogeneity and inter-component discontinuity. This is generally considered to be both the most important and the most challenging task in image processing. A decision if a point belongs to the object of interest, or not, is crucial for the quality of all following analysis steps and is often very difficult to make.

In an ideal case, a one-to-one correspondence between the set of image intensities and the set of image components exists and a partitioning can be based on a straightforward classification of pixel intensities. However, even in such an ideal case, discretization of the continuous image space leads to ambiguous situations where one pixel may be partly covered by more than one object in the image. The intensity assigned to such a pixel is a mixture of the intensities associated with the corresponding “pure” components.

However, segmentation is traditionally performed in a crisp way, where each image element is given only one label, i.e., a pixel is completely associated to one single image component. This type of crisp segmentation does not allow partial belongingness of a pixel to an object, and a hard decision of the belongingness has

to be made. Intuitively, some kind of thresholding is applied and the crisp classification of a “mixed” pixel as belonging to only one of the image components is performed. In less ideal and more general cases, the presence of noise excludes the possibility of a straightforward classification based on individual pixel intensities only, and more complex, sophisticated, and task dependent segmentation methods are applied, often utilizing spatial information and/or some type of a priori knowledge in addition. Different segmentation methods deal with (different types of) noise in different, more or less successful, ways; however, the issue of mixed border pixels remains, being caused by discretization itself.

Even for the simplest case, where the image only contains one object and the segmentation task reduces to that of defining what is foreground (object) and what is background, it starts to be clear that a segmentation which leads to a binary (two-valued) image as a result, where object points are mapped to one, or “white”, and nonobject points (background) to zero, or “black”, cannot handle uncertainties and heterogeneity of object properties very well. Despite the ability of the human visual system to provide an intuitive perception of an object as a whole, also in the presence of vague borders and “variability” in the image, it is observed when looking at small regions of an image, that humans can no longer make clear statements whether elements belong to an object or not. Our perception seems to define belongingness of image elements to an object not in a binary (crisp), but more in a graded, or fuzzy, manner. This observation can be successfully transferred and utilized in the field of digital image processing; to handle uncertainties and heterogeneity of object properties appropriately, the suggested methods should be fuzzy, as well, [1, 62]. More precisely, it is, in general, beneficial to perform a *fuzzy segmentation* of an image. Such an approach allows image elements to belong to an object to some extent, and therefore crisp decisions at this early analysis step are avoided. In this way, the risk of making early wrong decisions about object belongingness is reduced, and a larger amount of information is preserved and can be used later in the process.

The result of a fuzzy segmentation of an image containing a single object is a grey-level image of the object of interest, where object points are “white”, background is “black”, and grey-levels in between correspond to partial belonging of the points to the object, determined according to intensity, geometric, or other information available from the image.

To fully exploit the fuzzy framework, appropriate mathematical theories and algorithms for handling fuzzy discrete data are needed, not only for image segmentation, but in all steps of the image analysis process. There are many challenges to address and many questions to answer on the way of developing such. To list just a few: How are objects to be mathematically defined in fuzzy digital setting, to best address graded composition and hanging-togetherness of the image elements? How are fuzzy boundaries to be defined satisfying a Jordan boundary property? What are the appropriate algorithms to extract these entities from scenes in such a way to satisfy relevant definitions? After a discrete fuzzy spatial set (object) is extracted, how to proceed with the analysis and what analysis tools to use? How to, in the end, reach crisp nonambiguous results from the fuzzy data?

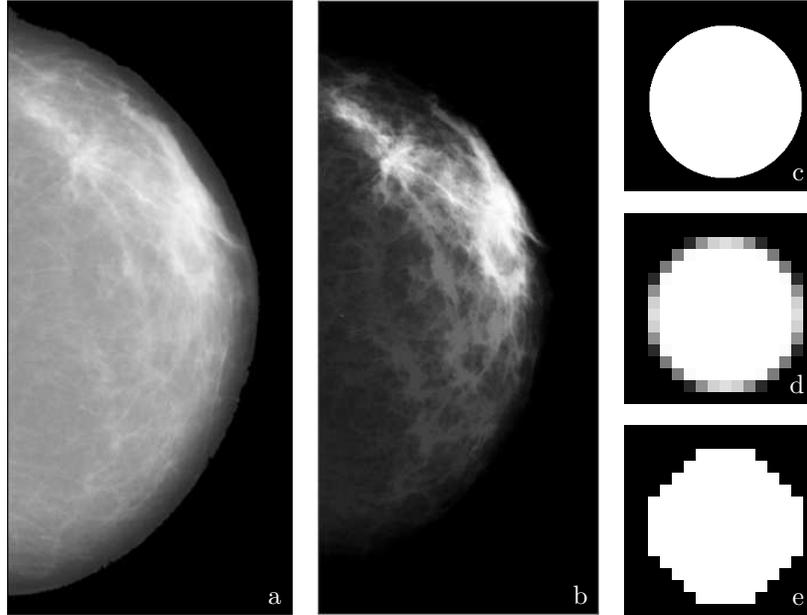


FIGURE 1. Examples of different object representations: (a) grey-level image showing a digitized X-ray mammogram; (b) fuzzy segmentation of a fibroglandular region in (a); (c) high resolution crisp representation of a disk; (d) low resolution coverage representation; and (e) low resolution crisp representation of the same disk.

Coverage representation. In our attempt to contribute to the development of this emerging image processing framework, we have focused our interest to one specific type of fuzzy discrete object representations. These are representations where membership function values correspond to *pixel coverage* (or, in higher dimensional images, *spel coverage*). Pixel values assigned in this model are equal to the relative area of a pixel covered by the imaged (presumably crisp continuous) object. For such images, pixel values (or, coverage values) range from 0 (assigned to pixels having empty intersection with the object) to 1 (pixels completely covered by the object) and the pixel values strictly between 0 and 1 appear only on the border of an object.

Starting from the idea of such a type of object representations, *coverage representations*, we are working on formulating and developing a general image processing framework that utilizes the benefits that come from appropriately treating the coverage information, while still respecting the discrete nature of digital images. We have conducted a number of studies which show many advantages of the proposed type of coverage representations, compared to crisp (binary) digital image representations. We have developed different feature estimators which utilize the coverage

information to improve the estimation precision and accuracy, [50, 52], and have proved that a possible lack of precision resulting from limited spatial resolution may be overcome by properly utilizing grey-level information contained in the images when estimating relevant features of the objects. The encouraging results, proven both theoretically and by empirical studies on synthetic objects, directed our interest to applications of the developed estimators on real images. The first step required for such use was the development of appropriate image segmentation methods that result in a coverage representation. We have suggested several such methods, appropriate for different applications. We further have proposed to utilize the high precision feature values obtained from the coverage information to generate high resolution reconstruction of the observed discrete object, and thereby to “improve” its visual appearance (in crisp representations), too.

In the following sections we describe in more detail our results related to the development of the *coverage model* and its applications in image processing. These results are based on the work presented in a number of publications, where additional details about the individual parts can be found, [28, 30, 33, 34, 50–55]. We will briefly mention some of the applications of the proposed methods, as well [29, 30, 50, 52, 53, 60]. Additionally, we will try to envision some of the possible future research and application directions. We believe they are numerous, since methods that provide results with sub-pixel precision are of highest importance in many fields where precision is a key factor. In addition, analysis of images at low (or simply insufficient) resolution is constantly a hot research topic; with the recent progress in imaging techniques, allowing imaging to reach nanometer scales, a previously inaccessible world of structures of sizes all the way down to molecular scale, opens up. Modern technology, together with humans’ curiosity and vision, constantly challenge science to push its limits ever further. Our wish is to be a part of this journey of exciting research.

2. Background and related work

Our work on coverage models is related to several research tracks within the field of image processing. This section, where we list and briefly introduce some of these tracks, aims at providing the reader a wider context for our research.

Initial studies, showing the usefulness of utilizing grey-level information, when analysing black and white 2D images obtained by a scanner, were presented in the early 1990s. Originating from that work, different methods for sub-pixel segmentation evolved. Due to often direct utilization of the image intensities in the algorithms, developed methods usually have strong ties to a specific method of image acquisition. Two sub-fields of image processing, where particularly refined methods for utilizing the intensity information in segmentation have emerged, are remote sensing and tomographic medical imaging.

The notions of *fractional pixels* and *partial volume (tissue fraction) effect* are often mentioned in remote sensing and in tomographic imaging, respectively. Both these notions relate to two distinct phenomena that influence intensity values in images in an undesired way. The first phenomenon which causes inconsistency

between ideal and achieved capturing of a signal is the blurring that is introduced by the limited resolving power of the imaging system, leading to a “leakage” of the signal from its actual source to the neighbouring regions in the image. The signal appears weaker, but is also spread over a wider area (the notion of *point spread function* is introduced). The second cause of the above mentioned phenomenon is image sampling. The signal from the imaged object is sampled on a discrete grid, but the contours of the image elements do not match the actual contours of the imaged intensity distribution. A number of spels therefore cover multiple image objects. This second effect is present in any digital image; no matter how high spatial resolution is used, discretization will always lead to that (some) image elements are covered by more than one object. If precision (particularly when it comes to measurements) is required, fractional/partial coverage has to be handled carefully. This challenge is exactly the one we address in our work.

Not only the tasks of image analysis, but also those related to visualisation and image generation impose the need for increased (i.e., sub-pixel) precision in image segmentation and careful handling of image intensities. In film-making, the technique to combine two or more images into a single one, referred to as *image compositing*, or *image matting*, dates back to the Lumière brothers. Since the mid-1980s, when advancements in computer graphics allowed matte painters to work directly in the digital realm, this technique has become less associated with double exposures and painted glass, and more with pixels and alpha channels. In chroma key compositing, commonly used for weather forecast broadcasts, wherein the presenter appears to be standing in front of a large map, which in the studio is actually a monochrome blue background, careful treatment of partially covered image elements, to avoid creation of a bluish aura around the presenter, is required. The field of computer graphics also includes a significant amount of work related to anti-aliasing, aiming to reduce the visual disturbance caused by representing smooth objects by square pixels on a screen. Also in this task, a careful treatment of partially covered pixels is most important. Both these techniques have connections to the work presented in this paper.

In our work on development of image analysis methods, our intention is to propose algorithms and approaches which are generally applicable in a range of situations and applications. Therefore, we try to avoid connecting our proposed methods to any particular way of image creation/acquisition. The foundation of the developed framework is, very suitably for the intended generality, in the fuzzy set theory, which provides a both flexible and powerful framework to represent and describe the methods at an abstract level, without ties to applications, or, if desired, even without ties to digital images. This text will, however, mainly stay close to conventional digital images.

In the following of this section, we give some more details about the background within the field of image processing, and, at the end, a brief overview of the concepts and notions of the fuzzy set theory, used in our work.

2.1. Grey-scale information. The relation between the spatial resolution, the grey-level quantization (grey-level resolution), and the achievable reconstruction

accuracy for certain types of images/objects, with a similar motivation as for our later research, is studied in [20]. It is shown that objects with straight edges can be reconstructed without error if grey-levels are not quantized, even though the spatial resolution of an image is limited. In other words, it is concluded that low-resolution grey-scale images of polygonal silhouettes induce less ambiguity than high-resolution bi-level images. This has served as an important inspiration for our work.

To improve the accuracy and precision of local estimators, estimation methods that utilize grey-level information in images have been suggested. In [11], an arc length estimation method that uses normal vectors computed from intensity values, at a number of pixels sampled along the boundary of an object, is presented; arc length is estimated as a cumulative sum of the length of short line segments, derived from the normal directions. A local step may in that way be assigned a variety of normal directions, instead of the very limited set of normal directions available for (Freeman style) local estimators on binary images. Our work on perimeter estimation (see Section 5.3) is based on the same idea that increased precision of normal directions estimation leads to increased precision of perimeter estimates. However, a thorough analysis of the performance of the method presented in [11] is not provided and no optimization of the local lengths is performed. Another approach to increase the precision of measurements is presented in [63]. The method is based on transformation of object boundaries in grey-level images into corresponding volumes, where the length estimation problem is converted into a (simpler) problem of volume estimation. The method relies on sampling theory and discrete approximations of analogue filters. The results are encouraging, but the evaluation is unfortunately only performed on discs of increasing radii, thereby somewhat limiting the possible conclusions. The method also includes some “practical choices” without full theoretical justification. Even though our approach to the same problem differs from the one in [63], it is important to notice that a possibility to increase precision of image analysis results by utilization of grey-levels has attracted attention of a number of researchers during last couple of decades. It still does.

2.2. Remote sensing—fractional pixels. It is not surprising that the issue of mixed pixels is thoroughly addressed in remote sensing applications. Pixels in remotely sensed images are of sizes ranging from a couple of meters to a couple of kilometers, which very often leads to individual pixels being covered by different classes/objects imaged on the ground. To assign the whole pixel to one class (even if that is the class mostly covering the pixel) leads to imprecision which is often intolerable. Estimation of partial coverage (also known as *soft classification*) of a pixel by all individual classes is preferred. Most often used approaches for such sub-pixel proportion estimation are linear mixture models, due to their simplicity. In more complex cases, e.g., due to multiple scattering leading to nonlinear mixtures, or requirements for advanced corrections of atmospheric distortions, rather involved and specialized methods for estimating partial pixel coverage values have been required. Most popular among them are based on neural networks [19].

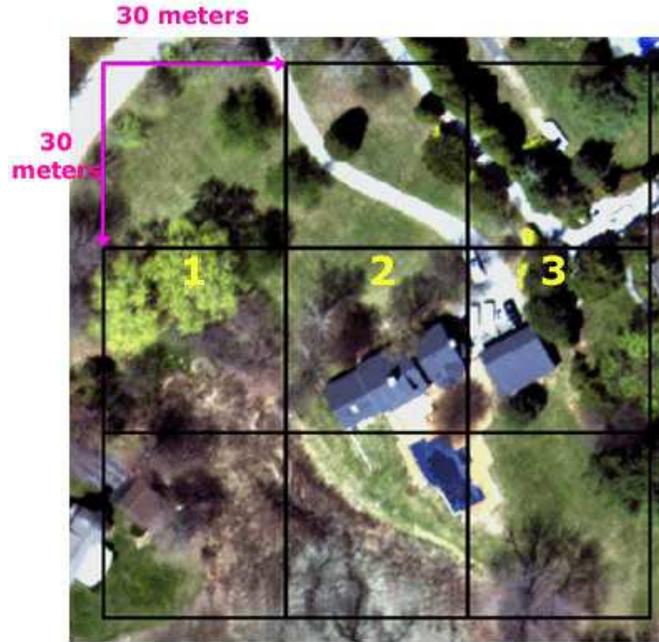


FIGURE 2. A high resolution aerial photograph, with a superimposed 30 meter resolution grid, illustrating the pixel size of a Landsat 7 satellite image. Pixel #1 is a fairly homogenous pixel, almost completely covered with trees, whereas pixels #2 and #3 are mixed pixels, partially covered by vegetation, buildings, and road.

An important characteristics of up-to-date remote sensing is utilization of spectral imaging systems. Spectral imaging for remote sensing of terrestrial features and objects arose as an alternative to high-spatial-resolution large-aperture satellite imaging systems. This type of imaging has evolved to include, instead of just one (grey-scale) band or a few colour bands, several hundred or more bands, encompassing not only the visible spectrum, but also parts of the surrounding electromagnetic spectrum, as well. Data coming from many wavelengths can provide very useful information about the materials in a scene, however extraction of such data usually requires sophisticated processing methods. This is, therefore, an important research direction in remote sensing. It does not fully coincide with our research interest, which is more focused on extraction of information from spatial, rather than from (multi)spectral data. The aim is, however, the same in both cases: to precisely determine the content of a pixel, at sub-pixel precision.

One observation, made in [13], adds additional connection between our work and research interests in remote sensing; it is emphasized that knowledge about the class composition of every pixel still does not provide any information about the spatial distribution of the classes within the pixel. This information can be important

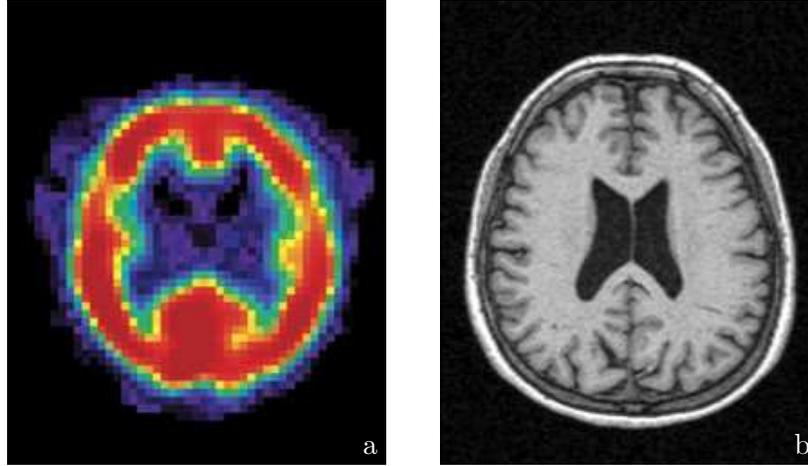


FIGURE 3. A 2D slice from a SPECT image (a), and an MR image (b), showing the same region of a brain.

and one way to acquire it is seen in a multiscale approach, i.e., in utilization of the possibility to observe the environment at a range of scales. This idea is to some extent explored in our work on object reconstruction (see Section 6), where a multiscale approach is taken in the task of high resolution object reconstruction.

2.3. Tomographic images—partial volume effect. Partially covered image elements attract significant attention in medical imaging, not only in cases of low spatial resolution (such as, e.g., SPECT or PET), but also in higher-resolution imaging, such as MRI or CT. Figure 3 illustrates difference in visual appearance due to, among other reasons, difference in spatial resolution, between (two 2D slices of) SPECT and MR images. To appropriately address this issue, in tomographic imaging known as *partial volume effect* (PVE), is particularly important when accurate measurements are required from the acquired images. The significance of this problem, and the need for sub-voxel precision, is well illustrated in [36], where it is shown that consistently misplacing the tissue borders in a brain volume having voxels of size 1 mm^3 by only a single voxel in each slice, resulted in volume errors of approximately 30%, 40%, and 60% for white matter, grey matter and cerebrospinal fluid (CSF), respectively. Negative effects of PVE on tumour detection and monitoring, and on therapy control based on PET images (an imaging modality ideal for this purpose) are thoroughly described in [56].

The complexity of the shape and structure of the human brain, specificities of imaging techniques, and demand for high quality visualization and high precision of (primarily volume) measurements have resulted in a significant number of studies and publications introducing a number of methods for partial volume effect correction in 3D medical imaging. First approaches were not focused on the PVE at a pixel level, but rather on improved estimation of total volume of each tissue in

the whole image, [46]. Further work led to approaches aiming at assigning, to each voxel in an image, an estimated portion of each of the tissues that is contained in it. Often used for that purpose are methods based on expectation-maximization, e.g. [25], scale-space approaches [65], wavelets [3], Markov random fields [5], fuzzy techniques, e.g. [57], etc. Different assumptions can be made, which leads to unmixing models of different complexities; a unifying framework for partial volume segmentation of brain MR images, presented in [25], gives a nice overview.

As opposed to remote sensing, medical tomographic imaging does not rely on a range of spectral bands, but more on the spatial distribution of grey-scale intensities, often in combination with a priori anatomical knowledge. In that sense research conducted to address PVE corresponds more to our main interests. However, our developed methods are more general and less tied to particular imaging situations than what is common for methods developed for handling PVE in medical imaging.

2.4. Fuzzy set theory in image processing. A fuzzy set is a collection of elements with a continuum of grades of membership; it is characterized by a membership function, which assigns a membership value between zero and one to each element. A fuzzy set is a generalization of a crisp set; while a crisp set either contains a given element, or it does not, which is described by the membership values one and zero, respectively (as given by the characteristic function of a set), belongingness of an element to a fuzzy set can be partial, and is therefore described by any value between zero and one. When introduced by Zadeh [68], the notion of a fuzzy set was intended to provide a starting point for the building of a conceptual framework, to exist in parallel with the framework of crisp (“ordinary”) sets, but to be more general and potentially provide increased applicability in different fields; image analysis became one of them. The framework provided a natural way of dealing with problems in which the source of imprecision is in the absence of sharply defined criteria for class membership.

Having on mind the difficulties in image segmentation, mainly caused by the existence of nonsharp boundaries between the objects in an image, it is not surprising that the comfortability of fuzzy sets, not forcing us to make hard (and possible wrong) decisions about object belongingness, became appreciated and well accepted in image analysis; for an overview of several applications, see [42].

A fuzzy membership function is defined as a mapping from an arbitrary set, the reference set, to, usually, the interval of real numbers $[0, 1]$. More formally, a *fuzzy subset* S of a reference set X is a set of ordered pairs $S = \{(x, \mu_S(x)) \mid x \in X\}$, where $\mu_S : X \rightarrow [0, 1]$ is the *membership function* of S in X [68].

The crisp set of points having strictly positive memberships to the set S is called the *support* of S , while the *core* of a fuzzy set S contains the points with memberships to S equal to 1 (it is sometimes referred to as the *kernel*). When defined on a discrete domain, the membership function is a discrete function, and a corresponding set is a discrete fuzzy set.

To represent an object in an image, we usually consider a fuzzy set defined on \mathbb{Z}^2 or \mathbb{Z}^3 , being typical spaces of discrete images. Such a set is called a *discrete spatial*

fuzzy set [1]. When represented in a computer, the number of different membership values is finite; integer values are often used to represent memberships, to increase the speed of computations. In this way, the range of a digital fuzzy function is not the interval $[0, 1]$, but rather the set $\{0, 1, \dots, \ell\}$. The value ℓ is often equal to 255, or 65 535, which corresponds to 8-, or 16-bit pixel depth (number of bits used to represent a pixel value).

Fuzzy set theory is nowadays rather rich and well developed. However, most of the theoretical results are derived for continuous, analytically defined, membership functions and often strongly rely on the properties and analytical expressions of these functions. On the other hand, membership values of image elements are derived from grey-levels, assigned to the image points during an imaging process, and sometimes, additionally, from a set of criteria designed to capture geometric, structural, and other properties of the imaged object. This makes the membership functions on an image (defining fuzzy objects observed) highly complex and practically never analytically defined. As a consequence, many of the well defined and thoroughly explored notions, relations, and properties of (analytically defined) continuous fuzzy sets become nonapplicable to the discrete fuzzy sets, which are most common in image processing. Therefore, it is often required to design new methods, which are more appropriate for the analysis of discrete fuzzy sets, and to develop mathematical theories and algorithms for handling fuzzy discrete data appearing in digital images.

A representation of a fuzzy set, which is often used as an alternative to representation by a membership function, is the one based on α -cuts. For a fuzzy set F , defined on a reference set X , the following two representations are equivalent [10]:

- a membership function $\mu_F : X \rightarrow [0, 1]$ which assigns to each $x \in X$ its membership grade $\mu_F(x)$ to the fuzzy set F ;
- the set of α -cuts $\{F_\alpha \mid \alpha \in (0, 1]\}$ of the set F , where $F_\alpha = \{x \in X \mid \mu_F(x) \geq \alpha\}$.

The connection between the membership function and the stack of α -cuts provides a common approach for extending functions defined on crisp sets, to functions defined on fuzzy sets. The so called fuzzification principle, based on one of the following equations:

$$(2.1) \quad f(S) = \int_0^1 \hat{f}(S_\alpha) d\alpha,$$

$$(2.2) \quad f(S) = \sup_{\alpha \in (0,1]} [\alpha \hat{f}(S_\alpha)]$$

can be used to extend a function \hat{f} to the domain of fuzzy sets. In this way, various properties defined for crisp sets (here, α -cuts) can be generalized to fuzzy sets, including the membership function itself; if the characteristic functions of the α -cuts are observed, the membership function of the corresponding fuzzy set can be obtained by either their integration (2.1), or by taking the supremum of their weighted values, over the “height” of the stack (2.2). This approach, and in particular equation (2.1), is often used in our work, when some relevant features of

objects represented by a coverage model are defined and extracted. Notably, area and perimeter of a spatial fuzzy set in 2D, can be defined based on (2.1) [41, 43].

One particular concept that is extended to the set of fuzzy sets, which we will use in the following, is that of a partition. Classically, a partition of a set S is a family of disjoint nonempty subsets of X whose union is equal to X . An often used definition of a fuzzy partition is the following one [45]: A *fuzzy partition* of a set X is a finite family $P = \{P_1, P_2, \dots, P_n\}$ of nonempty fuzzy subsets of X such that $\sum_{k=1}^n \mu_{P_k}(x) = 1$ for all $x \in X$.

3. The coverage model

The approach that we take to handle partially covered image elements and to best utilize intensity information in order to reach sub-pixel precision of estimates, differs quite significantly from most previously presented work with similar goals. As noticed in the background section, a lot of related work is based on more or less direct usage of the image intensities, leading to strong ties between developed methods and the specific imaging conditions. The path we take is, instead, to start from a well defined abstract theoretical model, with no connections with any particular application. For the proposed theoretical framework, we have developed feature estimators and derived exact results regarding their performance. The connection to specific imaging conditions is handled through a separate segmentation step, which serves the purpose of transforming the application dependent image information into an application independent form. This clear separation between the different parts of the presented approach is what provides generality of the developed image processing tools, where application specific information can be fully utilized in the segmentation step, while still not interfering negatively with the later steps.

The foundation of the proposed model lies in the fuzzy set theory, which provides a framework that has shown to be both powerful and flexible. Within the field of image processing, methods utilizing the concept of fuzzy/graded memberships have found a number of applications, and discrete fuzzy sets have shown to be a very good tool for representing image objects. The ability provided by fuzzy sets to represent uncertain and vague data facilitates development of robust methods which successfully handle noise and image artefacts. At the same time, these methods can be designed to enable high precision of measurements, overcoming well known problems originating from the discretization of the continuous space observed, unavoidably imposed by essential properties of computers and imaging devices involved.

The fuzzy set theory is very general and provides a lot of flexibility; the interpretation of the membership values can be adjusted to any need, property, or application. A great power of this large freedom lies in the possibility to, by adding well chosen restrictions, shape it so that it best fits a particular problem observed. Quite clearly, keeping full generality available in every applications is, in addition to being difficult, also hardly practically useful. Not surprisingly, the large freedom makes it difficult to derive well defined and strong statements about specific properties of the observed (general) fuzzy sets. We found that, by suitably restricting

the interpretation of membership values, strong theoretical results are more readily available.

Our research focuses on mathematical tools for representing and analyzing continuous crisp objects in digital n -dimensional images. For that specific task we have appropriately restricted the fuzzy membership function used, and have defined a particular type of fuzzy sets for representation of imaged objects. We refer to such representation of digital objects as *coverage representations*. We have shown that using such a coverage representation has many advantages compared to a traditional crisp representation.

3.1. Basic definitions. We start with a general definition of a coverage representation. We will then restrict it to better fit the application we have on mind, i.e., representing crisp objects in digital images.

Definition 3.1. Given a partition $\Sigma = \{\sigma_i\}_{i \in I}$ of a reference set X , a coverage representation of a set $S \subset X$ on Σ is a fuzzy subset $\{(\sigma_i, \alpha(\sigma_i)) \mid \sigma_i \in \Sigma\}$, such that $\alpha(\sigma_i) = |\sigma_i \cap S|/|\sigma_i|$.

In the context of digital image processing, we assume that the reference set X is the Euclidean space \mathbb{R}^n and that Σ is the Voronoi tessellation of \mathbb{R}^n defined by the set of integer points \mathbb{Z}^n . We refer to the Voronoi region of a grid point $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{Z}^n$ as the *spel* at \mathbf{x} (short for spatial element) and denote it with $\sigma(\mathbf{x})$. That is, $\sigma(\mathbf{x})$ contains the points of \mathbb{R}^n which are closer to \mathbf{x} , in terms of Euclidean distance, than to any other point in \mathbb{Z}^n (for points at equal distance we round upwards, i.e., the lower/left edge in each dimension is included in the spel). In other words, the set Σ^n of n D spels of an integer grid, consists of translations of the right open n -dimensional unit origin-centred cube by vectors $\mathbf{x} \in \mathbb{Z}^n$:

$$\Sigma^n = \{\sigma(\mathbf{x}) \mid \mathbf{x} \in \mathbb{Z}^n\}, \quad \sigma(\mathbf{x}) = \left[-\frac{1}{2}, \frac{1}{2}\right)^n + \mathbf{x}.$$

Based on the above we define the following digitization model:

Definition 3.2. For a given continuous object $S \subset \mathbb{R}^n$, inscribed into a grid \mathbb{Z}^n , the *coverage digitization* of S is

$$\mathcal{D}_c(S) = \{(\mathbf{x}, \alpha(\mathbf{x})) \mid \mathbf{x} \in \mathbb{Z}^n\}, \quad \alpha(\mathbf{x}) = \frac{|\sigma(\mathbf{x}) \cap S|}{|\sigma(\mathbf{x})|},$$

where $|X|$ here denotes the area/volume/Lebesgues measure of a set X .

Remark 3.1. To simplify notation, we utilize the one-to-one correspondence between elements $\mathbf{x} \in \mathbb{Z}^n$ and their respective spels (Voronoi regions) $\sigma(\mathbf{x}) \in \Sigma^n$, and consider a coverage digitization to be a set of ordered pairs from $\mathbb{Z}^n \times [0, 1]$, and not from $\Sigma^n \times [0, 1]$.

Definition 3.2 assumes assignment of nonquantized real coverage values α to the spels of the grid. However, when using digital approaches (computers) to represent, store, and analyze images, we are limited to a finite number of grey-levels to represent coverage of an individual spel. This leads to the following quantized version of coverage digitization:

Definition 3.3. For a given continuous object $S \subset \mathbb{R}^n$, inscribed into a grid \mathbb{Z}^n , the ℓ -level quantized coverage digitization of S is

$$\mathcal{D}_c^\ell(S) = \{(\mathbf{x}, \alpha^\ell(\mathbf{x})) \mid \mathbf{x} \in \mathbb{Z}^n\}, \quad \alpha^\ell(\mathbf{x}) = \frac{1}{\ell} \left\lfloor \ell \frac{|\sigma(\mathbf{x}) \cap S|}{|\sigma(\mathbf{x})|} + \frac{1}{2} \right\rfloor,$$

where $\lfloor x \rfloor$ denotes the largest integer not greater than x .

Clearly $|\alpha(\mathbf{x}) - \alpha^\ell(\mathbf{x})| \leq \frac{1}{2\ell}$. We denote the set of possible coverage values $\alpha^\ell(\mathbf{x})$ in ℓ -level quantized coverage digitization by $\mathcal{Q}_\ell = \{0, \frac{1}{\ell}, \frac{2}{\ell}, \dots, \frac{\ell}{\ell} = 1\}$. This set corresponds to the set of grey-levels available; e.g., $\ell = 1$ for a binary image, while $\ell = 255$ provides the set of grey-levels for an 8-bit representation. Similarly as using spatial resolution to denote the spatial sampling density, we let *coverage resolution* denote the number of (meaningful) coverage levels.

The coverage digitization model stands in contrast to the more common Gauss digitization model, where an object is represented by the set of integer grid points within the objects (or, mainly being a matter of notation, by the corresponding set of spels). More formally:

Definition 3.4. The *Gauss digitization* of $S \subset \mathbb{R}^n$ is $\mathcal{D}_g(S) = \{S \cap \mathbb{Z}^n\}$.

In our work, coverage of a spel is sometimes approximated by a super-sampling approach where a spel is split into several sub-spels, and a sample is taken from the centre of each. This facilitates easy approximation of coverage values for more complex synthetic objects, where true coverage may be difficult to compute analytically. Let the r -sampled spel $\hat{\sigma}^r(\mathbf{x})$ be the following set of r^n points within $\sigma(\mathbf{x})$:

$$\hat{\sigma}^r(\mathbf{x}) = \left\{ \sigma(\mathbf{x}) \cap \left(\frac{\mathbf{y} - \delta(r)}{r} \right) \mid \mathbf{y} \in \mathbb{Z}^n \right\},$$

where $\delta(r)$ is the vector $(\frac{r-1}{2}, \frac{r-1}{2}, \dots, \frac{r-1}{2})$.

Definition 3.5. For a given continuous object $S \subset \mathbb{R}^n$, inscribed into a grid \mathbb{Z}^n , the r -sampled coverage digitization of S is

$$\hat{\mathcal{D}}_c^r(S) = \{(\mathbf{x}, \hat{\alpha}^r(\mathbf{x})) \mid \mathbf{x} \in \mathbb{Z}^n\}, \quad \hat{\alpha}^r(\mathbf{x}) = \frac{|\hat{\sigma}^r(\mathbf{x}) \cap S|}{|\hat{\sigma}^r(\mathbf{x})|}.$$

Remark 3.2. Coverage values of a 1-sampled coverage digitization, $\hat{\alpha}^1(\mathbf{x})$, correspond to the characteristic function of a Gauss digitization.

Figure 4 illustrates the different digitization approaches and the output for one pixel, partly covered by a disk shaped imaged object S . The observed pixel is not included in the Gauss digitization (Fig. 4(a)), since the centre of the pixel is not covered. For the coverage digitization (Fig. 4(b)), the pixel is associated with a real number $\alpha(\mathbf{x}) \in [0, 1]$ indicating how large area of the pixel is covered by the object. In the ℓ -level quantized coverage digitization (Fig. 4(c)), the real coverage is approximated by its closest number in \mathcal{Q}_ℓ . For the r -sampled coverage digitization (Fig. 4(d)), the pixel is divided into sub-pixels, and the coverage is approximated by the number of covered sub-pixel centres, divided by the number of sub-pixels.

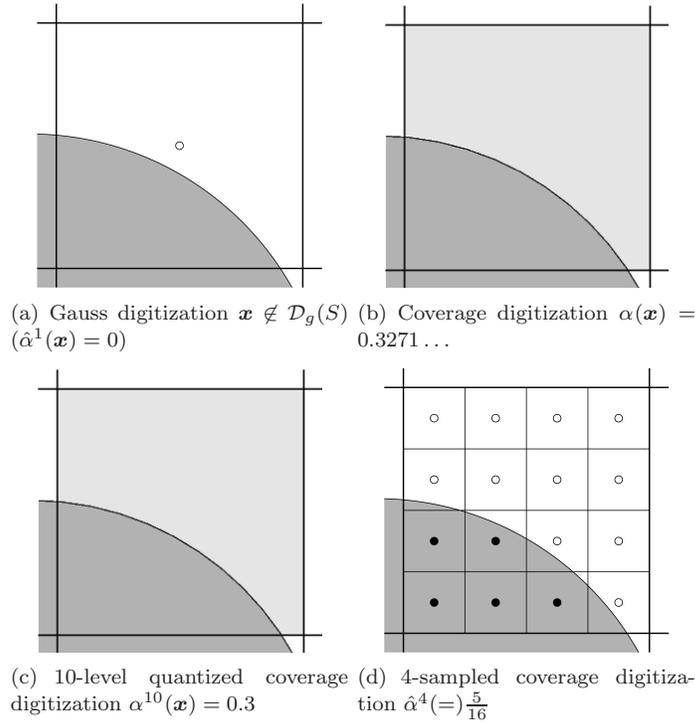


FIGURE 4. Different digitization approaches. Shown is one pixel $\sigma(\mathbf{x})$, partly covered by a disk shaped object S .

The r -sampled coverage digitization takes values from the same set \mathcal{Q}_ℓ as an ℓ -level quantized coverage digitization with $\ell = r^n$. However, the r -sampled digitization has lower accuracy in approximating the true coverage digitization values. For a spel \mathbf{x} , intersected by a straight edge of an image object, it holds that $|\alpha(\mathbf{x}) - \hat{\alpha}^r(\mathbf{x})| \leq \frac{1}{2r}$. Another important property to notice is that the quantized coverage digitization $\mathcal{D}_c^\ell(S)$, with $\ell = r^n$, is never a worse approximation of the true coverage digitization $\mathcal{D}_c(S)$ than the r -sampled coverage digitization $\hat{\mathcal{D}}_c^r(S)$. This holds by definition, since $\mathcal{D}_c^\ell(S)$ and $\hat{\mathcal{D}}_c^r(S)$ take values from the same set, and $\alpha^\ell(\mathbf{x}) = \arg \min_{\beta \in \mathcal{Q}_\ell} |\beta - \alpha(\mathbf{x})|$.

3.2. Properties of coverage digitizations. A coverage representation of a crisp real object with a well defined continuous border is, ideally, characterized by the presence of homogeneous connected regions of “pure” spels, completely covered by either object or background. Each two such regions are separated by a thin layer of “mixed” spels, i.e., those partially covered by both object and background. Pure spels are assigned coverage values 0 (background) or 1 (object), while mixed spels

are assigned values between 0 and 1, in accordance to their respective coverage by the image object.

Following the terminology of fuzzy sets, we define the *core* and the *support* of a coverage representation as the crisp set of spels with coverage values 1 and the crisp set of spels with nonzero coverage values, respectively:

$$\text{core}(C) = \{\mathbf{x} \mid \alpha(\mathbf{x}) = 1\}, \quad \text{supp}(C) = \{\mathbf{x} \mid \alpha(\mathbf{x}) > 0\}.$$

We note that there is a close connection between $\text{supp}(\mathcal{D}_c(S))$ and the *outer Jordan* or *super-cover digitization* of the set S (the union of all spels having a nonempty intersection with S), where a difference between the two notions appears only when the nonempty intersection of a spel and the set S has a measure zero. Such a spel is included in the outer Jordan digitization but not in $\text{supp}(\mathcal{D}_c(S))$. Similarly, for $\text{core}(\mathcal{D}_c(S))$ and the *inner Jordan digitization* of S , where the two notions differ only for spels with a nonempty intersection with the background of measure zero. Such a spel is excluded from the inner Jordan digitization but is not excluded from $\text{core}(\mathcal{D}_c(S))$. These differences, being mainly theoretical in the case of a real valued coverage digitization, become more prominent in the case of quantized or sampled coverage digitizations.

We call the set of “mixed” spels of a coverage representation C , with coverage values strictly between 0 and 1, the *fuzzy border* of C . This set is equal to the closed difference set between $\text{supp}(C)$ and $\text{core}(C)$. Formally:

Definition 3.6. The *fuzzy border* of a coverage representation $C = \{(\mathbf{x}, \alpha(\mathbf{x}))\}$ is

$$\partial_f C = \{\mathbf{x} \mid \alpha(\mathbf{x}) \in (0, 1)\}.$$

If the crisp set S has a reasonably smooth boundary (i.e., S is Jordan measurable) and digitized at a high enough resolution, then the fuzzy border $\partial_f \mathcal{D}_c(S)$, is not more than one spel thick.

An important property of coverage digitization is that it preserves partitions of the digitized space. Let $\{S_k \mid |S_k| \neq 0, k = 1, 2, \dots, m\}$ be a partition of a reference set X . Then $\{\mathcal{D}_c(S_k) \mid k = 1, 2, \dots, m\}$, is a fuzzy partition [45], i.e., $\forall k : \mathcal{D}_c(S_k) \neq \emptyset$, and $\forall \mathbf{x} \in X : \sum_{k=1}^m \alpha_k(\mathbf{x}) = 1$.

Proof. Follows directly from the additivity of the Lebesgue measure on \mathbb{R}^n : for disjoint sets $A, B \in \mathbb{R}^n$, $|A \cup B| = |A| + |B|$. \square

We call such a fuzzy partition, where membership values correspond to spel coverage, a *coverage partition* of X .

Remark 3.3. Note that a similar family of ℓ -level quantized coverage digitizations is not necessarily a coverage partition. For example, a 10-level quantized coverage digitization of a spel $\sigma(\mathbf{x})$ equally covered by three objects will lead to $\sum_{k=1}^3 \alpha_k(\mathbf{x}) = 0.9 \neq 1$. A sampled coverage digitization of a partition is, however, always a fuzzy partition.

4. Coverage segmentation methods

The definition of coverage digitization provides a way to compute coverage representations of relatively simple, or analytically defined, continuous subsets of \mathbb{R}^n such as, e.g., simple geometric objects. However, to extract similar information about more realistic and more complex objects in digital images, we instead need segmentation algorithms. To address the task of extracting coverage information about objects in images, we have proposed a number of *coverage segmentation* methods.

As already mentioned, image segmentation is commonly considered to be the process of assigning a label to every spel in an image, so that spels with the same label share certain visual characteristics. Traditionally, such an assignment is done in a crisp fashion, where a spel can only be associated with one single component of the image. More generally, a segmentation may be performed in a fuzzy manner, in which case each spel is associated with a membership, in the range zero to one, to each of the image components. The membership values assigned to one spel do not have to sum up to one, although this is assumed for many methods. For the coverage model it seems reasonable to assign memberships (coverage values) of a spel to different crisp nonoverlapping objects in \mathbb{R}^n , so that they do sum up to one.

Let \mathbb{A}_m denote the set of m -component (fuzzy) *segmentation vectors*

$$\mathbb{A}_m = \left\{ \boldsymbol{\alpha} = (\alpha_1, \alpha_2, \dots, \alpha_m) \in [0, 1]^m \mid \sum_{k=1}^m \alpha_k = 1 \right\},$$

and let $\mathcal{A}_m \subset \mathbb{A}_m$ be the corresponding set of crisp *segmentation vectors*

$$\mathcal{A}_m = \left\{ \boldsymbol{\alpha} \in \{0, 1\}^m \mid \sum_{k=1}^m \alpha_k = 1 \right\}.$$

A *coverage segmentation* of an image I into m components is a set of ordered pairs

$$\mathcal{S}_c(I) = \{(\mathbf{x}, \boldsymbol{\alpha}(\mathbf{x})) \mid \mathbf{x} \in I_D, \boldsymbol{\alpha}(\mathbf{x}) \in \mathbb{A}_m\}, \quad \alpha_k \approx \frac{|\sigma(\mathbf{x}) \cap S_k|}{|\sigma(\mathbf{x})|},$$

where $S_k \subset \mathbb{R}^n$ is the extent of the k -th (out of m) image component and $I_D \subseteq \mathbb{Z}^n$ is the discrete image domain. The continuous sets S_k are, in general, not known, and the values α_k therefore have to be estimated from the image data. We may also refer to a coverage segmentation as a coverage partition of I_D , i.e., a collection of nonoverlapping sets $\mathcal{S}_c(I)_k = \{(\mathbf{x}, \alpha_k(\mathbf{x})) \mid \mathbf{x} \in I_D\}$, $k = 1, 2, \dots, m$.

In many imaging situation, acquired image intensities correspond almost directly to coverage values. One such situation is, e.g., the integration of photons over finite sized sensor elements, as present in a digital camera. In the absence of object texture and large illumination variations, a suitable mapping of the image intensities may provide good enough coverage values. In Section 4.1 we present a method for automatically defining such a mapping by a double thresholding scheme. Direct mapping of intensity values does not work well in the presence of large scale intensity variations, however. In Sections 4.2 and 4.3 we present more elaborate methods for performing coverage segmentation.

4.1. Coverage segmentation based on double thresholding. Segmentation by thresholding is the most intuitive and simplest segmentation method. Successful application of any thresholding based segmentation method requires that the intensity distributions of the object and the background are well separable. If a separation between the image components is to be performed in a crisp manner, one threshold value is selected and two disjunct sets of intensities are determined; the corresponding spels are then accordingly classified into two classes.

Assume that we are imaging a bright object on a black background. This situation will, under reasonable conditions, result in two well separated sets of intensity values, and classification (segmentation) can be performed by thresholding. If the imaging device has a linear response, and the optical blurring is small compared to the pixel size, the partly covered pixels along the boundary of the object exhibit grey values between those of the background and the foreground, where the grey value assigned to a pixel is proportional to its coverage by the object. This observation makes it natural to try to develop a segmentation method that, to a high extent, utilizes the grey-levels of the graded transition between the two observed classes (object and background) for estimating coverage values. We notice that such a direct relation between grey-levels and spel coverage is a reasonable model for images where resolution is decided based on limited means for handling of the data rather than on the optical system (e.g. in high speed video), or when detector elements are intentionally grouped together (binned) to reduce photon/Poisson noise (e.g. in low dose CT).

In order to estimate coverage values from the image spel intensities, we need to first estimate the grey-levels of the completely covered foreground and background spels, respectively. Denote with f and b the intensity of the imaged foreground (object) and background, respectively. We then model the intensity $I(\mathbf{x})$ of a partly covered pixel $\sigma(\mathbf{x})$, with coverage $\alpha(\mathbf{x})$, as a convex combination of f and b :

$$(4.1) \quad I(\mathbf{x}) = \alpha(\mathbf{x}) \cdot f + (1 - \alpha(\mathbf{x})) \cdot b \quad \Rightarrow \quad \alpha(\mathbf{x}) = \frac{I(\mathbf{x}) - b}{f - b}$$

In [52] we presented a method which is based on the above model and which automatically finds two threshold values, f and b , that define the minimum intensity of the high intensity component (foreground) and the maximum intensity of the low intensity component (background). The threshold selection is based on the observation that coverage representations of crisp continuous objects are characterised by having fuzzy boundaries which are not more than one spel thick (see Section 3.2). This is, however, not enough to uniquely define the threshold; we additionally require the contrast between foreground and background to be as large as possible, giving a border with as rich intensity variations as possible. The initial reason to perform a coverage segmentation was to preserve the information given by the grey-levels; the more we keep and use, the better results of subsequent analysis.

To summarize, given a grey-scale image, we seek a threshold couple, b and f , where spels darker than b are considered to belong completely to the background, while spels brighter than f are considered to belong completely to the foreground,

such that the spels in between form a one spel thick separating layer. In addition, we want the contrast between foreground and background, i.e., the difference $\ell = f - b$, to be as large as possible.

The algorithm loops over all background thresholds b' ; each of them defines a possible support of the object. Instead of looping over all foreground thresholds, f' is found as the minimum grey-level within the core of the potential object. For an assumed support, the core should be not more than at one spel distance, which is conveniently expressed by mathematical morphology. For each potential support, the corresponding minimal core is found by eroding the support with a $3 \times 3 \times \dots \times 3$ binary structuring element. The foreground threshold f' is then found as the minimal grey-level in that potential core. However, we noticed that individual dark noise points make f' unnecessarily low. We therefore first perform a morphological closing before selecting the level f' . The difference $f' - b'$ is computed, and a new background threshold is tested. The thresholds leading to the largest intensity range $\ell = f' - b'$ is selected as the best one and, based on that, partial coverage values are computed according to (4.1). We also include an opening of the background, to avoid isolated bright noise points in the background appearing as object points. The processing is invariant w.r.t. intensity inversion; looping over the foreground thresholds and finding b' as the maximum value of the opened background, leads to the same result.

Instead of first performing thresholding, and then morphological operations, we speed up the processing by using grey-scale morphology (where erosion and dilation become min and max filters). The more time consuming morphological operations can then be done once for the whole process, and only threshold and min operations remain inside the loop.

Algorithm 1 summarizes the described steps (see also [52]). We denote the grey-scale erosion $I \ominus B$ by εI and the grey-scale dilation $I \oplus B$ by δI . The opening and closing of I by B are denoted $\delta \varepsilon I$ and $\varepsilon \delta I$, respectively.

An example of the use of this coverage segmentation method is presented in Section 7.1.

4.2. Coverage segmentation by local unmixing. Image segmentation is, as already stated, a difficult problem, which has been addressed more times than any other problem in image processing. As a consequence, very many segmentation algorithms have been proposed. A variety of methods, utilizing different theoretical concepts, and being more or less general, i.e., more or less tied to a particular application, are published. With an increasing complexity of the imaged scene, more information than bare grey-scale intensities of individual spels is required to perform successful (meaningful) separation of image components, which makes thresholding based methods nonapplicable. Spatial (geometric) information about the objects in the image is very often utilized, but other types of a priori knowledge about the imaged objects, if available, may be incorporated in a segmentation method, as well. Segmentation can be performed based on intensity homogeneity preservation (region based methods), or discontinuity detection (boundary based approaches); more and more often, a combination of these approaches is proposed,

Algorithm 1.

Input: A grey-scale image I with a bright object on a dark background.

Output: An approximate ℓ -level coverage segmentation C of the object in I .

```

 $b = 0; f = 0$ 
for each grey-level  $b'$ 
   $F' = \{\mathbf{x} \mid [\varepsilon I](\mathbf{x}) > b'\}$  /* Foreground */
  if  $F' \neq \emptyset$ 
     $f' = \min_{\mathbf{x} \in F'} [\varepsilon \delta I](\mathbf{x})$ 
    if  $f' - b' > f - b$  /* Better than previous */
       $f = f'; b = b'$ 
    endif
  endif
endif
endfor

 $\ell = f - b$ 
 $\alpha(\mathbf{x}) = \begin{cases} 0 & , \quad [\delta \varepsilon I](\mathbf{x}) \leq b, \\ 1 & , \quad [\varepsilon \delta I](\mathbf{x}) \geq f, \\ \frac{I(\mathbf{x}) - b}{\ell} & , \quad \text{otherwise.} \end{cases}$ 

```

to answer better to the high complexity of the tasks. If available, colour information is precious. Typically, three channels are used in colour imaging, which adds a lot of information compared to monochrome (grey-scale) images. The number of channels may, in some imaging techniques, be even higher (reaching hundreds), leading to so-called spectral images, typical in, e.g., remote sensing (see Section 2.2). Segmentation methods design to extract such type of information are accordingly developed.

Noticing this rich variety of different segmentation methods, more or less refined and adjusted to various imaging conditions, we propose to not “re-invent” a whole range of segmentation methods that should provide good coverage segmentation for a variety of tasks, but instead appropriately adjust already existing crisp segmentation methods. After all, our aim to perform segmentation so that it leads to a coverage representation of objects, is not far from the aim of “traditional” segmentation methods, providing a crisp segmentation. For all image spels which are completely covered by a single object class, there is no difference between a coverage segmentation and a crisp segmentation; for those spels there is no reason for a different output than what is achieved by any appropriately chosen “traditional” segmentation method.

In [53] we presented a method that, based on *any* existing crisp segmentation, enhances it to a coverage segmentation by identifying boundary spels and suitable re-evaluating their coverage values. By this approach we reached two important goals: (i) the segmentation results in a coverage representation of an object; (ii) all the advantages of well chosen crisp segmentation methods for a particular task are preserved and utilized. The methods is briefly described in the following.

To obtain a coverage segmentation, we propose a method composed of four steps:

- application of a crisp segmentation method, appropriately chosen for the particular task;
- selection of spels to be assigned partial coverage;
- application of a local linear mixture model for “un-mixing” of partially covered spels and assignment of corresponding coverage values;
- ordered thinning of the set of partly covered spels to provide one spel thin fuzzy borders (see Section 3.2) of mixed spels.

The first step in the proposed method is expected to provide correct assignment of class belongingness to pure spels. We suggest to utilize any appropriate existing segmentation method, and assume that the resulting segmentation provides a trustworthy result for all but boundary spels. Each spel $\sigma(\mathbf{x})$, inner (i.e., not neighbouring a spel of a different component) for the component k , is assigned crisp segmentation vectors $\boldsymbol{\alpha}(\mathbf{x})$ such that $\alpha_k = 1$ and $\alpha_{l \neq k} = 0$.

In the second step of the suggested segmentation method, spels possibly being intersected by the boundaries of continuous imaged objects are to be detected. Such spels are possibly mixed, with partial coverage by two or more image components. We define the set B to consist of all (n D) spels sharing an $(n - 1)$ -dimensional hyper-surface with a spel assigned a different segmentation label. In the sense of Definition 3.6, these spels are candidate mixed spels, and as such, they will be processed in the next steps of the algorithm. If continuous crisp objects are imaged at a reasonably high resolution, and the segmentation performed in step one correctly labels inner, completely covered spels, then the set of mixed spels will be a subset of the set B . Even though it is clear that, in the presence of noise, inner region spels are not of accurate reference intensity of a pure class, but are often exhibiting properties of mixed spels, the idea is to have confidence in the used crisp segmentation method up to the dichotomization into inner/pure and border spels. The spels detected as inner will, therefore not be revisited, or reassigned.

The third step in the coverage segmentation process is computation of partial coverage values of the (potentially mixed) spels of the set B . We suggest to use a linear model, due to its simplicity, and the fact that it corresponds to the ideal (noise-free) spel coverage assignment that arises when integrating spatially distinct signals over finite sized detector elements (e.g. in a digital camera). This model assumes that the value of a mixed spel is a convex combination of the values corresponding to the pure classes c_k covering the observed spel, where the coefficients in the combination correspond to the proportions of the pure classes appearing. Note that for imperfect imaging devices, the assumption that the value of a spel depends only on the content of that particular piece of the image domain, may not hold. Given a particular imaging situation, it is recommended to verify this assumption and possibly act accordingly, e.g. by incorporating a deconvolution step into the process.

In general, the intensity values of the pure classes are not known, but have to be estimated from the image data. We suggest to use a local approach when estimating the intensities characterizing a class k . For each spel observed in the process of

partial coverage assignment, the local pure class representation $c_k(\mathbf{x})$ is estimated as the mean value of the image intensities in a local neighbourhood of a suitable size, which are classified, according to the two first steps, as completely belonging to the observed class k . This approach, in our opinion, has two main advantages: 1) only the relevant classes –existing in the neighbourhood of the observed spel are considered for a mixture in that spel, and 2) sensitivity of the pure class description to intensity variations over the image is decreased; in general, the local within class variation is significantly smaller than the global one.

The image intensity values $I(\mathbf{x}) = (I_1, I_2, \dots, I_b)$ of a mixed spel $\sigma(\mathbf{x})$ (b being the number of channels (bands) of the image) are assumed, in a noise-free environment, to be a convex combination of the (locally estimated) m existing pure classes $c_k(\mathbf{x})$:

$$(4.2) \quad I(\mathbf{x}) = \sum_{k=1}^m \alpha_k c_k(\mathbf{x}), \quad \sum_{i=k}^m \alpha_k = 1, \quad \alpha_k \geq 0,$$

where each coefficient α_k corresponds to the coverage of the spel $\sigma(\mathbf{x})$ by an object of a class k . In a noise-free environment, and if the number m of classes (variables) is not bigger than the number $(b+1)$ of equations (including the equation $\sum_{k=1}^m \alpha_k = 1$), the problem of partial coverage is solved as a system of linear equations.

In real imaging conditions noise has to be considered. However, in the presence of noise, it is not certain that there exists a (convex) solution to the linear system (4.2). Therefore we reformulate the problem to the following minimization problem:

Find a vector I^ of the form $I^* = \sum_{k=1}^m \alpha_k^* c_k(\mathbf{x})$, such that I^* is a convex combination of $c_k(\mathbf{x})$ and the distance $d(I(\mathbf{x}), I^*)$ is minimal.*

The distance measure can be selected to appropriately fit the settings, e.g., a locally estimated Mahalanobis distance. For simplicity, we use the Euclidean distance in the following.

We solve the constrained optimization problem by using Lagrange multipliers method (leading to a least squares type of problem), and we minimize the function

$$F(\alpha_1, \dots, \alpha_m, \lambda) = \left\| I(\mathbf{x}) - \sum_{k=1}^m \alpha_k c_k(\mathbf{x}) \right\|_2^2 + \lambda \left(\sum_{k=1}^m \alpha_k - 1 \right)$$

over all $\alpha_k \geq 0$, for given intensity values of a spel $I(\mathbf{x})$ and local class intensities $c_1(\mathbf{x}), \dots, c_m(\mathbf{x})$. The obtained solution $\boldsymbol{\alpha}(\mathbf{x})^* = (\alpha_1^*, \dots, \alpha_m^*)$ provides estimated partial coverage of the spel $\sigma(\mathbf{x})$ by each of the observed classes $k \in \{1, 2, \dots, m\}$.

Coverage values, $\boldsymbol{\alpha}(\mathbf{x})$, are computed for all spels in the set B . However, since B is, in general, not a one spel thick set, it may happen that some of its elements, which should be pure, are assigned partial coverage due to presence of noise. To reduce the impact of noise, we, in the fourth step of the algorithm, perform thinning of the set of mixed spels. We iteratively assign back the simply connected elements of B which are at a smallest distance to one of the crisp class vectors. This continues until the resulting set of spels constitute a thin boundary of a coverage representation, in the sense of Definition 3.6.

By this, a coverage segmentation of the observed image is obtained. Performance of this method is illustrated in Section 5.3.

4.3. Graph based coverage segmentation. Several efficient methods for image segmentation have been formulated in the framework of edge weighted graphs. The graph theoretic approach to image processing naturally leads to methods that are applicable to images of any dimension, and images sampled on non-Cartesian or spatially variant grids [16, 58]. An image is often associated with a graph by identifying each spel with a vertex in the graph, and defining edges of the graph so that they represent local adjacency between spels. Each edge in the graph may also be associated with a (real-valued) weight, reflecting the image content [15]. A segmentation of a graph is formulated either as a mapping from the vertices of the graph to some set of object categories, or in terms of graph cuts. Informally, a graph cut is a set of edges such that, if they are removed, the graph is separated into two or more components. The two representations—classification of vertices and separation by cuts—are closely related, and the choice of one representation over the other is largely a matter of preference. In any of the cases, the graph structure utilized in the task of image segmentation provides generality and wide applicability of the designed methods.

We were interested in developing a graph based segmentation method which results in coverage representation, or at least in its approximation; a main interest is to enable subsequent precise feature estimation. Commonly, a segmentation of a graph is only defined on the vertices of the graph, and it is traditionally crisp. Our approach presented in [33, 34] is to interpret the edges of the graph as paths between the vertices, and to assign membership labels also to the points along the edges of the graph to one or more object classes. Thereby, we obtain an *edge segmentation* of the graph. In relation to this, we have also introduced the concept of *located cuts*, which are graph cuts defined with sub-edge precision. Via the concept of *induced edge segmentation*, located cuts provide a convenient way of extending a segmentation defined on the vertices of the graph to all points along the edges of the graph. Finally, we have defined *vertex coverage segmentation* as a graph theoretic equivalent of coverage segmentation, and have presented a method for its approximate computation.

In the following we describe this idea in more details. Further information can be found in [33, 34]

A framework for sub-pixel segmentation on graphs. A *graph* is defined as an ordered pair $G = (V, E)$, consisting of vertices $v \in V$ and edges $e \in E \subseteq V \times V$. An edge spanning two vertices v_i and v_j is denoted by e_{ij} . Edges can be assigned weights, in which case we refer to a graph as an *edge weighted graph*. If $e_{ij} \in E$, the vertices v_i and v_j are *adjacent*. The set of vertices adjacent to a vertex v is denoted by $\mathcal{N}(v)$. For undirected graphs, an edge is an unordered pair $\{v_i, v_j\}$, i.e., $e_{ij} \equiv e_{ji}$.

A *path* is an ordered sequence of vertices $\pi = \langle v_1, v_2, \dots, v_k \rangle$ such that $v_{i+1} \in \mathcal{N}(v_i)$ for all $i \in [1, k - 1]$. Two vertices v and u are *linked* in G if there exists a

path π in G that starts at v and ends at u ; we write $v \sim_{G'} u$. If all pairs of vertices in G are linked, then G is *connected*, otherwise it is *disconnected*.

Let $G = (V, E)$, $S \subseteq E$, and $G' = (V, E \setminus S)$. If, for all $e_{ij} \in S$, it holds that $v_i \not\sim_{G'} v_j$, then S is a (*graph*) *cut* on G . For any cut $S \neq \emptyset$, the graph $(V, E \setminus S)$ is disconnected, i.e., it consists of two or more components.

In order to introduce more formally the main concepts of the framework, we start with definitions of vertex and edge segmentations.

Definition 4.1. A *vertex segmentation* \mathcal{V} of a graph $G = (V, E)$ into m components is a mapping $\mathcal{V} : V \rightarrow \mathbb{A}_m$.

In the general case, this is a fuzzy segmentation and each vector component $\mathcal{V}(v)_k$ in $\mathcal{V}(v)$ represents the degree to which the vertex v belongs to the corresponding class k .

Vertex segmentations and graph cuts are closely related. If the *boundary*, $\partial\mathcal{V}$, of a vertex segmentation \mathcal{V} is defined as the set of edges $\partial\mathcal{V} = \{e_{ij} \in E \mid \mathcal{V}(v_i) \neq \mathcal{V}(v_j)\}$, then the boundary of a vertex segmentation determines a cut on G .

We interpret edges as connected paths between the vertices. Let a point on an edge e_{ij} be specified by a parameter $t \in (0, 1)$, and let the vertices v_i and v_j be associated with $t = 0$, and $t = 1$, respectively (for undirected graphs, we assume that the vertices are indexed, and use the convention to associate $t = 0$ with the vertex having lower index). If every $v \in V$ is included in E at least once (i.e., there are no isolated vertices), then every point on a graph can be specified by a pair (e, t) , where $e \in E$ and $t \in [0, 1]$. In particular, points corresponding to vertices are of the form $(e, 0)$ or $(e, 1)$.

Definition 4.2. An *edge segmentation* \mathcal{E} of a graph $G = (V, E)$ is a mapping $\mathcal{E} : E \times [0, 1] \rightarrow \mathbb{A}_m$.

An edge segmentation \mathcal{E} is said to be *consistent* if all segmentation vectors associated with a vertex point (by its different edges) are equal. A vertex segmentation \mathcal{V} and an edge segmentation \mathcal{E} are said to be *consistent* if $\mathcal{E}(e_{ij}, 0) = \mathcal{V}(v_i)$ and $\mathcal{E}(e_{ij}, 1) = \mathcal{V}(v_j)$ for all $e_{ij} \in E$. If \mathcal{V} is a vertex segmentation and \mathcal{E} is an edge segmentation such that \mathcal{V} and \mathcal{E} are consistent, then we may view \mathcal{E} as an extension of \mathcal{V} from the set of vertices to the points along the edges of the graph.

An important concept of the framework is the one of *located cuts*. The idea is to increase the precision of the separation of the objects in the graph by specifying a point along each edge of a cut, indicating where the transition between the different objects occurs. We refer to such a “precise” cut as a *located cut*. Such a cut can define a segmentation where also the points along the edges of the cut are assigned to the separated components, as opposed to a classical graph cut, where the cut edges are left unassigned. We denote a located cut on an edge e_{ij} (of the classical cut) by $\mathcal{T}(e_{ij})$ and, using the introduced edge parametrisation, the location of the cut (the point of transition between the components) is conveniently expressed as a real value in $[0, 1]$.

Located cuts provide a natural way to define a particular type of edge segmentation, consistent with a given vertex segmentation, via the concept of *induced edge segmentation*.

Definition 4.3. Given a vertex segmentation \mathcal{V} , and location \mathcal{T} such that $(\partial\mathcal{V}, \mathcal{T})$ is a located cut, the *induced edge segmentation* $\mathcal{I}_{\mathcal{V}, \mathcal{T}}$ is

$$\mathcal{I}_{\mathcal{V}, \mathcal{T}}(e_{ij}, t) = \begin{cases} \mathcal{V}(v_i) & \text{if } e_{ij} \in \partial\mathcal{V} \text{ and } t < \mathcal{T}(e_{ij}) \\ \frac{1}{2}(\mathcal{V}(v_i) + \mathcal{V}(v_j)) & \text{if } e_{ij} \in \partial\mathcal{V} \text{ and } t = \mathcal{T}(e_{ij}) \\ \mathcal{V}(v_j) & \text{otherwise.} \end{cases}$$

Essentially, edges that belong to the cut are “divided” into two parts, as determined by the location of the cut, and the parts are then assigned to the two components as determined by the vertices at the ends of the cut.

An edge segmentation contains, in general, more information than a vertex segmentation. Our interest is to utilize this additional information to obtain precise feature measurements of segmented objects on the graph. However, there are two issues that we have to consider: (i) existing feature estimators are defined for vertex segmentations only, and therefore extraction of relevant information from edge segmentation requires appropriate adjustments; (b) depending on criteria used to define initial (fuzzy) vertex segmentation, this segmentation may, or may not, be appropriate for extraction of geometric features. An appropriate model for this purpose is, as already shown, the coverage based one. It is therefore of high practical interest to convert an edge segmentation, which does contain information related to geometrical properties of the object, to an appropriate vertex segmentation that can be used within the existing framework for feature extraction. We have, in [34] introduced the concept of *vertex coverage segmentation*, a graph theoretic equivalent of the concept of coverage segmentation, which is highly appropriate for precise feature extraction. Finally we have proposed an approach for computing a vertex coverage segmentation, that corresponds to a given edge segmentation.

Assuming no isolated vertices in the graph (a reasonable assumption for the envisioned applications in image processing), we define the *domain* of a vertex v_i as the set of points on the “half-edges” adjacent to the vertex (this can be seen as a graph-theoretical counterpart of a spel). Let \mathcal{E} be an edge segmentation of G into m -components. The vertex coverage segmentation $\mathcal{C}_{\mathcal{E}}$ of a vertex v_i is a vector of \mathbb{A}_m defined as

$$(4.3) \quad \mathcal{C}_{\mathcal{E}}(v_i) = \frac{1}{|\mathcal{N}(v_i)|} \sum_{j, v_j \in \mathcal{N}(v_i)} 2 \int_0^{\frac{1}{2}} \mathcal{E}(e_{ij}, t) dt,$$

for all $v_i \in V$. For an induced edge segmentation $\mathcal{I}_{\mathcal{V}, \mathcal{T}}$, the integral in the numerator of Eq. (4.3) can be written in closed form:

$$(4.4) \quad 2 \int_0^{\frac{1}{2}} \mathcal{I}_{\mathcal{V}, \mathcal{T}}(e_{ij}, t) dt = \begin{cases} 2\mathcal{T}(e_{ij})\mathcal{V}(v_i) + (1 - 2\mathcal{T}(e_{ij}))\mathcal{V}(v_j) & \text{if } \mathcal{T}(e_{ij}) < \frac{1}{2}, \\ \mathcal{V}(v_i) & \text{otherwise.} \end{cases}$$

Equations (4.3) and (4.4) provide the final required items for the following proposed processing chain: starting from a given (fuzzy) vertex segmentation, compute

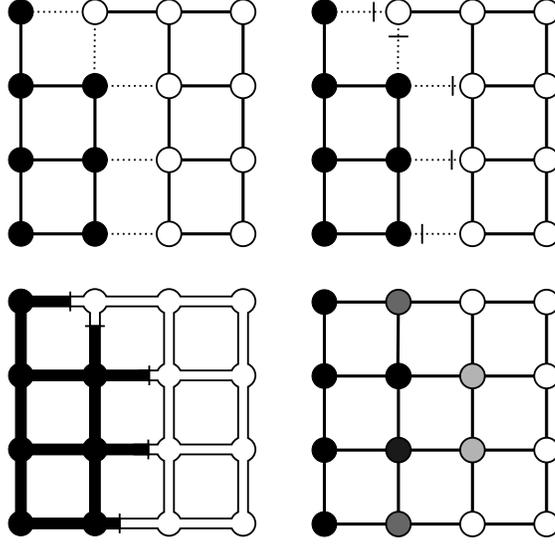


FIGURE 5. Notions of the proposed framework. (Top left) A crisp vertex segmentation \mathcal{V} of a graph. The boundary, $\partial\mathcal{V}$, of the segmentation is shown as dashed lines. (Top right) One possible located cut. (Bottom left) The edge segmentation $\mathcal{L}_{\mathcal{V},\mathcal{T}}$ induced by \mathcal{V} and \mathcal{T} . (Bottom right) One component of the corresponding vertex coverage segmentation $\mathcal{C}_{\mathcal{V},\mathcal{T}}$.

a located cut, extend the segmentation to an induced edge segmentation, from that compute a vertex coverage segmentation. Figure 5 illustrates different notions of the framework. The motivation for introducing these steps are, first of all, to reach a vertex coverage segmentation, providing highly improved feature estimates of imaged objects, but also, thanks to the division of the processing chain into individual and separately defined parts, to facilitate easy exchange of one step for another, providing flexibility of the approach and simplifying adjustments to fulfil task specific requirements.

A question that is not addressed here is how to compute located cuts, essential for defining an appropriate induced edge segmentation, and therefore for the final vertex coverage segmentation. One particular method for located cuts computation, applicable to *any* starting fuzzy segmentation, is suggested in [34]. It is based on an appropriate “reduction” of fuzziness of the initial vertex segmentation, so that only information relevant for the subsequent steps is preserved (i.e., precise location of the cut/object boundaries).

Evaluation of the proposed method contains results of area estimation of 2D synthetic objects obtained by the proposed segmentation method; area of an object is estimated as the sum of values assigned to vertices in the vertex coverage segmentation. Theoretical and empirical analysis of the results shows that, even

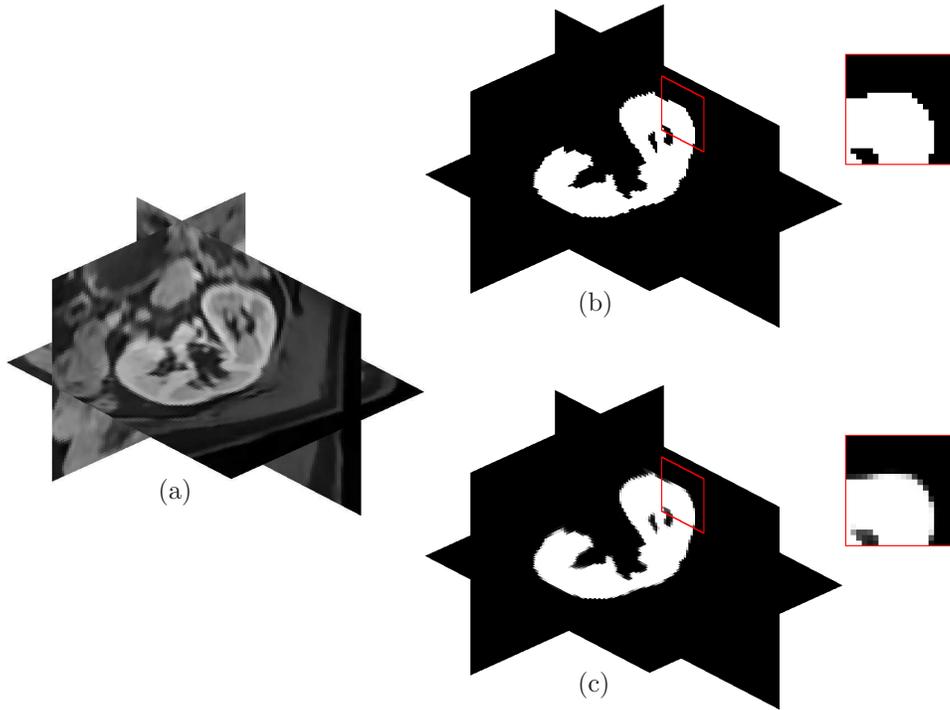


FIGURE 6. A kidney, segmented from an MR volume image of a human abdomen using the Relaxed IFT method [35]. (a) Original grey-scale image volume. (b) Crisp segmentation. (c) Vertex coverage segmentation.

though the convergence rate of the estimate is the same as in the crisp case, the area estimation error obtained by the proposed method is, for every given resolution, significantly smaller than the error obtained from a crisp representation. The method has also been used for segmentation of real medical images; an example of a kidney segmentation is shown in Figure 6, highlighting the difference between the classical crisp segmentation, and the proposed method.

Clearly, the true coverage model for object representation provides higher precision of geometric feature estimates, where convergence rate is proved to be higher than for the crisp case. The proposed graph-theoretical based method is, however, only an approximation of the coverage model, since information about fuzzy memberships over a finite number of one-dimensional sets of points (the edges) is used instead of information about memberships over an nD set of points (the spel). However, generality and applicability of the method are its appealing advantages, compensating for its somewhat lower accuracy.

5. Feature extraction

A coverage segmentation preserves more information about crisp original objects than a corresponding crisp segmentation. This additional information may be highly beneficial in subsequent processing steps, e.g., when estimating features of continuous imaged objects. It is reasonable to expect, however, that methods for feature extraction have to be more or less adjusted to be suited for application to coverage representations of objects.

It is important to notice that digital image analysis aims at measuring features of continuous (real world) objects on the basis of their digital images. Consequently, such measurements, derived from digital shapes, can only be estimates, since the interest is seldom in the features of the digital object, but rather in the original object that has been digitized. Consequently, an important task when designing an estimator is to evaluate its performance, preferably by providing some relevant error bounds.

The possibility to increase precision of estimates of various properties of a continuous original shape by utilizing the information available in a fuzzy representation was studied first for representations based on rather general fuzzy membership functions. Methods derived for estimation of perimeter, area, surface area, volume, geometric moments, signature of a shape, from a fuzzy representation, are presented in [4, 49, 55]. These publications contain statistical studies demonstrating improvement in precision of the estimates, as compared to estimates from a crisp representation. However, the generality of fuzzy membership functions considered, is prohibiting derivation of stronger theoretical statements about the developed methods. We therefore introduced the proposed constraints to the fuzzy membership assignment and the strict interpretation of membership values of the coverage model. This approach, which corresponds well to the outcome of many imaging devices, was, in our opinion, natural to be used. Most importantly, the imposed restriction enabled theoretical derivation of error bounds for a number of feature estimates.

Since a coverage representation is a special case of a fuzzy representation, the previously developed methods for feature estimation from a fuzzy representation are still applicable. For the case of estimation of geometric moments, the proposed method for general fuzzy sets works excellently also for the special case of coverage representations, and we were able to prove that the error decreases to zero at a rate faster than for the crisp case [50]. However, for the proposed perimeter and surface area estimators, despite providing on average much improved estimates, a general faster convergence to the true error was not possible to prove. Instead of starting from the definition of perimeter of a fuzzy set, but rather fully utilizing the knowledge that membership values correspond to pixel coverage, we have derived a method that does provide faster convergence to the true value [52].

We have analysed the accuracy of the estimation of geometric moments, when they are calculated from different representations of a shape. We show that the order of the error can be reduced if the estimation is based on a coverage representation; use of such a representation can therefore be an alternative to increasing the

spatial resolution of the image. Geometric moments of objects provide information about area, (hyper-)volume, centroid, principal axes, and a number of other features of the shape. In addition, we were interested in estimation of perimeter and surface area. The results obtained for perimeter estimation are published in [52]; surface area estimation remains our future work. (Note, however, that the method developed in [55] for general fuzzy sets, provides empirically very good surface estimates when applied to a coverage representation.)

Our main results, presented in [50, 52], and given in a fairly detailed description in the following subsections, are upper bounds for the estimation error as functions of spatial and coverage resolution. We have confirmed that inter-relations between these two types of resolutions affect the precision of estimation, and that one of the resolutions can, to some extent, be used to compensate for the other. It is usually the case that spatial resolution is given by the imaging device and cannot be changed, whereas improved intensity information, or simply better utilization of grey-levels, already at hand, may be much more easily accessible. Our main message is therefore that *appropriate utilization of intensity information available provides an excellent way to increase estimation precision*, and we specially notice that this applies also for fixed spatial resolution. We find this to be a very useful, applicable, and important result.

Before giving more detailed presentation of methods derived for estimation of geometric moments and perimeter from coverage representation of shapes, we give a brief introduction to main notions and tasks of shape description and analysis.

5.1. Shape analysis. The shape of an object is a representation of its geometric extent. It can be thought of as a silhouette of the object. It is often referred to as a region. The shape of an object is invariant to geometric transformations such as translation, rotation, (uniform) scaling, and reflection. Therefore, shape can be understood as an equivalence class in the set of objects; two objects are equivalent (i.e., have the same shape) if there exists a series of translations, rotations, scalings, and reflections that maps one of them to the other. There are many situations where image analysis can be reduced to the analysis of shapes, which gives high importance to the field of shape analysis.

There exist different classifications of shape analysis techniques, see, e.g., [31] for an overview. Depending whether only the shape boundary points are used for the description, or alternatively, the whole interior of a shape is used, the two resulting classes of algorithms are known as boundary-based (external) and region-based (internal), respectively. Examples of the former class are algorithms which parse the shape boundary and various Fourier transforms of the boundary. Their main advantages are reduction of data and that they may offer a compact description of complex forms. Region-based methods include, e.g., the medial axis transform, moment-based approaches, and methods of shape decomposition into the primitive parts. Their main advantages are easier characterisation and stability in practical applications, where there is unavoidable noise.

A *description* of a shape is data representing it in a way suitable for further computer processing. Such data can be low-dimensional (perimeter and moments),

or high-dimensional (medial axis and primitive parts). The first type of data is suitable for, e.g., shape classification, while the second, often called shape *representation*, provides good visual interpretation and compression.

The goal of a shape description is to uniquely characterize the shape. Desired properties of a shape description scheme are invariance to translation, scale, and rotation; these three transformations, by definition, do not change the shape of an object, and consequently should not change its descriptor. However, it should be noted that in the discrete case such invariance exists only up to discretization effects, and special care must often be taken in order to fulfil it.

Additional desired properties of a good shape description method are [31]:

- accessibility—How easy is it to compute a descriptor in terms of memory requirements and computational time; are the operations local or global?
- scope—How wide is the class of shapes that can be described by the method?
- uniqueness—Is the representation uniquely determined for a given shape?
- information preservation—Is it possible to recover the shape from its descriptor?
- stability and sensitivity—How sensitive is a shape descriptor to small changes of a shape?

Descriptors usually perform well regarding some of the listed properties, while failing regarding some others. Therefore, a common approach is to combine them in some appropriate way, to achieve a description that fulfils the requirements of a given task.

5.2. Geometric moments. The two-dimensional Cartesian moment, $m_{p,q}$ of a function $f(x, y)$ is defined as

$$m_{p,q} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) x^p y^q dx dy,$$

for integers $p, q \geq 0$. The moment $m_{p,q}(S)$ has the order $p + q$. Cartesian moments are often referred to as *geometric moments*. The geometric moment $m_{p,q}$ can be seen as the projection of $f(x, y)$ to the monomial basis set $x^p y^q$. A *complete moment set* of order n consists of all moments $m_{p,q}$ such that $p + q \leq n$.

Moments were made popular in image analysis by Hu, [18]. Hu's *Uniqueness Theorem* states that for a piece-wise continuous function $f = f(x, y)$, nonzero only within a bounded set in \mathbb{R}^2 , the moments of all orders exist. Moreover, the set of moments of a function f is uniquely determined by f , and conversely, the set of all moments of f uniquely determines f . In order to utilize geometric moments as shape descriptors, their behaviour under scaling, translation, rotation, and reflection has been studied. To provide shape descriptors which are invariant to scale, translation, and rotation of a shape, Hu defined seven nonlinear combinations of geometric moments up to order three, which are known as *absolute moment invariants*.

When used in image analysis, moments are calculated for discrete functions on discrete bounded domains. The definition of a geometric moment $m_{p,q}$ of a digital

image $f(x, y)$ is $m_{p,q} = \sum_i \sum_j f(i, j) i^p j^q$, where (i, j) are points in the (integer) sampling grid.

An image is always bounded and an image function is piece-wise continuous, which ensures that the Hu's statement holds for digital images. However, potentially very large number of moments (as many as there are pixels in the image itself) may be required for a unique representation and reconstruction of a digital image. To make the description practically feasible, a smaller set of moments has to be used, and consequently, only an approximate reconstruction can be provided. An important question is how to make an appropriate selection of moments, such that sufficient information is provided for a good enough characterization of the image.

Even though the nonorthogonality of the basis monomials $x^p y^q$ causes some undesired properties of geometric moments, which has initiated several alternative approaches to moment based shape description (e.g., Legendre and Zernike polynomials are defined on orthogonal basis sets, which provides more stable and simpler reconstruction), the Cartesian geometric moments are still well accepted shape descriptors, due to their simple definition, their uniqueness for a given shape, the possibility to derive descriptors invariant to rotation, translation, and scaling, and to express them as integers, their linearity, and the possibility to reconstruct a number of features of a shape from an appropriately chosen set of its moments. In addition, it is possible to express all other types of moments in terms of geometric moments.

The disadvantage of sensitivity to noise mostly applies to high-order moments. Their use in object description can be avoided (or, at least, reduced) by, e.g., first decomposing complex shapes into simpler and more regular parts, which can, then, be represented by a smaller set of lower order moments. Still, precision of estimated moments are of highest importance; decomposition of objects into smaller parts leads to moments computed for smaller objects, and, correspondingly, estimations based on fewer spels. If a special care is not taken, discretization errors may accumulate, cancelling the positive effect of the decomposition.

Our focus is on analysis of the errors that result from estimation of moments of a continuous shape from the corresponding moments of its crisp and different fuzzy digitizations. In particular, we have studied objects represented according to the pixel coverage model, with an aim to explore relations between coverage and spatial resolution and their influence on precision of object representation.

Moments estimated from a Gauss digitization.

Definition 5.1. The p_1, p_2, \dots, p_n -moment of a crisp bounded set S in the n -dimensional Euclidean space equipped with the Cartesian coordinate system is

$$m_{p_1, p_2, \dots, p_n}(S) = \iint \dots \int_S \prod_{i=1}^n x_i^{p_i} dx_1 dx_2 \dots dx_n,$$

for integers $p_1, p_2, \dots, p_n \geq 0$. The moment $m_{p_1, p_2, \dots, p_n}(S)$ has the order $q = \sum_{i=1}^n p_i$.

If the set S is inscribed into an integer grid and digitized, instead of the moments $m_{p_1, p_2, \dots, p_n}(S)$, the moments of the Gauss digitization $\mathcal{D}_g(S)$ are available.

Definition 5.2. The *discrete moment* $\tilde{m}_{p_1, p_2, \dots, p_n}(S)$ of a crisp set S is

$$\tilde{m}_{p_1, p_2, \dots, p_n}(S) = \sum_{\mathbf{x} \in \mathcal{D}_g(S)} \prod_{d=i}^n x_i^{p_i},$$

where $\mathbf{x} = (x_1, x_2, \dots, x_n)$, and $\mathcal{D}_g(S)$ is given by Definition 3.4.

Several features of a shape can be calculated from a sufficient number of its moments. In fact, a shape can be recovered from an appropriately chosen set of its moments. If continuous moments are replaced by their discrete counterparts, more or less good estimates of the observed features of a continuous shape can be obtained. An upper bound of the error introduced when approximating continuous moments by their crisp counterparts, can be derived from a theorem by Davenport [7].

To avoid problematic cases, the following restrictions are imposed on the set S , which is assumed to be an n -dimensional closed and bounded set of points.

- I Any line parallel to one of the n coordinate axes intersects S in a set of points which, if not empty, consists of at most h intervals.
- II The same is true (with m in place of n) for any of the m -dimensional regions obtained by projecting S on one of the coordinate spaces defined by equating a selection of $n-m$ of the coordinates to zero; and this condition is satisfied for all m from 1 to $n-1$.

Theorem 5.1. [7] *If S satisfies the conditions I and II, then*

$$|\mathcal{D}_g(S)| - |S| \leq \sum_{m=0}^{n-1} h^{n-m} S_m,$$

where S_m is the sum of the m -dimensional volumes of the projections of S on the various coordinate spaces by equating any $n-m$ coordinates to zero, and $S_0 = 1$ by convention.

It is desirable to know how the accuracy of approximation changes with a change of image resolution. Instead of increasing the resolution of the digitization grid, we keep the integer grid, and instead scale the set S . Let rS denote a scaling of the continuous set S about the origin by the factor r : $rS = \{(rx, ry) \mid (x, y) \in S\}$. Since h does not change with scale, and $V_m = \mathcal{O}(r^m)$, it is easy to see that $|S| = r^{-n}|\mathcal{D}_g(rS)| + \mathcal{O}(1/r)$, or, expressed in terms of moments, that

$$(5.1) \quad m_{0,0,\dots,0}(S) = \frac{1}{r^n} \tilde{m}_{0,0,\dots,0}(rS) + \mathcal{O}\left(\frac{1}{r}\right).$$

Observing that a first order moment of a continuous n -dimensional set S , can be expressed as the zero-order moment of a set S' in an $(n+1)$ -dimensional space, Davenport's theorem can, by induction, be generalized to moments of arbitrary order $q \in \mathbb{N}$.

Theorem 5.2. *If the closed and bounded set S satisfies the conditions I and II, then*

$$(5.2) \quad m_{p_1, p_2, \dots, p_n}(S) = \frac{1}{r^{n+q}} \tilde{m}_{p_1, p_2, \dots, p_n}(rS) + \mathcal{O}\left(\frac{1}{r}\right),$$

where $q = \sum_{i=1}^n p_i$, $q \in \mathbb{N}$.

Proof by induction. The base case, for zero order moments $q = 0$, is given by (5.1). Inductive step: Assume that (5.2) holds for an arbitrary moment of order q . We then show that (5.2) also holds for a moment of order $q + 1$, were we increase the exponent of the k th coordinate, p_k , by one.

$$\begin{aligned} m_{p_1, p_2, \dots, p_k+1, \dots, p_n}(S) &= m_{p_1, p_2, \dots, p_n}(S') \\ &= \frac{1}{r^{n+q+1}} \tilde{m}_{p_1, p_2, \dots, p_n}(rS') + \mathcal{O}\left(\frac{1}{r}\right) \\ &= \frac{1}{r^{n+q+1}} \tilde{m}_{p_1, p_2, \dots, p_k+1, \dots, p_n}(rS) + \mathcal{O}\left(\frac{1}{r}\right), \end{aligned}$$

where

$$S' = \{(x_1, x_2, \dots, x_n, x_{n+1}) \mid (x_1, x_2, \dots, x_n) \in S, x_{n+1} \in [0, x_k]\}.$$

The first equality holds from the definition of moments, the second is given by the assumption, and the third holds by noticing that for every integer point in rS , there are $rx_k + \mathcal{O}(1)$ integer points in rS' . \square

For a class of 2D shapes, a stronger statement related to error bounds is derived in [22].

Theorem 5.3. [22] *The moments of a planar 3-smooth convex set S , digitized in a grid with resolution r (the number of grid points per unit), can be estimated by*

$$m_{p_1, p_2}(S) = \frac{1}{r^{p_1+p_2+2}} \tilde{m}_{p_1, p_2}(rS) + \mathcal{O}\left(\frac{1}{r^{\frac{15}{11}-\varepsilon}}\right)$$

for $p_1 + p_2 \leq 2$

Remark 5.1. A planar 3-smooth convex set is a convex set in the Euclidean plane whose boundary consists of a finite number of arcs having continuous third order derivatives and a positive curvature at every point, except the end points of the arcs. These conditions exclude the existence of straight boundary segments.

Remark 5.2. Despite being given for convex sets, Theorem 5.3 also holds for finite unions and intersections of convex sets.

Moments estimated from a coverage digitization. Fuzzy moments and the centre of gravity of a fuzzy set are among the first defined fuzzy concepts.

Definition 5.3. The p_1, p_2, \dots, p_n -moment of a fuzzy subset S of a reference set $X \subset \mathbb{R}^n$ is

$$m_{p_1, p_2, \dots, p_n}(S) = \int \int \dots \int_X \mu_S(\mathbf{x}) \prod_{i=1}^n x_i^{p_i} dx_1 dx_2 \dots dx_n,$$

where $\mu_S(\mathbf{x})$ is the membership of the point \mathbf{x} to the set S .

The membership function μ , defining a fuzzy set, can be any mapping $X \rightarrow [0, 1]$. We further study the special case where the membership of a point \mathbf{x} is defined by the coverage $\alpha(\mathbf{x})$ of the corresponding spel $\sigma(\mathbf{x})$.

Replacing the Gauss digitization in Definition 5.2 with a coverage digitization, and replacing membership μ in Definition 5.3 with coverage values α , the following three definitions follow naturally.

Definition 5.4. The *discrete coverage moment* $\mathcal{M}_{p_1, p_2, \dots, p_n}(S)$ of a crisp set $S \in \mathbb{R}^n$ is

$$\mathcal{M}_{p_1, p_2, \dots, p_n}(S) = \sum_{(\mathbf{x}, \alpha(\mathbf{x})) \in \mathcal{D}_c(S)} \alpha(\mathbf{x}) \prod_{i=1}^n x_i^{p_i},$$

where $\mathbf{x} = (x_1, x_2, \dots, x_n)$, and $\mathcal{D}_c(S)$ is given by Definition 3.2.

Definition 5.5. The ℓ -*level quantized discrete coverage moment* $\mathcal{M}_{p_1, p_2, \dots, p_n}^\ell(S)$ of a crisp set $S \in \mathbb{R}^n$ is

$$\mathcal{M}_{p_1, p_2, \dots, p_n}^\ell(S) = \sum_{(\mathbf{x}, \alpha^\ell(\mathbf{x})) \in \mathcal{D}_c^\ell(S)} \alpha^\ell(\mathbf{x}) \prod_{i=1}^n x_i^{p_i},$$

where $\mathcal{D}_c^\ell(S)$ is given by Definition 3.3.

Definition 5.6. The r -*sampled discrete coverage moment* $\hat{\mathcal{M}}_{p_1, p_2, \dots, p_n}^r(S)$ of a crisp set $S \in \mathbb{R}^n$ is

$$\hat{\mathcal{M}}_{p_1, p_2, \dots, p_n}^r(S) = \sum_{(\mathbf{x}, \hat{\alpha}^r(\mathbf{x})) \in \hat{\mathcal{D}}_c^r(S)} \hat{\alpha}^r(\mathbf{x}) \prod_{i=1}^n x_i^{p_i},$$

where $\hat{\mathcal{D}}_c^r(S)$ is given by Definition 3.5.

In the following we derive error bounds for estimation of moments of a continuous shape from its discrete coverage moments. Theorems are formulated for the n -dimensional case, extending the results presented in [50].

Given a closed and bounded set $S \subset \mathbb{R}^n$, satisfying conditions I and II, we compare the r_f -sampled coverage moment $\hat{\mathcal{M}}^{r_f}(r_s S)$ of an r_s times dilated set S , with the crisp moment $\tilde{m}_{p_1, p_2, \dots, p_n}(r_s r_f S)$ of an r_f times further dilated set (corresponding to r_f times higher image resolution).

Theorem 5.4. *The discrete moments (Definition 5.2) of a set $r_s r_f S \subset \mathbb{R}^n$, can be estimated by the r_f -sampled coverage moments of a set $r_s S$ by*

$$\tilde{m}_{p_1, p_2, \dots, p_n}(r_s r_f S) = r_f^{q+n} \hat{\mathcal{M}}_{p_1, p_2, \dots, p_n}^{r_f}(r_s S) + \mathcal{O}(r_s^{q+n-2} r_f^{q+n}),$$

where $q = p_1 + p_2 + \dots + p_n$ is the order of the moment.

Combining this result with Theorems 5.2 and 5.3 leads to the following corollary (extending [50]):

Corollary 5.1. *The moments of a closed and bounded set $S \subset \mathbb{R}^n$, satisfying conditions I and II, can be estimated by*

$$(5.3) \quad m_{p_1, p_2, \dots, p_n}(S) = \frac{1}{r_s^{q+n}} \hat{\mathcal{M}}_{p_1, p_2, \dots, p_n}^{r_f}(r_s S) + \mathcal{O}\left(\frac{1}{r_s^2}\right) + \mathcal{O}\left(\frac{1}{r_s r_f}\right),$$

where $q = \sum_{i=1}^n p_i, q \in \mathbb{N}$. The moments of a planar 3-smooth convex 2D shape S can, for $p_1 + p_2 \leq 2$, be estimated by

$$(5.4) \quad m_{p_1, p_2}(S) = \frac{1}{r_s^{p_1+p_2+2}} \hat{\mathcal{M}}_{p_1, p_2}^{r_f}(r_s S) + \mathcal{O}\left(\frac{1}{r_s^2}\right) + \mathcal{O}\left(\frac{1}{(r_s r_f)^{\frac{15}{11}-\varepsilon}}\right).$$

We note that Theorem 5.4 and Corollary 5.1 also hold for estimations based on ℓ -level quantized discrete coverage moments $\mathcal{M}_{p_1, p_2, \dots, p_n}^\ell(r_s S)$, with $\ell = r_f^n$. This follows from the fact that the coverage values of an ℓ -level quantized coverage digitization, with $\ell = r_f^n$, do not differ from the values assigned by a real coverage digitization more than the (corresponding) values assigned by an r_f -sampled coverage digitization (see Section 3.1). We conclude that once the spatial resolution is high enough to fully “exploit” the coverage values of spels, i.e., when $r_s > Cr_f$, where C is a constant derived from the asymptotic expression for the error bound, using r_f^n coverage values provides the same accuracy of moment estimation as increasing the (crisp) spatial resolution of the image r_f times.

Proof of Theorem 5.4. Without loss of generality, we assume that S is fully contained within $[0, 1]^n$. The $r_s r_f$ times dilated shape $r_s r_f S$, then fits in $I_D = [0, r_s r_f]^n$. Let us partition this region (image domain) into r_s^n nonoverlapping blocks of size r_f^n . Each such part of the image space can be expressed as an r_f times dilated spel, such that $\bigcup_{\mathbf{x} \in I_d} r_f \sigma(\mathbf{x}) = I_D$, where $I_d = \{\frac{1}{2}, \frac{3}{2}, \dots, \frac{2r_s-1}{2}\}^n$ is the set of half-integer points of the r_f times smaller domain $[0, r_s]^n$. The moment \tilde{m} can be computed as a sum of the moments of all such parts of the image space.

$$(5.5) \quad \tilde{m}_{p_1, p_2, \dots, p_n}(r_s r_f S) = \sum_{\mathbf{x} \in I_d} \tilde{m}_{p_1, p_2, \dots, p_n}(r_s r_f S \cap r_f \sigma(\mathbf{x}))$$

Assume that for a block $r_f \sigma(\mathbf{x}), \mathbf{x} \in I_d$, there are k out of the r_f^n spels which have their centroids within the continuous crisp shape $r_s r_f S$. The moment of such a block is

$$\tilde{m}_{p_1, p_2, \dots, p_n}(r_s r_f S \cap r_f \sigma(\mathbf{x})) = \sum_{j=1}^k \prod_{i=1}^n (r_f x_i + r_f \Delta_{i,j})^{p_i} = r_f^q \sum_{j=1}^k \prod_{i=1}^n (x_i + \Delta_{i,j})^{p_i},$$

where $r_f \Delta_{i,j}$, denotes the i th coordinate of the displacement of the j th covered spel with respect to the centre of the block, $r_f \mathbf{x}$. $\Delta_{i,j}$ takes values in the range $[-\frac{1}{2}, \frac{1}{2}]$.

We consider two different cases: (i) the block $r_f \sigma(\mathbf{x})$ is completely covered by the set $r_s r_f S$ and $k = r_f^n$; (ii) the block is partly covered and $1 \leq k \leq r_f^n$. The moment of an empty block, $k = 0$, is correctly estimated as zero, and does not contribute to any estimation error.

Case (i): $k = r_f^n$. Since the spels are symmetrically distributed around \mathbf{x} , then for each spel j , there is a corresponding spel j' such that $\Delta_{i,j'} = -\Delta_{i,j}$, for all i . We count both contributions each time and divide by two.

$$\tilde{m}_{p_1, p_2, \dots, p_n}(r_s r_f S \cap r_f \sigma(\mathbf{x})) = \frac{1}{2} r_f^q \sum_{j=1}^{r_f^n} \left(\prod_{i=1}^n (x_i - \Delta_{i,j})^{p_i} + \prod_{i=1}^n (x_i + \Delta_{i,j})^{p_i} \right)$$

For $x_i = \mathcal{O}(r_s)$, and $a_i \in \mathbb{R}$, such that $x_i \gg a_i$, it holds that

$$\begin{aligned}
& \prod_{i=1}^n (x_i - a_i)^{p_i} + \prod_{i=1}^n (x_i + a_i)^{p_i} \\
&= \prod_{i=1}^n \left(x_i^{p_i} - a_i x_i^{p_i-1} + \mathcal{O}(r_s^{p_i-2}) \right) + \prod_{i=1}^n \left(x_i^{p_i} + a_i x_i^{p_i-1} + \mathcal{O}(r_s^{p_i-2}) \right) \\
&= \prod_{i=1}^n x_i^{p_i} - \sum_{j=1}^n a_j x_j^{p_j-1} \prod_{\substack{i=1 \\ i \neq j}}^n x_i^{p_i} + \mathcal{O}(r_s^{q-2}) + \prod_{i=1}^n x_i^{p_i} + \sum_{j=1}^n a_j x_j^{p_j-1} \prod_{\substack{i=1 \\ i \neq j}}^n x_i^{p_i} + \mathcal{O}(r_s^{q-2}) \\
&= 2 \prod_{i=1}^n x_i^{p_i} + \mathcal{O}(r_s^{q-2})
\end{aligned}$$

And, therefore,

$$\begin{aligned}
\tilde{m}_{p_1, p_2, \dots, p_n}(r_s r_f S \cap r_f \sigma(\mathbf{x})) &= r_f^q \sum_{j=1}^{r_f^n} \left(\prod_{i=1}^n x_i^{p_i} + \mathcal{O}(r_s^{q-2}) \right) \\
&= r_f^{q+n} \prod_{i=1}^n x_i^{p_i} + \mathcal{O}(r_s^{q-2} r_f^{q+n}) \\
&= r_f^{q+n} \hat{M}_{p_1, p_2, \dots, p_n}^{r_f}(r_s S \cap \sigma(\mathbf{x})) + \mathcal{O}(r_s^{q-2} r_f^{q+n})
\end{aligned}$$

Case (ii): $k = \mathcal{O}(r_f^n)$. To cover the worst case, we assume that all covered spels are at maximal distance from the centre of the block; $\Delta_{i,j} = \frac{1}{2}$.

$$\begin{aligned}
\tilde{m}_{p_1, p_2, \dots, p_n}(r_s r_f S \cap r_f \sigma(\mathbf{x})) &\approx r_f^q \sum_{j=1}^k \prod_{i=1}^n \left(x_i + \frac{1}{2} \right)^{p_i} \\
&= r_f^q \cdot k \prod_{i=1}^n x_i^{p_i} + \mathcal{O}(r_s^{q-1} r_f^{q+n}) \\
&= r_f^{q+n} \hat{M}_{p_1, p_2, \dots, p_n}^{r_f}(r_s S \cap \sigma(\mathbf{x})) + \mathcal{O}(r_s^{q-1} r_f^{q+n})
\end{aligned}$$

For a closed and bounded set S , satisfying conditions I and II, there are $\mathcal{O}(r_s^n)$ blocks of type (i) and $\mathcal{O}(r_s^{n-1})$ blocks of type (ii) in the sum of Eq. (5.5), leading to the final result:

$$\tilde{m}_{p_1, p_2, \dots, p_n}(r_s r_f S) = r_f^{q+n} \hat{M}_{p_1, p_2, \dots, p_n}^{r_f}(r_s S) + \mathcal{O}(r_s^{q+n-2} r_f^{q+n}).$$

□

Statistical study on synthetic test images. We perform a statistical study to examine the properties of moments estimated at low resolutions. Multigrid resolution is expressed by dilations of the observed objects. Tests are performed on squares and disks of increasing size. For each of the observed real-valued side lengths a large number of randomly positioned squares (in various rotations) are considered. Similarly, for each of the observed real-valued radii, a large number of disks with random centre position, are observed. The continuous objects are digitized using

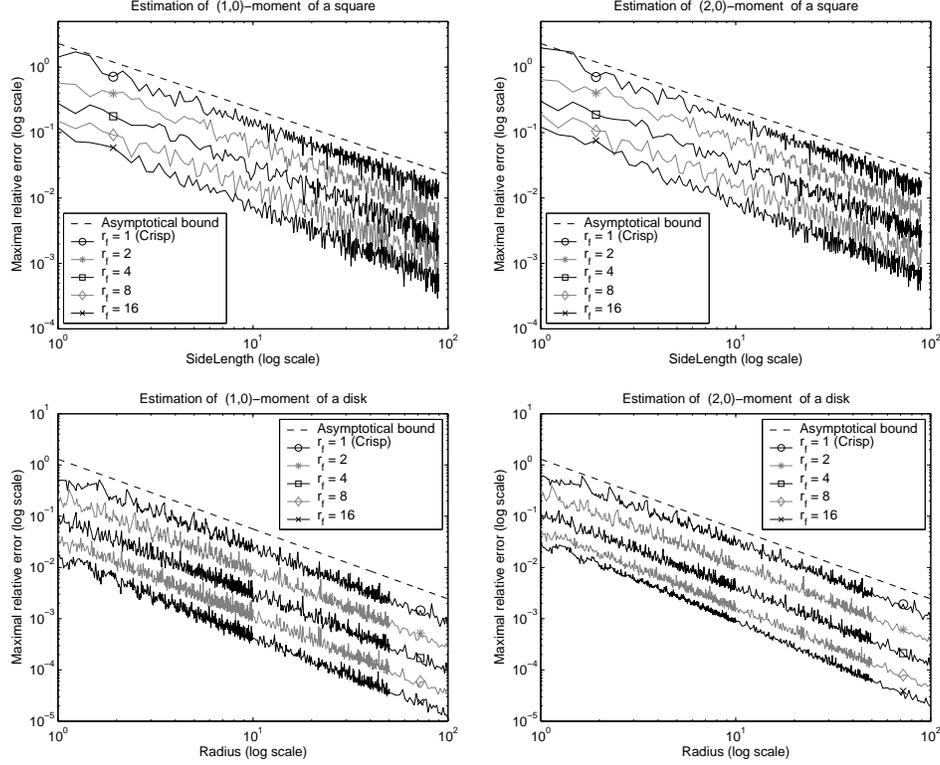


FIGURE 7. Plots of maximal observed errors for first and second order moments estimation for different spatial and coverage resolutions. *Top*: Moment estimation of a square. *Bottom*: Moment estimation of a disk.

r_f -sampled coverage digitization, with different super-sampling levels. Note that $r_f = 1$ corresponds to crisp segmentation, and that $r_f = 16$ approximates the upper limit for the coverage resolution of 8-bit pixel depth.

For each size of an object, we determine the maximal relative estimation error for moments up to the order two. We present the results for $m_{1,0}$ and $m_{2,0}$ moments estimation, both for squares and for disks, in Figure 7. The estimation errors for squares show asymptotic behaviour in accordance with expression (5.3). Disks are 3-smooth convex objects, and the corresponding estimation bounds agree with expression (5.4).

Plots are presented in a logarithmic scale so that the “slopes” of the curves correspond to the order of estimation error, and can be compared with the plotted straight line which has a slope equal to the theoretically derived order of error (-1 for squares and $-\frac{15}{11}$ for disks). Note that the plots show accordance with the asymptotic bounds also at low spatial resolutions. The relative positions of

the curves clearly show how the estimation error becomes smaller both with the increase of spatial and coverage resolution.

5.3. Perimeter. The length of the boundary (perimeter) of an object in a two dimensional (2D) image is an essential object features in image analysis. Despite its apparent simplicity, it is a feature that is very difficult to accurately compute from the information provided in a digital image. Accordingly, a large number of publications have addressed the issue of achieving accurate and precise estimates of object perimeter.

Most methods presented in the literature deal with binary images, where pixels either have a value one, being assigned to the object, or zero, if they are assigned to the background; for an overview, see, e.g., [6] and [9]. The binary model corresponds well with the output of the Gauss centre point digitization scheme. It, however, discards a large amount of useful information, especially along the object boundary.

In this section we show how perimeter of a continuous object can be accurately estimated from its coverage digitization. For the case of a quantized coverage digitization, we derive optimal scale factors, minimizing the maximal estimation error for straight edge segments. Both proven theoretically, for straight edges, and observed empirically, for more general shapes, a significant improvement in the accuracy and precision of perimeter estimates is achieved by utilizing the coverage information.

Background. The length of the boundary of a digitized object can be estimated as the cumulative sum of the lengths of local steps along the border of the object. Such estimates are straightforward to accomplish by summing the distances between pixel centres as determined from the Freeman chain code [14]. Doing so, however, results in rather big overestimates; the (intuitive) local step weights, 1 for isothetic and $\sqrt{2}$ for diagonal steps, are not optimal when measuring digitized line segments, this is illustrated in Figure 8.

Starting from an assumption that the boundary of an object is locally planar, optimal weights for the local steps along the border have been derived, [23, 39], leading to improved perimeter estimates. Weights for the 2D case, optimized to provide an unbiased estimator with minimal mean square error for straight lines with length tending to infinity, have been proven to perform even better for curved contours [9]. This last property is important to notice since it provides much more general applicability of the estimator.

In addition to the local type of estimators mentioned above, different nonlocal perimeter estimators have been developed; see e.g [6] for an overview. By basing the estimate on information from larger regions of the image, nonlocal estimators can be made to ensure convergence toward the true value, as the spatial grid resolution increases [22]. Such estimators are often referred to as multigrid convergent estimators. A common approach for nonlocal perimeter estimation is to recognize straight boundary segments, and to perform a polygonalization of the object. In spite of the fact that local methods can not be made multigrid convergent in a general sense (see e.g., [59]), they are still often preferred to nonlocal ones, due to their several advantages. Local methods are relatively easy to implement, parallelizable

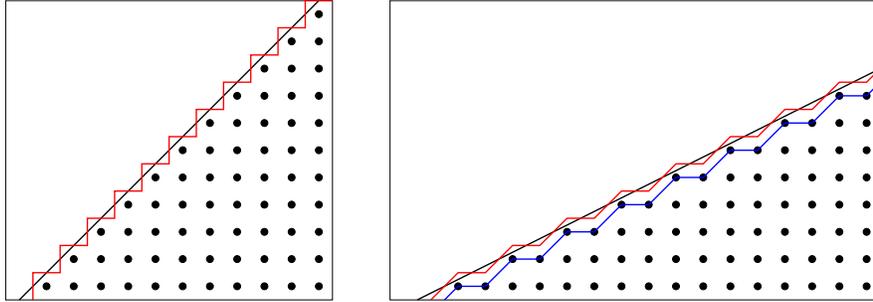


FIGURE 8. Estimation of the length of a straight edge using local steps, (left) using 4 directions, and (right) using 8 directions. Direct use of Euclidean lengths of the steps leads to estimates which are too large for certain directions, due to the shown staircase effect. Errors can be reduced by scaling the estimate with an appropriate factor. However, the estimate remains variant to rotation, and the maximal error (for the better, 8 direction version) is still almost 4%. This error does not decrease with increased image resolution.

(enabling very fast implementations), and inherently stable (in a sense that a small change in an image causes only a small change of the estimate). These important properties do not in general hold for nonlocal methods.

In this section we present an algorithm, proposed in [67], for estimating the boundary length of a continuous object from its coverage digitization. The method uses only local data and a parallel implementation is straightforward. Effects of quantization of coverage values are considered and the optimal scale factor for the (Freeman-style) cumulative sum of local steps is derived, as a function of the number of coverage levels available. The maximal error (difference from the Euclidean length of the original continuous line segment) is minimized for digital straight segments with the length tending to infinity. The method is applied and evaluated on objects with nonstraight boundaries as well. The issue of a trade-off between spatial and coverage resolution for a good performance of the estimator is explored by observing the performance of the method on shapes digitized at increasing spatial resolution, for a range of coverage resolutions. The results show that the accuracy and precision of estimates rapidly increase with the increase of coverage resolution, once a “reasonable” spatial resolution is provided.

Edge length estimation based on difference of column sums.

Non-quantized case A well known formula for computing the arc length of a function $y = f(x)$ over an interval $[a, b]$ is $l = \int_a^b \sqrt{1 + [f'(x)]^2} dx$. Applied to a linear function, $y = kx + m$, with $k \in [0, 1]$, this formula gives the length l of a line segment for $x \in [0, N]$, $N > 0$ as

$$(5.6) \quad l = N\sqrt{1 + k^2} .$$

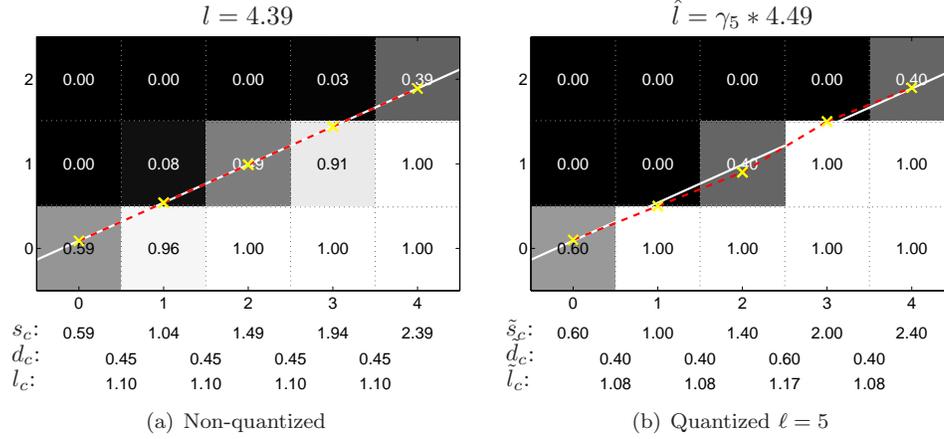


FIGURE 9. Example illustrating edge length estimation based on the difference d_c of column sums s_c for a segment ($N = 4$) of a halfplane edge given by $y \leq 0.45x + 0.78$. The true halfplane edge is shown as a solid white line. The approximations of the edge segment using local steps of slope d_c and \tilde{d}_c , respectively, are shown as dashed lines with \times marking the ends of each step.

Given a halfplane H defined by $y \leq kx + m$, $k \in [0, 1]$, $m \geq 0$, we can use Eq. (5.6) to compute the length of the edge segment $y = kx + m$, $x \in [0, N]$. For the straight edge $y = kx + m$ of H , the slope k can be expressed as $k = \frac{y(x+\Delta x) - y(x)}{\Delta x} = y(x+1) - y(x)$. Observing integrated function values over a unit interval, $s_c = \int_{c-1/2}^{c+1/2} y(x) dx + \frac{1}{2}$, and denoting $d_c = s_{c+1} - s_c$, we conclude that $k = d_c$ for all $c \in \mathbb{R}$.

Assume $N \in \mathbb{Z}^+$. If we observe $c \in \{0, 1, \dots, N-1\}$, then each unit-wide interval used in the integration above defining s_c corresponds to one column of pixels in a digital grid. More precisely, given an image I of width N , being a coverage digitization of the halfplane H , $I = \mathcal{D}_c(H)$, the length l in Eq. (5.6) can be computed as

$$(5.7) \quad l(I) = \sum_{c=0}^{N-1} l_c, \quad \text{for } l_c = \sqrt{1 + d_c^2},$$

where the value $d_c = s_{c+1} - s_c$ is the difference of two consecutive column sums of the pixel values of the image. This is illustrated in Figure 9(a). The corresponding results for $k \notin [0, 1]$ follow by symmetries of the square grid.

Quantized case If the observed image I is, instead, an ℓ -level quantized pixel coverage digitization $I = \mathcal{D}_c^\ell(H)$, then the differences d_c are computed from quantized pixel coverage values. We denote such column differences, derived from a quantized coverage digitization, by \tilde{d}_c . These values are from the set \mathcal{Q}_ℓ and an edge with real valued slope $k \in [0, 1]$ is thereby approximated using local steps with

slopes from \mathcal{Q}_ℓ . An illustration is given in Figure 9(b). Due to the quantization, and the edge line only being approximated, an error is unavoidable. With an aim to minimize the maximal error, we introduce a scale factor γ_ℓ , providing the following formula for the estimation of the length of the edge present in the image:

$$(5.8) \quad \hat{l}(I) = \sum_{c=0}^{N-1} \hat{l}_c, \quad \text{where} \quad \hat{l}_c = \gamma_\ell \sqrt{1 + \tilde{d}_c^2}.$$

In the next section we derive a formula for the optimal value of the scale factor γ_ℓ as a function of the number of quantization levels ℓ so as to minimize the estimation error of (5.8) and we show that $\lim_{\ell \rightarrow \infty} \gamma_\ell = 1$; this is when formula (5.8) reduces to formula (5.7).

Binary case. Let us observe the special case of a 1-level quantized pixel digitization, that is, a binary image. The differences \tilde{d}_c of column sums for a binary image of a given edge with $k \in [0, 1]$ belong to the set $\mathcal{Q}_1 = \{0, 1\}$; the difference value $\tilde{d}_c = 0$ corresponds to a horizontal step (with slope $k = 0$) and the value $\tilde{d}_c = 1$ corresponds to a diagonal step (with slope $k = 1$). In this way, any edge with a real valued slope $k \in [0, 1]$ is approximated by a sequence of steps with slopes $k = 0$ or $k = 1$.

The estimation error for this situation is studied in e.g. [23, 39]. As already mentioned, to approximate the length l with $\tilde{l} = \sum_{c=0}^{N-1} \sqrt{1 + \tilde{d}_c^2}$ leads to an overestimate of the true edge length, in all cases when the slope k of the edge is not equal to 0 or 1. By scaling the step length with a properly chosen factor γ_1 , estimates with a minimal error are achieved. In [64] a value $\gamma_1 \approx 0.9604$ is shown to minimize the maximal error for the binary case.

Optimization to minimize maximal error. Let a set of linearly independent vectors $\mathbf{S} = \{\mathbf{S}_i = (1, \frac{i}{\ell}), i \in \{0, 1, \dots, \ell\}\}$ be given. Their slopes are $\frac{i}{\ell} \in \mathcal{Q}_\ell$ and they correspond to the possible slopes of local steps, \tilde{d}_c , as derived in the previous section. The length S_i of the vector \mathbf{S}_i is $S_i = \sqrt{1 + (i/\ell)^2}$.

The edge segment $y = kx + m$, $k \in [0, 1]$, on the interval $[0, N]$, represented by the vector $\mathbf{l} = (N, kN)$, can be expressed as a linear combination of two vectors \mathbf{S}_i and \mathbf{S}_j from the set \mathbf{S} , having slopes $\frac{i}{\ell}, \frac{j}{\ell} \in \mathcal{Q}_\ell$ such that $\frac{i}{\ell} \leq k \leq \frac{j}{\ell}$, as follows:

$$(5.9) \quad \mathbf{l} = \frac{(j - \ell k)N}{j - i} \mathbf{S}_i + \frac{(\ell k - i)N}{j - i} \mathbf{S}_j.$$

The length of \mathbf{l} can be estimated by using Eq. (5.9) as

$$(5.10) \quad \hat{l} = \gamma_\ell^{(i,j)} \left(\frac{(j - \ell k)N}{j - i} S_i + \frac{(\ell k - i)N}{j - i} S_j \right).$$

This corresponds to an edge segment such that $\tilde{d}_c \in \{\frac{i}{\ell}, \frac{j}{\ell}\}$ for all c , for which Eq. (5.8) is equivalent to Equation (5.10).

We derive $\gamma_\ell^{(i,j)}$ as a function of ℓ to minimize the maximal error of estimation formula (5.10). In the given context, the coefficients $\frac{(j - \ell k)N}{j - i}$ and $\frac{(\ell k - i)N}{j - i}$ of Eq. (5.10) represent the nonnegative number of repetitions of each of the local steps

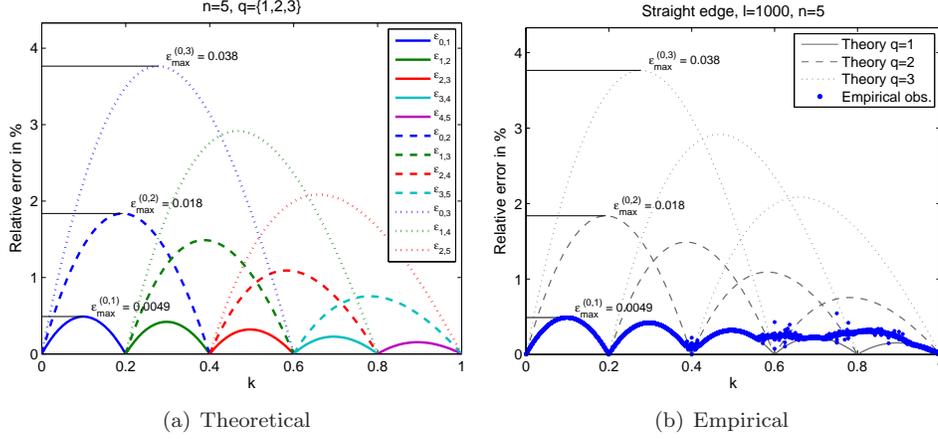


FIGURE 10. Relative error $\varepsilon_\ell^{(i,j)}(k)$, for $k \in [0, 1]$, $\gamma_\ell = 1$, $\ell = 5$. The values $\varepsilon_{\max}^{(i,j)}$ are indicated. (a) Theoretically derived behaviour of $\varepsilon_\ell^{(i,j)}(k)$, for $i \in \{0, 1, 2, 3, 4\}$, and $q = j - i \in \{1, 2, 3\}$. (b) Empirically observed values of $\varepsilon_\ell^{(i,j)}(k)$ for straight edges $y = kx + m$ of length $l = 1000$ for 10 000 values of k and random m , superimposed on the theoretical results shown in (a).

\mathbf{S}_i and \mathbf{S}_j in approximation of \mathbf{l} and are therefore required to be integers. This condition is, however, rather difficult to impose in the general case. We avoid the problem of integer valued coefficients by deriving the theory for segments of infinite length ($N \rightarrow \infty$) (see also [64]).

The *signed relative error* of the length estimate \hat{l} of an edge segment with slope k , such that $k \in [\frac{1}{\ell}, \frac{2}{\ell}]$, is given by the formula

$$(5.11) \quad \varepsilon_\ell^{(i,j)}(k) = \frac{\hat{l} - l}{l} = \gamma_\ell \frac{(j - \ell k)S_i + (\ell k - i)S_j}{(j - i)\sqrt{1 + k^2}} - 1.$$

To get a visual impression of the error function given by Eq. (5.11), we plot $\varepsilon_\ell^{(i,j)}(k)$ as a function of k in Figure 10(a), for the case $\gamma_\ell = 1$, $\ell = 5$ and for a number of combinations of i and j .

The best trade-off to minimize $|\varepsilon_\ell^{(i,j)}(k)|$ is found when

$$\max_{k \in [\frac{1}{\ell}, \frac{2}{\ell}]} \varepsilon_\ell^{(i,j)}(k) = - \min_{k \in [\frac{1}{\ell}, \frac{2}{\ell}]} \varepsilon_\ell^{(i,j)}(k)$$

which gives the following optimal value for $\gamma_\ell^{(i,j)}$

$$\gamma_\ell^{(i,j)} = \frac{2(j - i)\ell}{(j - i)\ell + \sqrt{(\sqrt{\ell^2 + i^2}\sqrt{\ell^2 + j^2} - (\ell^2 + ij))^2 + \ell^2(j - i)^2}},$$

where, for two given vectors S_i and S_j , the maximal error is $|\varepsilon_\ell^{(i,j)}| = 1 - \gamma_\ell^{(i,j)}$. The derivation of this result is given in [51].

If we assume that $j - i = q$ is constant, i.e., that vectors S_i and S_{i+q} are used for the approximation of a given line, it can be derived, [51], that the error $|\varepsilon_\ell^{(i,i+q)}|$ is the largest for $i = 0$. To minimize the maximal error for $i \in \{0, 1, \dots, \ell\}$, it is sufficient to observe $\gamma_\ell^{(0,q)}$:

$$(5.12) \quad \gamma_\ell^{(0,q)} = \frac{2q}{q + \sqrt{(\sqrt{\ell^2 + q^2} - \ell)^2 + q^2}} = \frac{2}{1 + \sqrt{(\sqrt{(\ell/q)^2 + 1} - \ell/q)^2 + 1}}.$$

The value $\gamma_\ell^{(0,q)}$ is denoted γ_ℓ and is used as optimal scale factor in the length estimation defined by Equation (5.10). The corresponding maximal estimation error, $|\varepsilon_\ell|$, is given by

$$|\varepsilon_\ell| = |\varepsilon_\ell^{(0,q)}| = 1 - \gamma_\ell^{(0,q)}.$$

Observing that $\sqrt{(\ell/q)^2 + 1} - \ell/q$ decreases as the ratio $\frac{\ell}{q}$ increases, we conclude that $\gamma_\ell^{(0,q)}$ increases, and consequently the length estimation error decreases, with $\frac{\ell}{q}$ increasing. In other words, by either increasing ℓ or decreasing q the maximal estimation error is reduced. This supports our main motivation for this work: using more coverage levels reduces the length estimation error.

It can be noticed that for $\ell = 1$, corresponding to a binary image, and with $q = 1$, being the only option for $q \leq \ell$, Eq. (5.12) provides that the scale factor γ_1 that minimizes the maximal length estimation error is

$$\gamma_1 = \gamma_1^{(0,1)} = \frac{2}{1 + \sqrt{(\sqrt{2} - 1)^2 + 1}} \approx 0.9604$$

which is a well known optimal result, [64]. The corresponding estimation error is $|\varepsilon_1| = 1 - \gamma_1 \leq 4\%$.

More generally, observing the estimation error corresponding to $\gamma_\ell = \gamma_\ell^{(0,q)}$ as a function of ℓ , as given by Eq. (5.12), we conclude that for any constant $q \ll \ell$

$$(5.13) \quad |\varepsilon_\ell| = \mathcal{O}(\ell^{-2}),$$

which gives an asymptotic upper bound for the estimation error as $\ell \rightarrow \infty$. The derivation of this result is given in [51]. Empirical studies performed and presented in [67] are in agreement with this theoretical results.

The value of $q = j - i$, appearing in the factor γ_ℓ , deserves some more attention. It reflects the difference in slope of the vectors S_i and S_j used in the linear combination, Eq. (5.9). Their slopes $\frac{i}{\ell}$ and $\frac{j}{\ell}$ correspond to the column differences \tilde{d}_c of the image $I = \mathcal{D}_c^\ell(H)$. A larger value of q leads to a larger error (according to Eq. (5.12)). For the purpose of minimization w.r.t. the maximal error, we observe the worst case situation, i.e., the value q should reflect the largest range of possible column differences \tilde{d}_c for any ℓ -level digitization of a halfplane edge for any given slope $k \in [0, 1]$

We have proved [51] that $q = 3$ corresponds to the worst case situation. According to this observation and Equation (5.12), the optimal scale factor is:

$$\gamma_\ell^{(0,3)} = \frac{6}{3 + \sqrt{(\sqrt{\ell^2 + 9} - \ell)^2 + 9}} .$$

This leads to the maximal estimation error for the general case: $|\varepsilon_\ell| = 1 - \gamma_\ell^{(0,3)}$. Combining this result with the asymptotic convergence as $\ell \rightarrow \infty$, Equation (5.13), the quantization level convergence of the length estimate follows straightforwardly:

Theorem 5.5. [52] *Let $I = \mathcal{D}_c^\ell(H)$ be an ℓ -level quantized coverage digitization of a halfplane $H : y \leq kx + m$, $k \in [0, 1]$. The length l of the straight edge segment $y = kx + m$ for $x \in [0, N]$, $N \in \mathbb{Z}^+$ can be estimated from I by using Formula (5.8) and it holds that $l = \hat{l} + 1/\ell^2$.*

However, for a coverage digitization with few coverage levels ℓ , the worst case situation with $q = 3$ does not appear. As already noticed, for $\ell = 1$ only two different slopes are available, and thus q cannot be greater than 1. We proved that q relates to ℓ as follows:

$$(5.14) \quad q \leq \begin{cases} 1, & \ell \leq 2, \\ 2, & 3 \leq \ell \leq 8, \\ 3, & \ell \geq 9 \end{cases}$$

We have also found examples where the upper bounds of q , in relation (5.14), are reached, i.e., for $\ell \geq 9$, $q = 3$ has to be considered for the (theoretical) maximal error bound estimation.

However, despite explicit examples that $q = 3$ is required for $\ell \geq 9$, empirical studies show that the situations where more than two different slopes appear in the estimation are very rare in practise. Observing the plot in Figure 10(b), we see that very few edges have an unscaled error greater than $\varepsilon_{\max}^{(0,1)}$, which supports the idea that computing γ_ℓ using $q = 1$ may be a better choice in practise, than using the theoretically derived worst case value $q = 3$. To use $\gamma_\ell^{(0,1)}$ rather than $\gamma_\ell^{(0,3)}$ is, therefore, recommended in general case, since it is observed as a better choice in our empirical test performed on other shapes, as well.

Local computation of length. Local computation of the edge length relies on local computation of the \tilde{d}_c values of Equations (5.8). How to estimate the slope of the observed edge from a small neighbourhood is a very important question which was answered in details in [51, 52]. The answers addressed issues related to the effects of quantization on the estimation algorithm, as well as what step are required to obtain an algorithm generally applicable, i.e., handling cases when $k \notin [0, 1]$.

Here, we only briefly summarize the most important observations and steps of the algorithm.

First observation is that for lines of a slope $k \in [0, 1]$, each value \tilde{d}_c depends on at most six pixels, located in a 3×2 rectangle; the remaining pixels in the column

pair do not contribute to the difference, neighbouring pixels in each row being the same. The difference \tilde{d}_c of two columns can thus be computed using information only from one, appropriately selected, 3×2 region—a subset of the two observed columns. We denote such a 3×2 configuration $D_{(c,r)}$, where the left pixel of the middle row is located at (c, r) , and with $\tilde{d}_{(c,r)}$ the difference of column sums within $D_{(c,r)}$. We then formulate a criterion to detect which $D_{(c,r)}$, for each observed pair of neighbouring columns, is intersected by the straight edge. Carefully treating the effects of quantization we prove that the following holds (see [51]):

For $\tilde{u}(c, r)$ defined as

$$\tilde{u}(c, r) = \frac{1}{2} \sum_{\tilde{p} \in D_{(c,r)}} \tilde{p} + (r - \frac{3}{2}),$$

then

$$\tilde{u}(c, r) \in [r - \frac{1}{2}, r + \frac{1}{2}] \Rightarrow \tilde{d}_c = \tilde{d}_{(c,r)}.$$

Based on this, we derive an estimation formula for the length of a line segment, with a slope $k \in [0, 1]$, from a local neighbourhood of a size 3×2 .

Further, we observe that if the slope k of the observed line is greater than one, then the differences calculated to determine $\tilde{d}_{(c,r)}$ should be taken row-wise instead of column-wise, and a 2×3 region should therefore be used instead of a 3×2 region. To simplify application of the method, we suggest to not use two different region sizes for the two situations ($|k| \leq 1$ and $|k| > 1$), but instead to use a 3×3 region in all the cases. Assigning a local edge length contribution to the central pixel of each 3×3 configuration also provides a more appealing output of the algorithm, where edge length values are associated with pixels instead of with edges in between pixels. We denote a 3×3 configuration, with the central pixel located at (c, r) , by $T_{(c,r)}$, with (quantized) pixel values \tilde{p}_i , $i = 1, 2, \dots, 9$, indexed row-wise, from left to right and top to bottom. A 3×3 configuration $T_{(c,r)}$ contains (if we consider lines of slope $k \in [0, 1]$) two 3×2 sub-configurations, $D_{(c-1,r)}$ and $D_{(c,r)}$.

Based on the above, we derive the following local edge length $\hat{l}_{(c,r)}^T$, assigned to one 3×3 configuration $T_{(c,r)}$:

$$\hat{l}_{(c,r)}^T = \hat{l}_l + \hat{l}_r, \text{ where } \begin{cases} \hat{l}_l = \begin{cases} \frac{\gamma_\ell}{2} \sqrt{1 + \tilde{d}_{(c-1,r)}^2}, & \tilde{u}_{c-1} \in [r - \frac{1}{2}, r + \frac{1}{2}] \\ 0, & \text{otherwise,} \end{cases} \\ \hat{l}_r = \begin{cases} \frac{\gamma_\ell}{2} \sqrt{1 + \tilde{d}_{(c,r)}^2}, & \tilde{u}_c \in (r - \frac{1}{2}, r + \frac{1}{2}] \\ 0, & \text{otherwise.} \end{cases} \end{cases}$$

The proposed edge length estimate over the whole image I is

$$(5.15) \quad \tilde{l}^T(I) = \sum_{T_{(c,r)} \subset I} \hat{l}_{(c,r)}^T.$$

Finally, it remains to observe a general situation—a 3×3 configuration $T_{(c,r)} \subset I$, where the image $I = \mathcal{D}_c^\ell(H)$ is an ℓ -level quantized coverage digitization of a halfplane $H : y \leq kx + m$ or $H : y \geq kx + m$, where the slope k is in $[-\infty, \infty]$. We present a method to isometrically transform every general configuration $T_{(c,r)}$ so that the transformed configuration $T'_{(c,r)}$ corresponds to that of a halfplane

$H' : y \leq k'x + m'$, where $k' \in [0, 1]$. In that way we extend application of the estimation formula (5.15) to the general case. We have defined criteria for selection of the appropriate transformation to be applied to an observed configuration $T_{(c,r)}$, and proved their correctness also for the quantized case.

Combining all of the above, the following algorithm is presented to compute the edge length contribution $\hat{l}_{(c,r)}^T$ for a given 3×3 configuration. To compute the complete edge length \hat{l}^T , the algorithm is applied to all pixels (or, alternatively, only to those adjacent to the object edges, if information about edges of the object is available), and the total length is obtained as a sum of the local edge length contributions, according to formula (5.15). For quantized coverage digitizations of straight edges, this algorithm provides the accuracy guaranteed by Theorem 5.5.

Algorithm 2.

Input: Pixel coverage values \tilde{p}_i , $i = 1, \dots, 9$, from a 3×3 neighbourhood $T_{(c,r)}$.

Output: Local edge length $\hat{l}_{(c,r)}^T$ for the given 3×3 configuration.

<pre> if $\tilde{p}_7 + \tilde{p}_8 + \tilde{p}_9 < \tilde{p}_1 + \tilde{p}_2 + \tilde{p}_3$ /* $y \geq kx + m$ */ swap(\tilde{p}_1, \tilde{p}_7) swap(\tilde{p}_2, \tilde{p}_8) swap(\tilde{p}_3, \tilde{p}_9) endif if $\tilde{p}_3 + \tilde{p}_6 + \tilde{p}_9 < \tilde{p}_1 + \tilde{p}_4 + \tilde{p}_7$ /* $k < 0$ */ swap(\tilde{p}_1, \tilde{p}_3) swap(\tilde{p}_4, \tilde{p}_6) swap(\tilde{p}_7, \tilde{p}_9) endif if $\tilde{p}_4 + \tilde{p}_7 + \tilde{p}_8 < \tilde{p}_2 + \tilde{p}_3 + \tilde{p}_6$ /* $k > 1$ */ swap(\tilde{p}_2, \tilde{p}_4) swap(\tilde{p}_3, \tilde{p}_7) swap(\tilde{p}_6, \tilde{p}_8) endif $\tilde{s}_1 = \tilde{p}_1 + \tilde{p}_4 + \tilde{p}_7$ $\tilde{s}_2 = \tilde{p}_2 + \tilde{p}_5 + \tilde{p}_8$ $\tilde{s}_3 = \tilde{p}_3 + \tilde{p}_6 + \tilde{p}_9$ </pre>	<pre> $\tilde{u}_l = (\tilde{s}_1 + \tilde{s}_2)/2$ $\tilde{u}_r = (\tilde{s}_2 + \tilde{s}_3)/2$ if $1 \leq \tilde{u}_l < 2$ $\tilde{d}_l = \tilde{s}_2 - \tilde{s}_1$ $\hat{l}_l = \frac{\gamma_\ell}{2} \sqrt{1 + \tilde{d}_l^2}$ else $\hat{l}_l = 0$ endif if $1 < \tilde{u}_r \leq 2$ $\tilde{d}_r = \tilde{s}_3 - \tilde{s}_2$ $\hat{l}_r = \frac{\gamma_\ell}{2} \sqrt{1 + \tilde{d}_r^2}$ else $\hat{l}_r = 0$ endif $\hat{l}_{(c,r)}^T = \hat{l}_l + \hat{l}_r$ </pre>
---	---

Estimator performance on synthetic test images. To study the accuracy and stability of the method applied to (both convex and nonconvex) curves, we evaluate the presented algorithm with respect to the accuracy of length estimation on a set of synthetic objects digitized using coverage digitization. We use the set of test shapes proposed in [21] (also used in e.g. [6]), containing convex and nonconvex objects with known perimeter, see Figure 11. The test shapes are digitized at a

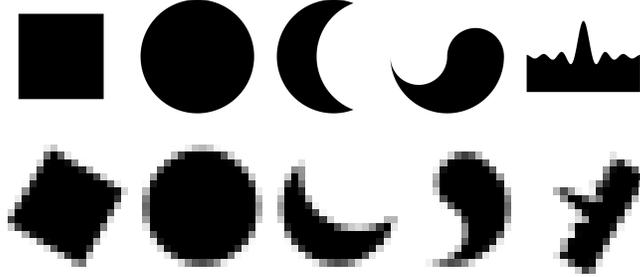


FIGURE 11. Pixel coverage digitizations of the test data set; one high resolution (grid resolution 512) (top) and one rotated low resolution (grid resolution 20) digitization (bottom).

range of resolutions, with random alignment in the digitization grid. Results of this evaluation are presented in Figure 12. We use $q = 1$ when computing the scale factor $\gamma_\ell = \gamma_\ell^{(0,q)}$, since that shows to provide empirically better results than using $q = 2$ or $q = 3$. For each test shape and for a number of resolutions the coverage digitizations for several different rotations and positions of the shape in the digital grid are computed. A number of quantization levels are observed and the nonquantized digitization, indicated with $\ell = \infty$, is also included. The average performance of the method is plotted as a function of resolution. The true pixel coverage digitizations of the test shapes are approximated by 256-sampled coverage digitizations, $\hat{D}_c^{256}(S)$.

As noticed in [21], the thin elongated peak of the yin-yang curve slows down the convergence to the true value quite significantly. For complementary comparison we also show Figure 13, where the yin-yang shape is removed from the test material.

As an additional test (also performed in [6]), we estimate the perimeter of a rotated square and plot, in Figure 14, the estimate as a function of angle. As can be seen the rotational variation decreases rapidly with the increase of number of coverage levels, and is for $\ell \geq 3$ within $\pm 1\%$ of the mean estimate. The slight overall underestimate of the perimeter of the square (less than 0.5%) is attributed to its four corners.

In general, we observe that the presented *local* estimator for boundary length estimation performs very well in comparison with the nonlocal multigrid convergent estimators evaluated in [6]. Note that this holds also for estimates based on relatively few coverage levels.

6. Defuzzification and high resolution reconstruction

As described in previous sections, there are several reasons to consider discrete fuzzy representations as a useful way to for represent objects in images. Among first mentioned were advantages of fuzzy representations in handling noise and intensity variations in images, as well as imprecision of various types. No matter what physical property is imaged (reflection of light, density of a material, intensity

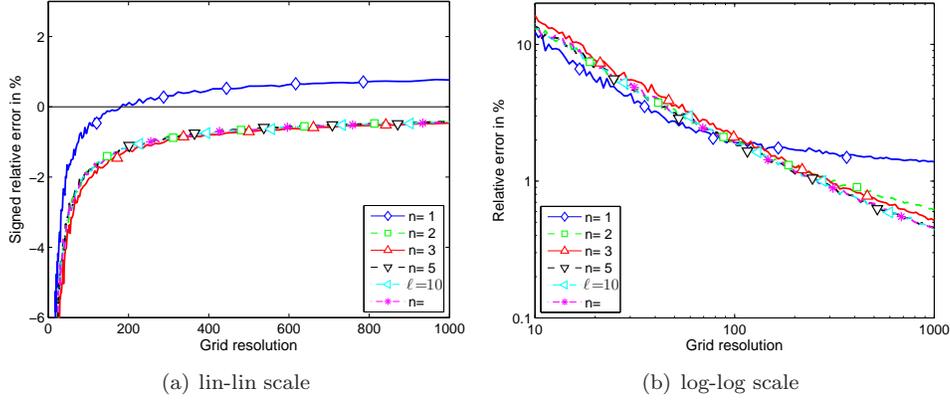


FIGURE 12. Relative errors (in percent) of perimeter estimates for the test shapes (shown in Figure 11) digitized at increasing resolution for 5 different degrees of quantization and for nonquantized ($\ell = \infty$) coverage digitization.

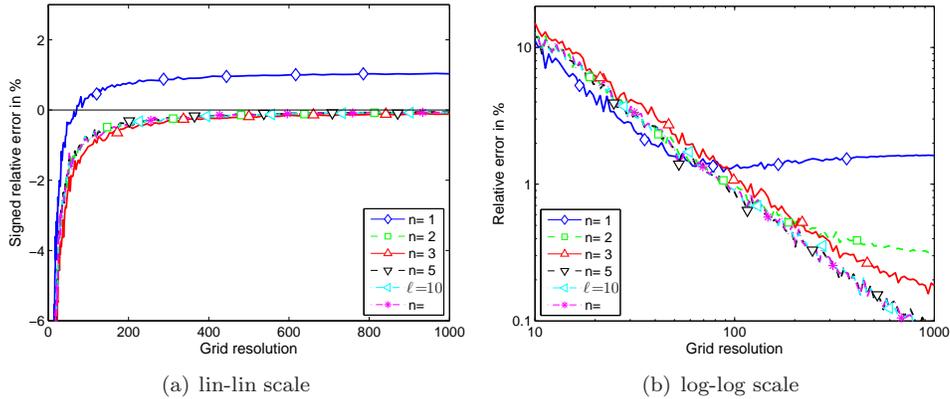


FIGURE 13. Relative errors (in percent) of perimeter estimates when the yin-yang shape is removed from the set of test shapes.

of a flow, or amount of movement), and independently of if a crisp real object or a naturally fuzzy one (such as a cloud, or a flame, for example, but also properties like blood flow, or activity of cells) are represented in an image, an appropriately chosen membership function can always be chosen so that a corresponding fuzzy representation provides better preservation of information relevant for the imaged object and better treatment of appearing imprecision, than the crisp one.

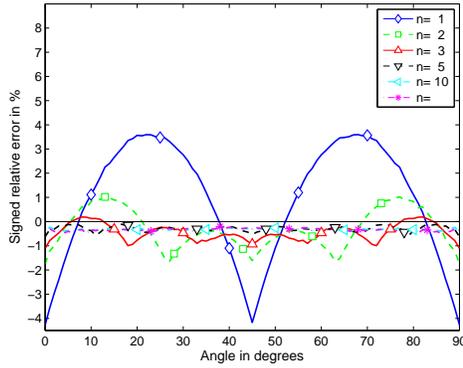


FIGURE 14. Relative errors (in percent) of perimeter estimates for a rotating square of size 128×128 , from images with 5 different degrees of quantization and for the nonquantized ($\ell = \infty$) case.

Our focus has been on preservation of geometric features of the crisp continuous imaged objects; for that purpose we find the coverage representation most appropriate. Features such as perimeter, area, and other geometric moments are good examples for that.

In spite of mentioned advantages of utilizing fuzzy segmented images, a crisp representation of objects may still be needed. Reasons for that are, e.g., to facilitate easier visualization and interpretation. Even though it contains less information, a crisp representation is often easier to interpret, understand, and manipulate, especially if the spatial dimension of the image is higher than two. Moreover, analogues for many tools available for the analysis of binary images are still not developed for fuzzy images. This may force us to perform at least some steps in the analysis process by using a crisp representation of the objects, requiring the ability to “switch to” an appropriate crisp representation at any point in the process.

In our work, presented in [28, 30, 54], we explored possibilities to generate crisp representations of image objects, starting from fuzzy ones. The process of replacing a fuzzy set with an appropriately chosen crisp set is, in fuzzy set theory, referred to as defuzzification. It can be performed either as an inverse of fuzzification, [40], with the intention to recover a fuzzified crisp original, or as a process independent of any fuzzification, but based on some pre-defined conditions that should be fulfilled for a crisp set to be the representation of a given fuzzy set [24, 44]. In image analysis the fuzzification function is rarely known, and practically never analytically defined; as mentioned earlier, fuzzification of an image is a consequence of a combination of properties of the continuous original, discretization effects, and imaging conditions. Therefore, the inverse of a fuzzification function cannot, in general, be used to define defuzzification. Defuzzification is rather performed so that certain predefined criteria are respected in the process. It seems both natural and beneficial to, for this purpose, impose criteria that reflect properties of a (possible) continuous crisp original. By that, the two approaches to defuzzification are combined. We refer

to such defuzzification, which is defined based on some set of imposed criteria, but with “awareness” of the crisp original, as *object reconstruction*.

It can be observed that defuzzification, following fuzzy segmentation, is an alternative to crisp segmentation. We found reasonable to expect that such an approach to crisp segmentation can be tuned so that it enables preservation of the most relevant information (features) of the observed object (and available in the given fuzzy representation) for the application in question. In other words, loss of information is inevitable when defuzzification is performed, but preservation of some features can be prioritized, if appropriate for the application.

The reconstruction that we propose is based on preservation of geometric properties of an object; they are estimated with high precision from a fuzzy object representation and imposed as defuzzification criteria. We determine a crisp representative of the given fuzzy set to be a crisp set which has the selected features as similar as possible to the corresponding features of a given fuzzy set. In this way, defuzzification is defined as an optimization process, where the distance between the given fuzzy set and its crisp reconstruction (defuzzification) should be as small as possible. Formal definition, together with further details related to choice of features to consider, choice of distance to minimize, and choice of optimization method to perform, are given below, in accordance with [28, 30, 54].

Optimal defuzzification. Let $\mathcal{F}(X)$ be the set of fuzzy subsets of a reference set X , and $\mathcal{P}(X)$ be the set of crisp subsets of X , also known as the *power set* of X .

Definition 6.1. Given a fuzzy set $A \in \mathcal{F}(X)$, an *optimal defuzzification* $\mathcal{R}(A)$ of A , with respect to the distance measure d , is

$$(6.1) \quad \mathcal{R}(A) = \arg \min_{B \in \mathcal{P}(X)} [d(A, B)] .$$

Distance Measure. For any injective mapping Φ from $\mathcal{F}(X)$ into a metric space H , we can define a metric on $\mathcal{F}(X)$ by requiring that Φ is an isometry. Assuming a mapping $\Phi : \mathcal{F}(X) \rightarrow H \subset \mathbb{R}^n$, where the vector $\Phi(A)$ contains different features of a fuzzy set A , we define a *feature distance* between fuzzy sets.

Definition 6.2. The *feature distance* $d_p^\Phi(A, B)$ between fuzzy spatial sets A and B , on the same reference set X , is the Minkowski distance d_p between the representations $\Phi(A)$ and $\Phi(B)$ of the sets A and B in the feature space $H \subset \mathbb{R}^n$:

$$(6.2) \quad d_p^\Phi(A, B) = d_p(\Phi(A), \Phi(B)).$$

For $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, the Minkowski distance is defined as $d_p(\mathbf{x}, \mathbf{y}) = \sqrt[p]{\sum_{i=1}^n |x_i - y_i|^p}$.

By suitably designing the mapping Φ , i.e., by considering suitable relevant features of the observed sets, the above distance measure can be tuned to provide defuzzifications where both shape characteristics and membership values are taken into account. This enables defuzzification that fits the individual problem well, and provides a powerful family of defuzzification methods.

Features. Preservation of geometric features of shapes in fuzzy representations gives good motivation to include them into the representation Φ of a fuzzy set. In the 2D case, area, perimeter, and geometric moments of a continuous shape are

shown to be preserved with a high precision in a fuzzy representation. They are used in reconstruction as global features, providing information about global geometric properties of the object. Memberships of all the points should be considered as well; they are referred to as local features of the object. Gradient in each point is another relevant feature considered. In addition to the local and global features, a range of meso-scale features can be considered. Our work, presented in [30], introduces meso-scale area measurements into the defuzzification procedure.

Different combinations of features were included in representations of the observed sets and their influence on reconstruction is evaluated in [54]. Even though the selection of features is highly dependent on the requirements imposed by the task, and often on the type of objects, it is clear that inclusion of features of different scales (local, meso, and global) improves the reconstruction.

Tests performed in [54] gave insight in behaviour of Minkowski distances depending on values of p . In most cases, choices $p = 1$ and $p = 2$ showed to be best suited for our needs, and requirements of the observed tasks.

Optimization. In general, the optimization problem (6.1) cannot be solved analytically. In addition, the search space $\mathcal{P}(X)$ is too big to be exhaustively traversed. As a consequence, we are forced to rely on some numerical optimization method, to minimize the distance between the fuzzy set and its crisp counterpart. In [54], two methods, floating search and simulated annealing, are used to find an approximate solution for Eq. (6.1). Simulated annealing performs very well for the task. It is, however, nondeterministic, while at the same time the trade-off between computation efficiency and performance may lead to long computation time required. The optimization task is a well separated problem, so many other search methods can be used to approximately solve Eq. (6.1). We applied DC (Difference of Convex functions) in [32], and SPG (Spectral Projected Gradient) based optimization in [27]. These methods are deterministic and fast, however less flexible with respect to inclusion of features into the feature vector representations Φ .

Reconstruction by optimal α -cut. In general, we do not impose any topology related constraint to defuzzification, even though the proposed defuzzification algorithm allows inclusion of such constraints and control of, e.g., the number of connected components of the resulting reconstruction. Further, we do not, in general, impose criterion of preservation of monotonicity of membership values, i.e., we do not require that an obtained reconstruction must be an α -cut of the starting fuzzy set. The optimization method used in the process is allowed to “decide” about the most appropriate selection of points included in the crisp representation; it may therefore happen that for two points of a fuzzy set one with lower membership is included in defuzzification, while the one with higher membership is not, if that leads to overall better optimization result.

If appropriate, monotonicity preservation can be imposed. Such an approach leads to defuzzification by α -cutting; the defuzzified set is found by thresholding the fuzzy membership function at an appropriate level. This is an appealingly simple method, however, the selection of a threshold is to be determined in some way (in most cases depending of an application) and this is often a rather difficult task. We notice that the simplicity of the method somewhat restricts its performance

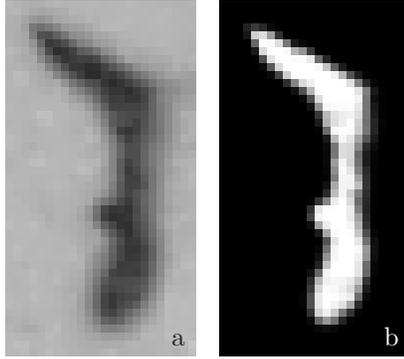


FIGURE 15. Example of fuzzy segmented 2D image. (a) A part of a microscopy image of a bone implant. (b) Fuzzy segmentation of a bone region in (a).

and applicability; this is a commonly known disadvantage of thresholding as a segmentation method in image processing. However, defuzzification by α -cutting fits well, as one specific case, within the proposed framework of defuzzification by feature distance minimization, and we have tested and compared its performance with other, less restricted, defuzzifications.

With the constraint to preserve monotonicity added, the task of minimization of differences between relevant selected (geometric) features of the fuzzy set and its defuzzification is restricted to the selection of *the optimal α -cut*, i.e., the α -cut at the smallest distance from the fuzzy set in terms of feature distance. The search space in this type of optimization is much smaller (there are as many α -cuts to explore as there are different membership levels in the fuzzy set), and exhaustive search can easily be performed to select the optimal α -cut.

Figure 16 presents examples of different defuzzifications; threshold at $\alpha = 0.5$, optimal α -cut, and defuzzification without monotonicity constraint are observed. Achieved minimal distances, as well as reconstructed sets are presented, for a part of a histological image shown in Figure 15.

Reconstruction at increased spatial resolution. A fuzzy representation, in general, contains a lot more information than a crisp representation at the same spatial resolution. If defuzzification is performed at a given spatial resolution, i.e., the crisp representative is generated at the same spatial resolution as the given fuzzy set, this additional information is, to a high extent, lost. It is reasonable to pose the question whether, instead, this information can be utilized to provide a crisp reconstruction at an increased spatial resolution. We explored this issue in [28, 30].

If r_F is the spatial resolution of the given fuzzy set, and r_C is the spatial resolution of the crisp (defuzzified) set, then increase of a spatial resolution is expressed by a factor $r = \frac{r_C}{r_F}$. This factor is required to have an integer value. In that case,

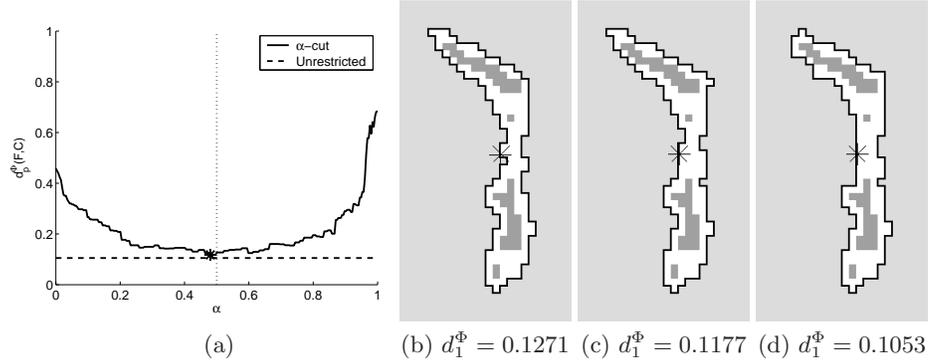


FIGURE 16. Different defuzzification approaches, and their respective feature distances. (a) Plots of feature distance p_1^Φ as a function of α for defuzzification by α -cutting. The minimum is indicated with a star (*), while the vertical line is positioned at $\alpha = 0.5$. (b) Defuzzification by α -cutting at $\alpha = 0.5$. (c) Defuzzification by α -cutting at optimal α . (d) Defuzzification by simulated annealing, starting from the optimal α -cut.

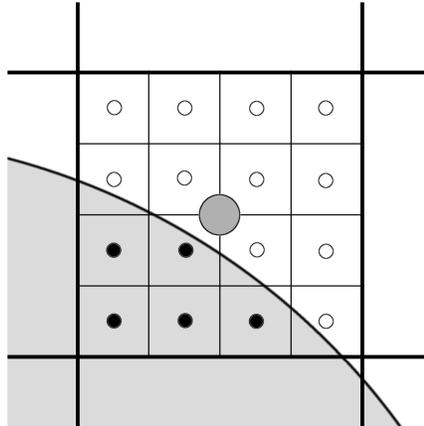


FIGURE 17. One pixel in a low resolution (fuzzy) image, and the corresponding block of 4×4 pixels of a 4 times higher resolution (crisp) reconstruction.

each spel in the low resolution representation corresponds to a block of $r \times r$ spels in the high resolution representation; a 2D illustration is shown in Figure 17.

As in Section 5.2, we recall that there are two approaches to perform multigrid studies: one is to observe the (r times) dilated object in the unchanged grid, whereas

TABLE 1. The contribution of the different features to the feature distance, and the total distance, without (Dist 1), and with (Dist 2), the meso-scale area features.

Figure	Perimeter	Area	Centroid	Membership	Meso-scale	Dist 1	Dist 2
18(b)	0.0000	0.0000	0.0000	0.0957	0.3828	0.0957	0.4785
18(c)	0.0015	0.0381	0.0000	0.0957	0.1758	0.1353	0.3111

the other is to observe the unchanged object inscribed in the (r times) refined grid. These two approaches are dual. We use the first one, which implies that the size of the spel is equal to 1 in all the observed grids, whereas the object features calculated in different grids are resolution-variant.

A main idea is to interpret the membership value of a (n -dimensional) spel as an additive property which is distributed over the whole (hyper)-volume of the spel. As such, membership can be summed over blocks of spels, or divided into parts if a spel is divided into sub-spels. This corresponds to the area/volume of a fuzzy set, which by definition is the sum of membership values of all elements of the set [41]. Therefore, instead of comparing pairs of corresponding spels in the fuzzy and the crisp image, we relate the membership value of a spel in the fuzzy image to the sum of membership values of the corresponding block of r^n spels (i.e., for the crisp set, the number of object covered (sub-)spels) in the high resolution representation. This approach can be further generalized so that blocks of spels can be observed in both fuzzy and crisp representations. In that way an additional range of features can be incorporated in defuzzification; local sums of membership values (i.e. area/volume/Lebesgue measure) computed for blocks of spels and interpreted as meso-scale features, ranging from the local (one spel size) to the global (whole object size). This scale space approach provides an appropriate treatment of details in images, where the details are usually relevant only in some range of scales. We addressed it in [30].

An illustrative example is given in Figure 18, where defuzzifications without, and with, meso-scale features are shown, for a synthetic example. The result of defuzzification of the object in Figure 18(a), using the proposed scale space approach, is shown in Figure 18(c). Even though the global features are perfectly matched in the solution presented in Figure 18(b) (obtained without meso-scale features), we consider the solution in Figure 18(c) to better preserve the properties of the original set. A problem which we refer to as “transportation of area” over the image appears if no meso-scale features are used (Figure 18(b)), and is avoided if such features are included in feature vector and considered in defuzzification (Figure 18(c)). The contributions of the different features to the overall distance are, for this example, given in Table 1. From the presented data, it is clear that selection of features considered in defuzzification has very high influence on the result. If understanding of the scene, or a priori knowledge about the object are available, this can be used to obtain the most appropriate reconstruction of an observed object.

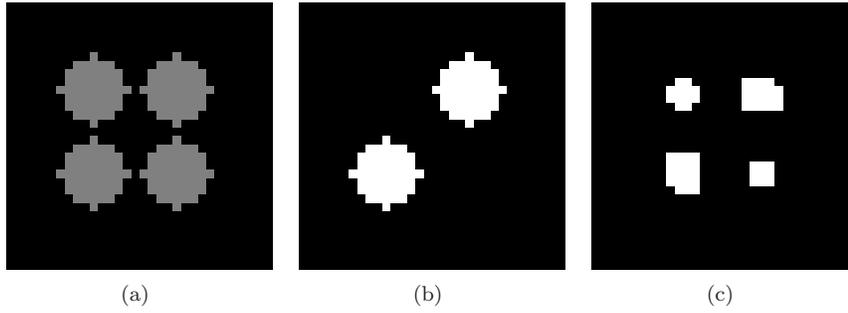


FIGURE 18. (a) Four discrete disks of radius 4 and membership 0.5. (b) Optimal defuzzification using feature distance without meso-scale area components. (c) Defuzzification using feature distance including meso-scale area components.

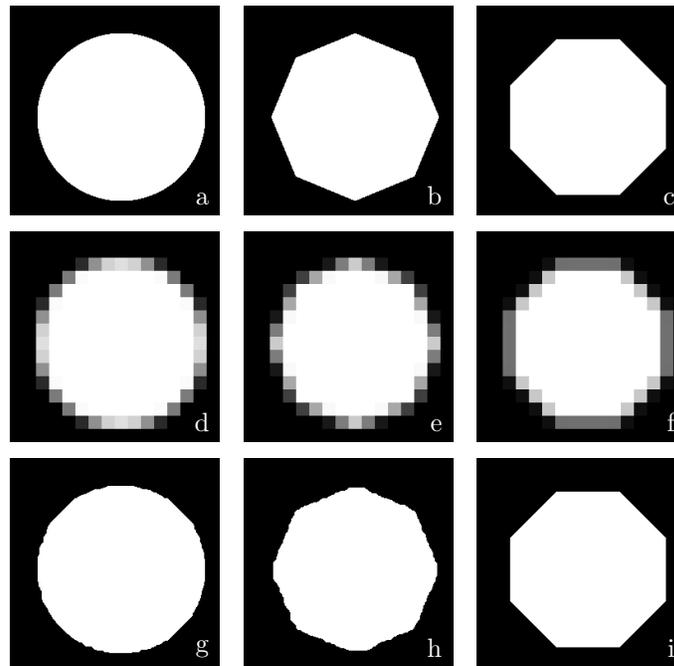


FIGURE 19. Top row: High resolution crisp representations of three crisp continuous shapes. Middle row: Fuzzy representations of the same crisp continuous shapes at relatively low resolution. Bottom row: Defuzzifications of the sets given in the middle row, at 16 times higher resolution.

High resolution reconstruction of a shape by the suggested method is illustrated in Figure 19. Defuzzifications of the fuzzy sets in the middle row of Figure 19, at 16 times higher resolution, are shown in the bottom row of Figure 19. The defuzzification performed is based on minimization of differences between area, perimeter, centroid and membership values of the fuzzy and crisp representations. The amount of information in the images shown in the top row, considered to constitute “ground truth” for this example, as the best possible crisp representations of the observed objects, is $256 \times 256 \times 1$ bit while the amount of information in the fuzzy images in the middle row, and consequently, in the high resolution reconstructions in the bottom row, is $16 \times 16 \times 8$ bit. This means that, without increasing the amount of information by a factor of 32 beyond initially available, an exact reconstruction (shown in the top row) is not possible. Some artefacts are therefore visible in the images in the bottom. However, considering the amount of available information, the visual appearance of the reconstruction result is, in our opinion, rather appealing and facilitates judgement on the original continuous crisp shape.

In order to simplify the notation, we describe the necessary steps in the process of scale-space high resolution reconstruction in the 2D case. Generalization to nD is straightforward, up to availability of appropriate feature estimates.

Weighting of Features. In order to provide that the effect of the total contribution of all measures of one (type of) feature, observed at one particular scale, is of approximately the same size as the effect of one global feature, features of multiplicity h are scaled with $\frac{1}{\sqrt[p]{h}}$, where p is the exponent of the Minkowski distance in Eq. (6.2).

To compare features calculated at different scales, measures also have to be rescaled with respect to the spatial resolution of the image and the dimensionality of the particular feature; e.g., perimeter of an object increases linearly with the spatial resolution, whereas area increases quadratically. To get resolution invariant global features we divide each feature with the feature value of the observed reference set X , that is, for an arbitrary feature F , we observed the resolution invariant feature $\tilde{F}(S) = \frac{F(S)}{F(X)}$.

Feature Vector Representation. For a given fuzzy set $S \in \mathcal{F}(X)$ of size $2^m \times 2^m$ pixels, we generate $(m + 1)$ partitions of the set into square blocks of $2^{m-i} \times 2^{m-i}$ pixels, for $i = 0, \dots, m$. Each partition i consists of 2^{2i} blocks. Let B_j^i represent the j th block of $2^{k-i} \times 2^{k-i}$ pixels, where $j = 1, \dots, 2^{2i}$, $i = 0, \dots, m$. A feature of highest interest is area of a set, at all levels (block sizes). Block B_1^0 is equal to the set S and, correspondingly, $\tilde{F}(S) = \tilde{F}(B_1^0)$, for all the observed features F . The membership values of all the pixels are included in such a representation, being local areas of one-pixel-size blocks ($i = m$). In addition, the perimeter of the set S , as well as the coordinates of its centroid, are included in the feature representation.

This leads to the following form of the feature representation $\Phi_m(S)$ of S :

$$\Phi_m(S) = \left(\frac{1}{\sqrt[p]{2^{2m}}} \tilde{A}(B_1^m), \dots, \frac{1}{\sqrt[p]{2^{2m}}} \tilde{A}(B_{2^{2m}}^m), \right. \\ \left. \frac{1}{\sqrt[p]{2^{2(m-1)}}} \tilde{A}(B_1^{m-1}), \dots, \frac{1}{\sqrt[p]{2^{2(m-1)}}} \tilde{A}(B_{2^{2(m-1)}}^{m-1}), \right. \\ \left. \dots \frac{1}{\sqrt[p]{2^0}} \tilde{A}(B_1^0), \tilde{P}(S), \tilde{C}_x(S), \tilde{C}_y(S) \right).$$

A convenient way to efficiently implement and utilize scale dependent features in defuzzification is to use a resolution pyramid. We use two resolution pyramids for storing the areas of the blocks B_j^i of the fuzzy original set, and of the crisp defuzzification. Pyramids are built by grouping 2×2 neighbouring (*children*) pixels in the image at the current resolution level, and create one (*parent*) pixel at the next, lower, resolution level, where the value of the parent pixel is assigned to be the sum of the values of the children pixels. The process is repeated at every newly created resolution level, until the lowest possible resolution.

Defuzzification. For a given fuzzy set S , containing $2^m \times 2^m$ pixels, a resolution pyramid representation with $m + 1$ resolution levels is built. For reconstruction at $r = 2^k$ times increased resolution a resolution pyramid for the crisp set K , with $m + k + 1$ resolution levels, is created and defuzzification is performed by minimizing the feature distance $d^\Phi(S, K) = d(\Phi_m(S), \Phi_m(K))$, where d is the Minkowski distance for appropriate choice of p .

Depending on the optimization method selected, a starting configuration may be an important issue; performance of simulated annealing, e.g., highly depends on the selection of the initial configuration. We suggested to use the optimal α -cut of S , i.e., the α -cut at minimal distance d^Φ to S , as the starting configuration for defuzzification. In order to obtain the initial configuration K at r times increased resolution, each pixel in the α -cut is subdivided into 2^r sub-pixels.

Scale space defuzzification of 3D fuzzy sets. The defuzzification method, suggested for 2D discrete spatial fuzzy sets is straightforwardly generalized to the 3D case. The features selected to be included in the feature distance are local, meso-scale, and global volumes, obtained by iterative grouping of blocks of $2 \times 2 \times 2$ voxels, and surface area and centroid, as additional global features.

Once when the feature representation is generated, the defuzzification process is exactly the same as in the 2D case. However, due to a rapid increase of data, compared to 2D images, some practical implementation related issues may become relevant. An important one is certainly the choice of optimization strategy; in our work presented in [27, 32] we have addressed utilization of optimization methods (DC based, SPG based) applicable to large scale optimization problems to the task of defuzzification. Examples of application of a 3D high resolution reconstruction on real medical images are given in Section 7.4.

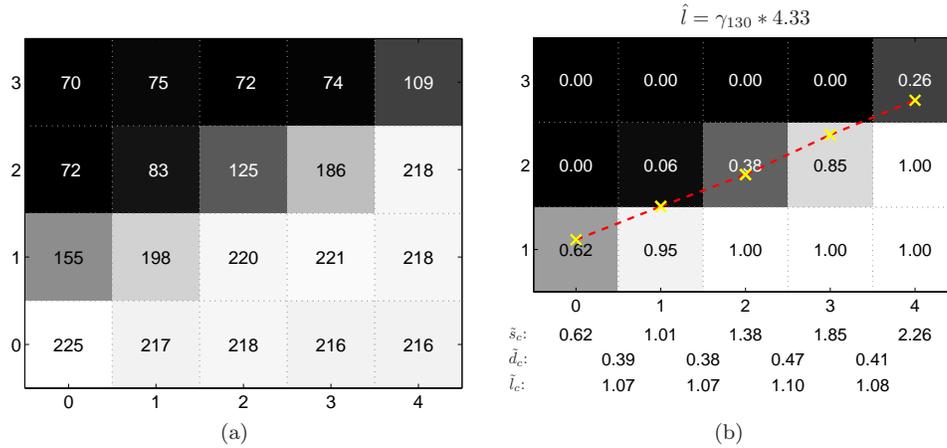


FIGURE 20. (a) Close up of the straight edge of a white paper imaged with a digital camera. (b) Segmentation output from Algorithm 1, utilizing 130 positive grey-levels. Approximating edge segments are superimposed (dashed lines) on the image (compare with Figure 9).

7. Performance of the presented approach

The coverage model, in our opinion, offers an appealing way to improve information preservation in image processing. With fast advancement in imaging techniques, which naturally comes with development of technology, it becomes less and less acceptable to lower the quality of image processing results and subsequent conclusions by suboptimal methods, not suited to handle the available information and achieved precision in imaging. The coverage model is one of the potential answers to the challenge of “keeping up with the technological development”. However, a lot of work still remains, in order to have a complete tool-box for image processing with the coverage model. We have presented in this paper our results obtained along the way; they include segmentation methods, methods for feature extraction, and methods for object reconstruction. Some properties and theoretical evaluation of the proposed approaches are given, separately for each of the methods. Our main interest is in having a processing “chain” consisting of methods developed for the coverage model. We therefore evaluated some of the combinations of the described methods, applied on different real tasks. In some cases, we have used real images obtained in controlled environment, where ground truth is known and the performance evaluation of the method is direct, whereas in some cases we showed applicability of the method in real conditions, where no ground truth is available and conclusions related to performance are derived more implicitly.

7.1. Comparison of different perimeter estimation methods. We created a controlled real environment for testing performance of the perimeter estimator presented in Section 5.3. This test is also presented in [52]. We took a number of photos of a straight edge of a white paper on black background at a number of angles, using a digital camera in grey-scale mode. All images were cropped to the same width and the slope of each edge was computed using moments [48]. The edge lengths were computed according to Eq. (5.6). We consider these estimates to be correct edge lengths and we use them as the ground truth.

We have evaluated the perimeter estimator presented in Section 5.3 on the task of edge length estimation for the described images, by comparing its performance with performances of several methods previously presented in the literature. In order to confirm that utilization of grey-levels indeed improves the performance, we compared the proposed estimator with most known and best performing estimators for bi-level (crisp) images; we considered the method in [64] (which is equivalent with the proposed one if binary case is observed), and the corner count method, [66]. In order to confirm superiority of the proposed method over previously existing ones applicable to grey-level images, we considered the method of Eberly and Lancaster [11]. Observing high noise-sensitivity of the original approach described in [11], where the gradient components are computed only based on the difference of two pixels, we complemented this method with an additional smoothing step. Two Gaussian filters of different sizes are used to smooth the images prior to length estimation. This improved the performance of the method [11] when applied to straight edges estimation. However, Gaussian filtering is expected to reduce performance on nonstraight edges.

Figure 20(a) shows a part of one image taken as described above, where an edge of a half-plane is presented. The correct value of the slope, $k = 0.42$, is computed. The original image presented in Figure 20(a) is segmented by the coverage segmentation method described in Section 4.1. Figure 20(b) shows the output of the coverage segmentation. Double thresholding is performed, and the number of grey-levels preserved in a one pixel thick border of the object is found to be between 90 and 140 (out of 255) for the different photos. Values of estimated slopes, \bar{d}_c , and lengths, \bar{l}_c , for local steps, are indicated, as well.

Binary segmentation, required for testing binary estimators, is performed by using Otsu's thresholding method [38], which works very well on the high contrast scene.

Evaluation results are presented in Figure 21. Relative estimation errors for six considered methods are shown, for digital straight segments with different slopes. The maximal errors for the observed methods are presented in Table 2. These results confirm superior performance of the proposed coverage model, and in particular, the proposed perimeter estimation method.

7.2. Coverage segmentation followed by feature estimation for noisy data.

The segmentation method presented in Section 4.2 provides exact coverage values if continuous pixel coverage values and a noise-free environment are ensured. These ideal conditions, however, never exist in practise; quantization errors and presence

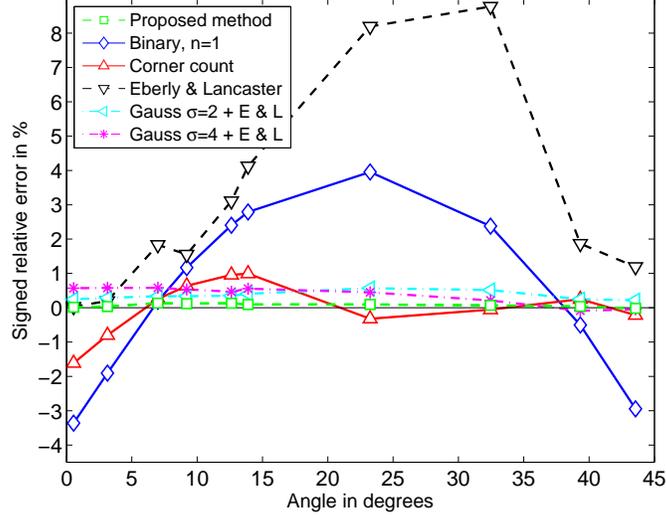


FIGURE 21. Relative errors for different methods when used to estimate the length of the edge of a white paper photographed at different angles with a digital camera.

TABLE 2. Maximal perimeter estimation error when using different approaches.

Method	Max error
Proposed coverage based method [52]	0.14%
Binary method [64]	3.95%
Corner count [66]	1.61%
Eberly & Lancaster, [11]	8.78%
Smoothing ($\sigma = 2$) & Eberly & Lancaster	0.57%
Smoothing ($\sigma = 4$) & Eberly & Lancaster	0.58%

of noise are unavoidable in real images. Performance tests of the method in such environment are certainly of interest.

Being interested in coverage segmentation primarily for its further use for precise and accurate feature estimation, we have, in addition to directly testing the performance of the proposed segmentation method, also evaluated feature extraction based on such segmentations. We have observed perimeter estimates computed by the method presented in [52] (see Section 5.3), and area estimates computed according to [50] (Section 5.2).

We have performed two types of tests, also presented in [53]. First, we have observed synthetic objects with known feature values, affected by simulated noise. We have evaluated both coverage values assignment in the segmentation method

and subsequent feature extraction. Second test is performed on real histological colour images.

7.2.1. Synthetic noisy images. A synthetic object (Fig. 22(a)) of known dimensions is randomly placed (rotated and translated) at a number of different positions in the square grid and digitized using coverage digitization. A zoomed-in part of the resulting object, with a (one-pixel thick) partial coverage at its boundary, as well as its superimposed crisp discretization, are shown in Fig. 22(b). Each digitization is subsequently corrupted by increasing levels of additive uncorrelated Gaussian noise, which provides our observed set of test images.

Crisp digital representations of continuous objects are created by a Gauss centre point digitization. This digitization is considered to be equivalent to an ideal, error free, crisp segmentation, and is used as a starting crisp segmentation, required for our proposed coverage segmentation method described in Section 4.2. In that way, evaluation of the method is not dependent of the properties of any particular segmentation method, and is therefore more objective. The same (Gauss centre point) crisp digitization is also used as a reference in comparisons. The neighbourhood required for estimating pure class values is defined by using an appropriate 2D Gaussian mask.

To evaluate the pixel coverage segmentation, the assigned coverage values are, per pixel, compared with the true ones, for increasing amounts of added noise. The average absolute error for coverage values of boundary pixels,

$$\varepsilon = \frac{1}{N} \sum_{p \in B} |\hat{\alpha}(p) - \alpha(p)|$$

where B is the set of evaluated boundary pixels, N is the cardinality of B , and $\hat{\alpha}(p)$ and $\alpha(p)$ are, respectively, assigned and true coverage for a pixel $p \in B$, is computed and presented in Fig. 22(d). A number of random displacements of the object are observed for each level of noise. In Fig. 22(e) we show the relative error of perimeter estimation for the observed synthetic object for increasing levels of noise, whereas Fig. 22(f) presents the area estimation on the same test object, under the same conditions.

As it is visible from the plots in Fig. 22(d-f), the improvement when using the suggested method, compared to the results obtained for an ideal (noise free) crisp segmentation, is significant when the standard deviation of the present noise does not exceed 20%. Above that level, the suggested method does not provide any improvement in terms of accuracy. It is, however, worth noting that the precision of the feature estimates (exhibited as low variation of the obtained results) is significantly higher for the proposed method, and in the case of area estimation provides improvement of the result for all the observed noise levels (i.e., for up to 40% of noise). These observations confirm applicability and excellent performance of the proposed methods—coverage segmentation and feature extraction—for analysis of noisy images.

7.2.2. Quantitative analysis of a histological image. After performing tests on synthetic images, we have tested applicability of the coverage segmentation

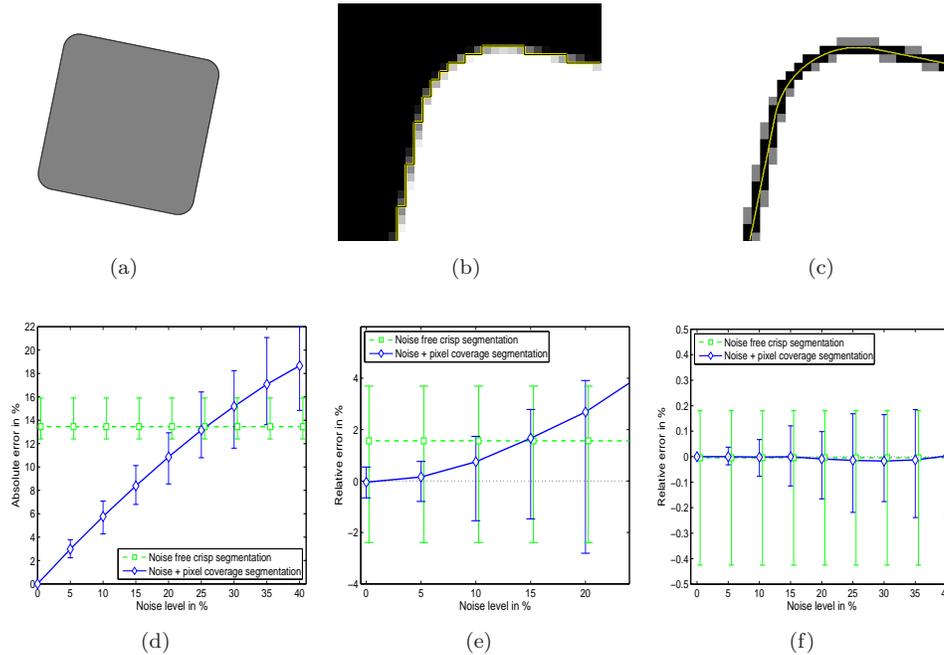


FIGURE 22. (a) Continuous synthetic test object. (b) Part of a pixel coverage segmentation of (a), with superimposed border of its crisp segmentation. (c) Part of the set B . Grey pixels are removed in the thinning step. Superimposed is the border of the original continuous object. (d) Average absolute error of pixel coverage values in B for increasing levels of noise, lines represent means of 100 observations and bars show max and min errors. (e) and (f) Relative error for the estimated perimeter and area of the object, respectively.

method followed by feature extraction methods on a real example. We use the coverage segmentation method presented in Section 4.2 to segment a microscope slide from a histomorphometrical study. Starting from the obtained coverage segmentation, we compute feature estimates. Comparison with results obtained by previously existing methods, and with results considered to be the ground truth, confirm that the proposed coverage model provides estimates with increased precision. Details of this work are presented in [53]. Here we give a brief summary.

The image shown in Fig. 23(a) is a part of material used in a histomorphometrical study described in [47]. It contains three regions: a screw-shaped implant, bone region, and soft tissue. Quantification is performed by measuring the length of the contact between the implant and the bone region, relative to the overall length of the implant border, and by measuring the percentage of bone area in the vicinity of

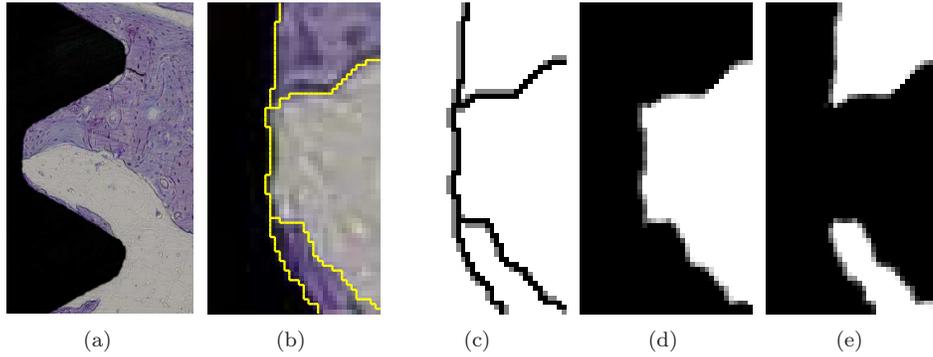


FIGURE 23. (a): The screw-shaped implant (black), bone (purple with a number of hollow spaces) and soft tissue (light blue). (b) Part of a crisp (manual) segmentation of (a) into the three regions. (c) The set B of (grey and black) pixels. Partial coverage values are assigned to the black pixels. (d) and (e) Pixel coverage segmentations of the soft tissue and the bone region, respectively.

the implant; for a detailed description see [47]. Measurements obtained manually, by an expert using integrated microscope software and with higher magnification available, are used as a ground truth.

We applied the proposed coverage segmentation method to segment the RGB image; this example is an illustration of applicability of the proposed method to multi-channel images. First step was to perform a crisp segmentation of the image. For this illustrative example, manual crisp segmentation is used, to get a good starting segmentation and to not mix errors from the crisp segmentation process with errors from the pixel coverage estimation. A part of the manual segmentation is presented in Fig. 23(b). The extracted set B of pixels to be re-evaluated in the process is shown in Fig. 23(c) as the union of black and grey pixels. The grey pixels are detected as pure in the thinning step and only the one-pixel thick 4-connected region in black is assigned partial coverage values. The result of the suggested pixel coverage segmentation method is presented in Fig. 23(d) and 23(e). The first presents segmented soft tissue, whereas the second shows segmented bone. Grey values, visible on the borders between the regions, correspond to partial coverage of pixels.

The aim of the study is to obtain bone-implant contact length estimates, as well as bone area estimates, which provide an improvement in terms of accuracy and precision, compared to those obtained in [47]. We apply the length estimation method presented in [52], and the area estimation method presented and analysed in [50], to the coverage segmentation. While the estimation of area is straightforward, some adjustments of the length estimation method are required. The method, as presented in [52], is applicable for estimation of the border between two classes, whereas in the observed example there may exist pixels which are partly covered

by three classes. To adjust the method appropriately, we observe the border line as being between the two classes—implant and nonimplant (soft tissue and bone together)—where the existing method is directly applicable. After the border line within a pixel is estimated, it is distributed to the two nonimplant classes proportionally to their coverage of that pixel. This approach is attractive due to its simplicity, and is acceptably accurate. Due to very low number of pixels in the image which are covered by three classes, potentially introduced errors have minimal impact on the result. Pixels covered by more than two classes are presumably rare in most applications.

TABLE 3. Feature estimation results for manual (using integrated microscope software), crisp [47], and the herein suggested method. See [47] for notation details.

Method	Contact length	Bone area R	Bone area M
Manual	79%	48%	78%
Crisp	88%	50%	81%
Suggested	85%	49%	81%

The results of feature estimates (based on length and area measurements) are presented in Table 3. It is clear that the proposed method provides feature estimates closer to the manual measurements, compared to the previously used methods. This, again, confirms the high performance and applicability of the proposed pixel coverage approaches.

7.3. Estimation of affine deformations of shapes. Image registration is an important task of image processing. Its goal is to find the geometric correspondence between images. Many approaches have been proposed for a wide range of problems in the past decades [69]. Shape matching, that is object registration based on geometry alone, and not on radiometric information, is a viable model when the image intensities are only weakly related and the relation between intensities of two images is hard to model; this happens in, e.g., multimodal registration (between images acquired by rather different imaging modalities) or when the image intensities undergo strong nonlinear deformations, e.g., in case of X-ray imaging. Shape matching requires an initial segmentation step, where the same region is segmented in the two images to match. This segmentation can be performed in a crisp or in a fuzzy way. In the following we present shape matching with improved precision based on coverage representations of shapes. This approach is described in details in [60, 61]. It gives additional evidence about advantages of using the coverage model to improve quality of image processing steps.

Domokos et al. proposed an extension [8] to the parametric estimation method of Francos et al. [17] to handle affine matching of crisp shapes. Estimation methods of this type have the advantage of providing accurate and computationally simple solution, avoiding both the need for finding point correspondences in the images as well as the need for computationally demanding optimization. Appropriate (global)

low-dimensional representations of a shape are instead utilized and correspondence between such representations is found by a direct computation. We have extended this approach to the case when the segmentation method is capable of producing a coverage segmentation instead of a classic crisp segmentation [60]. We know that the information preserved by using coverage representation may be successfully utilized to improve precision and accuracy of several shape descriptors. Precise moment estimation is essential for a successful application of the object registration method presented in [8] and the advantage of coverage representations is clearly noticeable in the represented study.

Observation that all individual correspondences between pairs of points of the two shapes, related by an affine transformation to be recovered, can be observed at the same time (instead of, e.g., only correspondences between some selected pairs), and can be integrated over the shape, is what the approach, developed for binary shape registration, and described in details in [8], relies on. Extending the established correspondences by applying appropriately chosen mappings to them, and integrating, a system of independent equations can be created. Unknowns of the system are the parameters of the applied unknown affine transform. For 2D shapes, six parameters are to be recovered, and therefore six equations are required in the system. It is observed that polynomial equations, for polynomials of at least second order, are simplest ones being at the same time linearly independent. The suggested system is therefore:

$$(7.1) \quad |\mathbf{A}| \int_{\mathcal{F}_t} x_k^n d\mathbf{x} = \sum_{i=0}^n \binom{n}{i} \sum_{j=0}^i \binom{i}{j} q_{k1}^{n-i} q_{k2}^{i-j} q_{k3}^j \int_{\mathcal{F}_o} y_1^{n-i} y_2^{i-j} d\mathbf{y},$$

where $k = 1, 2$; $n = 1, 2, 3$ and q_{ki} denote the unknown elements of the inverse transformation \mathbf{A}^{-1} with Jacobian $|\mathbf{A}|$.

This polynomial system is derived in the continuous space. However, digital image space provides only limited precision for these derivations and the integral can only be *approximated* by a discrete sum over the pixels. In [8] the Gauss centre point digitization is used. We explored whether using a coverage digitization would improve the registration performance. The coefficients of the system of equations in Eq. (7.1) are the first, second and third order geometric moments of the *template* and *observation*. Replacing them with their corresponding discrete approximations, we expect that the increased precision achieved from a coverage representation [50] will lead to improved registration performance.

Following the definition of discrete moments, the approximating discrete system of polynomial equations corresponding to Eq. (7.1) can now be produced:

$$|\mathbf{A}| \sum_{\mathbf{x} \in X_t} \mu_{F_t}(\mathbf{x}) p_k^n = \sum_{i=0}^n \binom{n}{i} \sum_{j=0}^i \binom{i}{j} q_{k1}^{n-i} q_{k2}^{i-j} q_{k3}^j \sum_{\mathbf{x} \in X_o} \mu_{F_o}(\mathbf{x}) p_1^{n-i} p_2^{i-j}.$$

Clearly, the spatial resolution of the images affects the precision of this approximation. We note that sufficient spatial resolution may be unavailable in real applications or may lead to too large amounts of data to be successfully processed within

the time constraints. On the other hand, it was shown in [50] that increasing the number of grey-levels ℓ , representing pixel coverage, by a factor r^2 provides asymptotically the same increase in precision as an r times increase of spatial resolution. This makes the suggested approach, utilizing increased membership resolution, a very powerful way to compensate for insufficient spatial resolution, while still preserving desired precision of moments estimates.

Evaluation of the transformation estimation—synthetic tests. Evaluation tests are first performed on a database of synthetic binary shapes. We examine the effect of the number of quantization levels on the precision of registration and compare results with the binary case. Pairs of corresponding synthetic fuzzy shapes are obtained by applying known affine transformations and the presented registration results for synthetic images are neither dependent nor affected by a segmentation method. This also means that the ground truth is available.

The data set consists of a number of different shapes and their transformed versions, a total of 2000 images. The transformation parameters (including rotations, translations, shear, and scaling) were randomly selected from uniform distributions. The templates are binary images, i.e., pixels in them are assigned coverage values either 0 or 1 (this corresponds to 1-bit representation). The coverage representations of the observation images are quantized and represented by integer values using k -bit ($k = 1, \dots, 8$) representation. Some typical examples of these images and their registration accuracies are shown in Figure 24.

In order to quantitatively evaluate the results, we use two error measures. The first error measure (denoted by ϵ) is the average distance in spels between the true (\mathbf{Ax}), and recovered ($\widehat{\mathbf{A}}\mathbf{x}$) positions of the transformed spels over the template. This measure can be used for evaluation only if the true transformation is known; this is the case in the tests on synthetic images. Another error measure is the absolute difference (denoted by δ) between the *registered* template image and the *observation* image

$$\epsilon = \frac{1}{|T|} \sum_{\mathbf{x} \in T} \|\mathbf{Ax} - \widehat{\mathbf{A}}\mathbf{p}\|, \quad \text{and} \quad \delta = \frac{|R \triangle O|}{|R| + |O|},$$

where $|T|$ is the number of *template* spels, \triangle denotes the symmetric difference, while R and O are the set of spels of the *registered* shape and the *observation* respectively. Before computing the errors, the images are binarized by taking the α -cut at $\alpha = 0.5$.

The medians of errors for both ϵ and δ are presented in Table 4 for different quantization levels. Experimental data confirm the theoretical results on increased precision of moments estimation based on coverage representation. Consequently, the registration results, compared to the binary case, are improved. It is important to notice that registration based on coverage representation may be applied for lower image resolutions, i.e. where the binary approach becomes unstable.

An important property of the proposed registration method is that, although based on solving a system of polynomial equations, it provides the result without any iterative optimization step. Its performance is based on the precision and

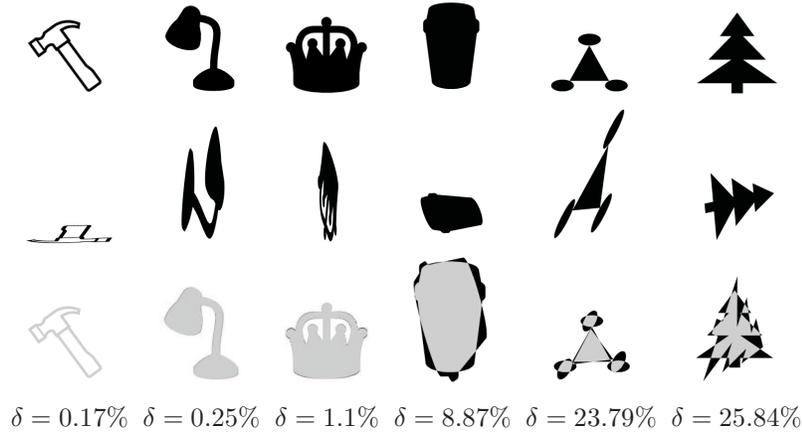
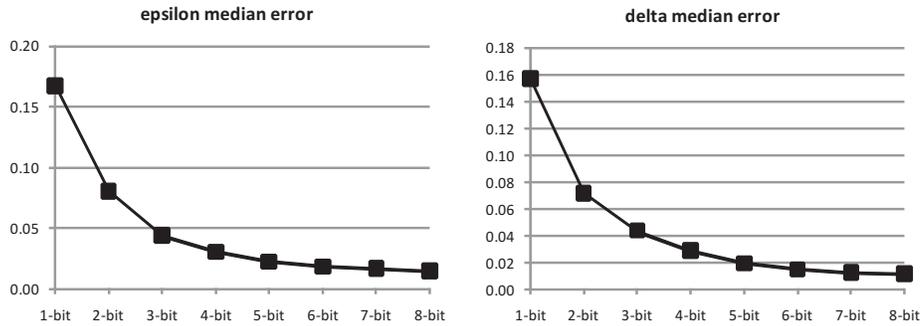


FIGURE 24. Examples of *templates* (top row) and *observations* (middle row) images. In the third row, grey pixels show where the registered images matched each other and black pixels show the positions of registration errors.

TABLE 4. Registration results of 2000 images using different quantization levels of the fuzzy boundaries.

	Fuzzy representation							
	1-bit	2-bit	3-bit	4-bit	5-bit	6-bit	7-bit	8-bit
ϵ median (pixels)	0.168	0.080	0.0443	0.0305	0.0225	0.0186	0.0169	0.0147
δ median (%)	0.157	0.072	0.0439	0.0292	0.0196	0.0151	0.0125	0.0116
Registered	1905	1919	1934	1943	1933	1929	1925	1919
Not registered	95	80	66	57	67	71	75	81



accuracy of moments estimates. The time complexity of the method is $\mathcal{O}(N)$, where N is the number of the pixels of the image, enabling real time registration of shapes.

Experiments on real X-ray images. Hip replacement is a surgical procedure in which the hip joint is replaced by a prosthetic implant. In the short post-operative time, infection is a major concern. An inflammatory process may cause bone resorption and subsequent loosening or fracture, often requiring revision surgery. In current practise, clinicians assess loosening by inspecting a number of post-operative X-ray images of the patient’s hip joint, taken over a period of time. Obviously, such an analysis requires the registration of X-ray images. Even visual inspection benefit from registration, as clinically significant prosthesis movement can be very small.

There are two main challenges in registering hip X-ray images: One is the highly nonlinear radiometric distortion [12] which makes any grey-level-based method unstable. Fortunately, the segmentation of the prosthetic implant is quite straightforward [37] so shape registration is a valid alternative here. Herein, we used an appropriate coverage segmentation method to segment the implant. The second problem is that the true transformation is a projective one which depends also on the position of the implant in 3D space. Indeed, there is a rigid-body transformation in 3D space between the implants, which becomes a projective mapping between the X-ray images. Fortunately, the affine assumption is a good approximation here, since the X-ray images are taken in a well defined *standard position* of the patient’s leg.

For the diagnosis, the area around the implant (especially the bottom part of it) is the most important for the physician. It is where the registration must be the most precise. Fig. 25 shows some registration results. Since the best aligning transformation is not known, only the δ error measure can be evaluated. We also note, that in real applications the δ error value accumulates the registration error and the segmentation error.

The preliminary results obtained on real X-ray images of hip prosthetic implants taken during post-operative controls are, in our opinion, very encouraging; they show that our approach using coverage segmentation and subsequent registration as described above can be used in real applications. Further research on possible improvements, generalizations, and thorough evaluation of these initial studies is in progress.

7.4. High resolution reconstruction—two examples. Some examples of the performance of our proposed high resolution reconstruction method are shown already in Section 6. 2D objects are observed, and one illustrations of influence of selection of features in feature representation to the reconstruction result is given. More elaborate evaluation of various relevant issues of defuzzification is presented in [54]. In this section we present two examples of application of the proposed method on 3D images. Both examples are within medical imaging, imaging modalities are different (CT and X-rays), sizes and complexity of data differ (first one is a solid object, of a rather simple structure, second is an object of a high complexity, but smaller in data size). Issues of interest are choice of an optimization method that gives a good result at a reasonable cost, selection of a starting position for the optimization, decision about topology preservation constraint, selection of (scales of)

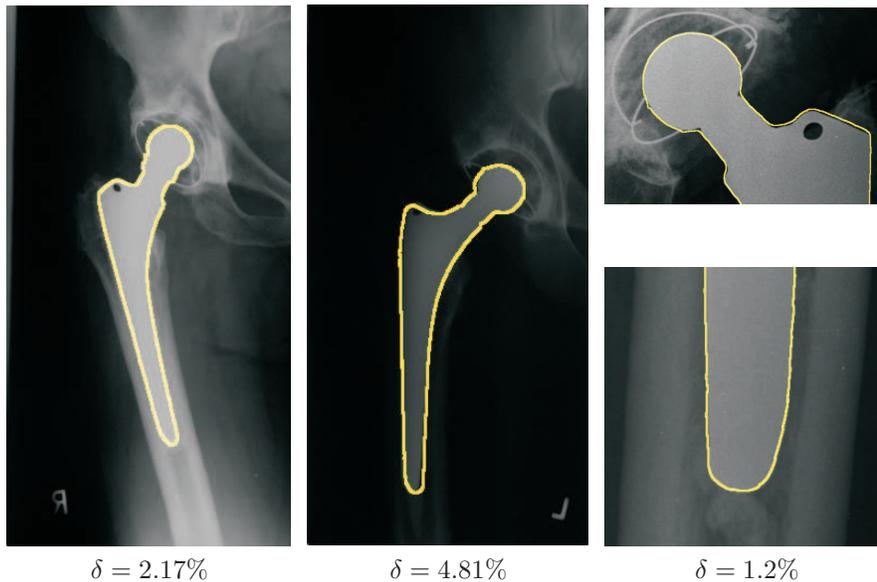


FIGURE 25. Real X-ray registration results. (a) and (b) show full X-ray observation images and the outlines of the registered template shapes. (c) shows a close up view of a third study around the top and bottom part of the implant.

features to be used in feature vector representation. Both examples show applicability and very good performance of the proposed method. Topological constraints, size of the data, complexity of the task are all addressed well by the presented method.

3D CT image of a bone implant. First example of high resolution reconstruction is performed on a 3D object presented in Figure 26. The data volume is a CT image of a bone implant (inserted in a leg of a rabbit). We applied the method to a part of the image (Figure 26(a) shows a slice through the volume) containing a connected piece of bone area (dark grey), surrounded by a nonbone area (light grey). Figure 26(b) shows a slice through a 3D fuzzy set representing the bone region.

All features are matched well in this example; there are no large regions of high fuzziness, and the global features do not provide any reason for “transportation” of volume as in the example in Section 6. Defuzzifications with or without meso-scale features are therefore practically identical. No topological constraints are required, since this object appears to be simple enough for the process to handle the topology automatically. Main challenge can be seen in the size of a data set (even though this particular example is made small enough); regarding this issue a selection of an optimization method is of rather high importance. The result presented here is obtained by simulated annealing, but increasing difficulties in optimization,

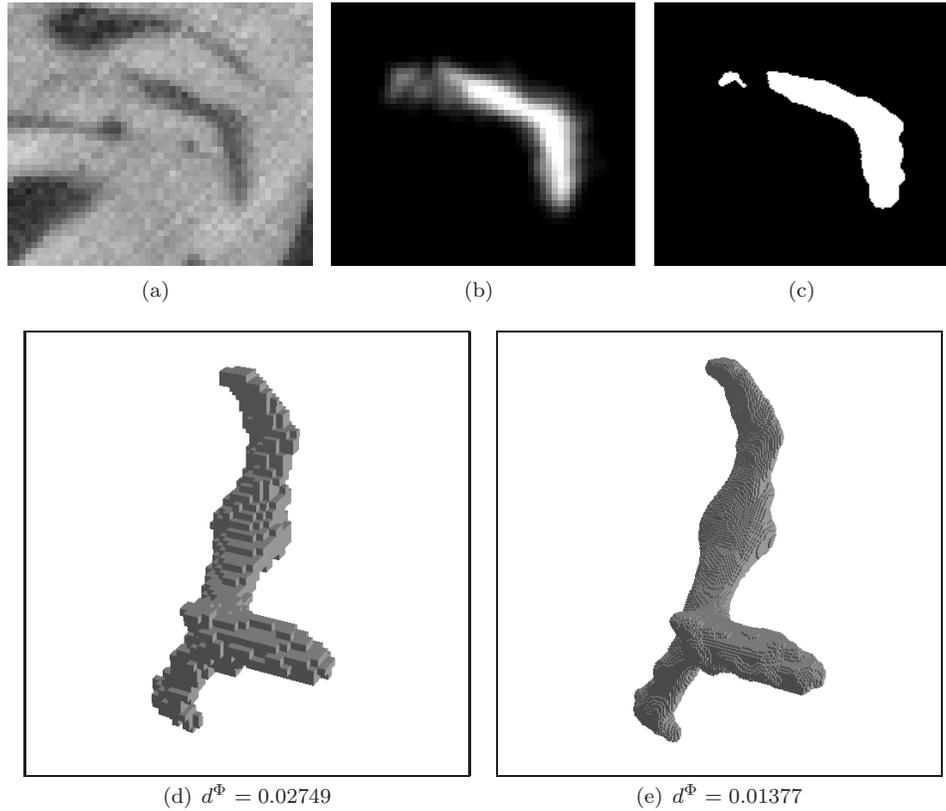


FIGURE 26. Defuzzification of a part of a 3D image of a bone implant. (a) Slice through the image volume. The dark grey area is bone, the light parts are nonbone areas. (b) Slice through a fuzzy segmentation of the bone region in the image volume. (c) Slice through a defuzzification, using meso-scale volume features, of the fuzzy segmented image volume. (d) 3D rendering of the α -cut at smallest feature distance to the fuzzy object. (e) 3D rendering of a high resolution defuzzification of the fuzzy segmented object. A four times scaled up version of the best α -cut (d) was used as starting set for the simulated annealing search.

caused by increasing size of data, motivated our research on alternative options for optimization strategy. Results are presented in [27, 32]. Starting position can influence result significantly; a good choice, used in the example shown in Figure 26, is an optimal α -cut.

High resolution reconstruction of X-ray image of vessels. Fuzzy representations, and coverage representation in particular, of image objects are especially

useful when the spatial resolution is too low to provide a good crisp representation. One such situation can be seen in Figure 27(b), which displays a maximum intensity projection of a part of a rotational b-plane X-ray scan of the arteries of the right half of a human head (provided by Philips Research, Hamburg, Germany), shown in Figure 27(a). A contrast agent is injected into the blood and an aneurism is shown to be present. The intensity values of the image voxels correspond fairly well with partial volume coverage, and are therefore used directly as coverage membership values.

This example image violates the sampling theorem; the vessels imaged are not resolved since they are less than one voxel thick. This fact causes a number of problems related to information extraction. Using a priori knowledge about the image, it is still possible to obtain a reasonable high resolution reconstruction. One such a priori piece of information is the knowledge that the vessel tree is simply connected. Starting from one simply connected component, and preserving topology [2] throughout the search, it is provided that the obtained crisp representation (reconstruction) is also simply connected.

Centroid position is not an intuitive feature to use for defuzzification of a vessel tree. It may interfere in undesirable ways with the topology preservation during the search procedure, so we exclude centroid from the feature representation in this example.

It is clear that high resolution reconstruction is really needed here; any crisp representative at the same resolution as the original image would be a rather bad representation; to preserve the volume of the fuzzy image, many parts of the vessel tree would not be included in the crisp set.

Performing reconstruction at two times the original resolution, we get the result presented in Figure 27(c). The result is not visually appealing, due to severe underestimation of the surface area of a crisp thin (less than one voxel thick) structure by the surface area of the fuzzy set. This problem is not present for a crisp object whose fuzzy representation is obtained at sufficiently high resolution and contains points with memberships equal to one in the interior of the object. In the case presented in Figure 27, however, the reconstruction using the inaccurate surface area estimate fails to preserve the vessel structure.

It would be of high interest to have a better surface area estimate for the defuzzification; study about this feature estimate is certainly included in our future work. In the absence of such, we perform reconstruction without the surface area feature. Using only volume based information (at a range of scales) the high resolution reconstruction is fairly unconstrained, which leads to the rather jagged result of Figure 27(d). Dropping the meso-scale feature from the feature representation, we get the result presented in Figure 27(e).

We note that, although not visible in Figure 27, the topology is in deed preserved; all the resulting objects are simply connected. However, the vessels may not always be connected in a correct way, so some additional information on how vessels branch and bend may be required in this case.

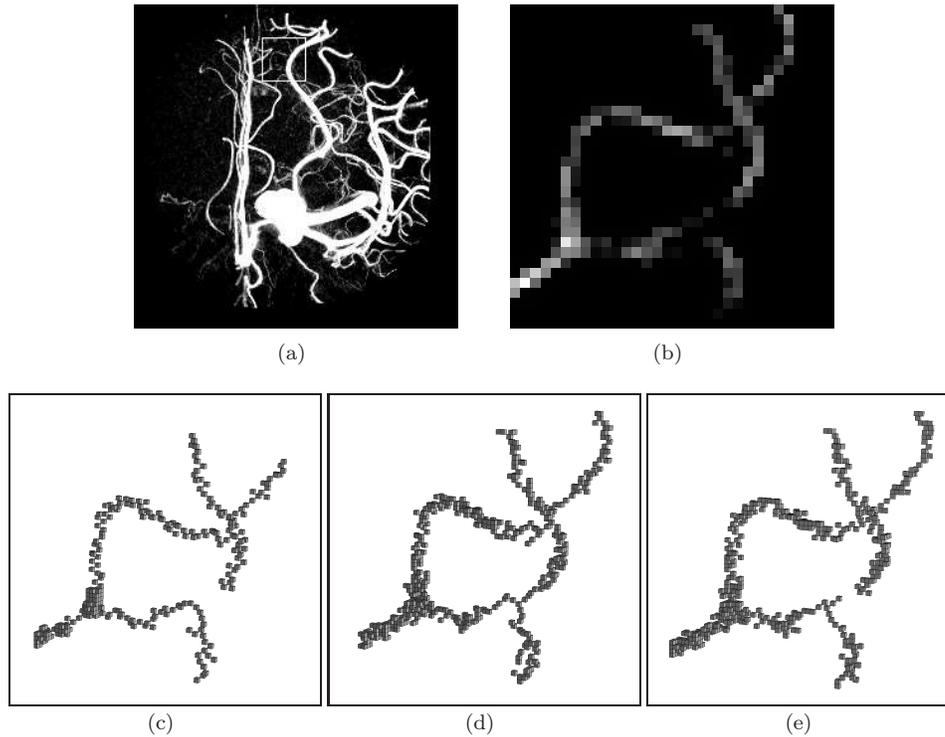


FIGURE 27. Defuzzification of a selected part of an angiography 3D image, showing the arteries of the right half of a human head. (a) Maximum intensity projection through the image volume; the white square in the upper part of the image indicates the location of the selected part of the volume that is defuzzified in this example. (b) Maximum intensity projection through the selected part the volume. (c) 3D rendering of a defuzzification at twice the resolution using volumes of all scales and surface area. (d) 3D rendering of a defuzzification at twice the resolution using only volumes of all scale. (e) 3D rendering of a defuzzification at twice the resolution using only global and local volumes.

8. Conclusions and future work

In the field of computerized image processing and analysis, increasing attention is lately given to methods that provide results with sub-pixel precision. Such methods are especially important for applications where precision is a key factor, for example in medicine. When working with 3D images (CT, MR), systematic errors of the size of a pixel may accumulate to unacceptably large overall errors in the final results. In addition, analysis of images at low resolution is constantly a hot research topic;

with the recent progress in imaging techniques, allowing imaging to reach nanometer scales, a previously inaccessible world of structures of sizes all the way down to molecular scale, opens up.

The work presented in this chapter is a summary of our research on development of image processing methods that provide results at sub-pixel precision; this research was conducted during last couple of years. In our opinion, it is of highest importance for the field of image processing to go on with this type of research and with further development of high precision image processing tools and algorithms; a call imposed by the technological development is already shown to be well addressed by this very promising research track with obvious high applicability.

Object representations that facilitate utilization of sub-pixel precision methods, and particularly those based on spel coverage by an object, were in our research focus throughout the recent years. We developed segmentation methods characterized by generality and wide applicability, as well as some important feature estimators. We explored crisp object reconstruction methods based on the derived feature estimators and we showed that information preserved in coverage representation can be successfully used to compensate for lacking spatial resolution, both in estimations, and in reconstructions. The following natural step in research is development of more methods for analysis of images segmented by some of the coverage segmentation methods, in an attempt to make a complete tool-box of processing methods for images providing sub-pixel information precision. Tasks related to development of methods for feature estimation and object description, primarily in 3D, will be first addressed. Our intention is to develop strong theoretical background for every method we suggest, and to test and prove their applicability on real world tasks and challenges.

Application fields for our developed methods, as well as those that are to be developed, are numerous. Medical applications have already been addressed, and increased precision of the methods has shown to be of high importance and benefit for them. We intend to continue with development of image registration methods; an additional research track will be improvement of distance measures so that they are applicable to objects represented at sub-pixel precision. Developed appropriate distances, together with new high precision features (descriptors) will find applications in content retrieval. Studies and utilization of sub-pixel methods in the field of optical character recognition (OCR), being of high interest in a wide range of fields, including, e.g., digitization of cultural heritage, are also envisioned in our future work.

Finally, an application field that is particularly in our interest for future work, is the field of biometrics. Biometrics, which in a broad sense refers to the science and technology of measuring and statistically analyzing biological data, and therefore already concerns the presented medical applications of our work, is attracting more and more attention also outside the fields of medicine and biomedicine, e.g., in (information) security, where it finds use for identification and authentication purposes. Biometrics, provides a variety of research challenges; increased precision

of representations and analysis methods are certainly of highest importance for biometrics applications, and we feel that our approaches, involving coverage (sub-pixel precision) models, may provide improvements for methods used in biometrics tasks.

Biometrics considers methods for uniquely recognizing humans based upon one or more intrinsic physical or behavioural properties/characteristics. In particular, biometrics is used as a form of identity access control, being of highest importance in (information) security related issues. Applications of biometrics recognition systems include computer systems security, secure electronic banking, mobile phones, credit cards, secure access to buildings, health and social services. As opposed to several traditionally used access control tools, such as passwords and ID cards, biometric can not be borrowed, stolen, or forgotten, which makes it increasingly popular in security systems. Nowadays, most often used biometrics characteristics and methods are fingerprint, face, DNA, palm print, hand geometry, and iris recognition.

Biometric recognition assumes representation of individuals by a feature vector derived from their physiological and/or behavioural characteristic; automatic recognition of these individuals becomes recognition of their representations. Feature vectors are usually stored in appropriate databases. Searching procedures and matching algorithms are typical tasks involved. Our so far developed methods involve feature extraction algorithms, feature selection and object representations, geometric matching of objects, as well as studies related to distance measures, [26]. Performance of biometrics systems strongly depends on distinctiveness of the representations, as well as their accuracy and precision. It should be possible to measure/estimate the selected features accurately every time when data is required and acquired, which demands high precision of estimation algorithms used: it is also important to have distance/similarity measures and matching algorithms offering high precision performance, to ensure full discriminative power of the system. All these requirements are met by our proposed concept of sub-pixel (coverage) model. We find very promising to adjust existing algorithms and propose new ones to address the tasks of this rapidly developing and highly important field.

Acknowledgement. This work is partially supported by the Serbian Ministry of Education and Science, the projects ON174008, “Advanced Techniques of Cryptology, Image Processing and Computational Topology for Information Security”, and III44006, “Development of new information and communication technologies, based on advanced mathematical methods, with applications in medicine, telecommunications, power systems, protection of national heritage and education”.

References

- [1] I. Bloch, *Fuzzy spatial relationships for image processing and interpretation: A review*, Image Vis. Comput. **23** (2005), 89–110.
- [2] G. Borgfors, I. Nyström, and G. Sanniti di Baja, *Connected components in 3D neighbourhoods*, in: *Proc. of Scand. Conf. on Image Analysis*, Pat. Rec. Soc. of Finland, 1997, 567–572.
- [3] N. Bousion, M. Hatt, A. Reilhac, and D. Visvikis, *Fully automated partial volume correction in PET based on a wavelet approach without the use of anatomical information*, in: *Proc. of IEEE Nuclear Science Symp.*, IEEE, 2007, 2812–2816.

- [4] J. Chanussot, I. Nyström, and N. Sladoje, *Shape signatures of fuzzy sets based on distance from the centroid*, Pattern Recognit. Lett. **26(6)** (2005), 735–746.
- [5] H. S. Choi, D. R. Haynor, and Y. Kim, *Partial volume tissue classification of multichannel magnetic resonance images—a mixed model*, IEEE Trans. on Medical Imaging **10** (1991), no. 3, 395–407.
- [6] D. Coeurjolly and R. Klette, *A comparative evaluation of length estimators of digital curves*, IEEE Trans. Pattern Anal. Mach. Intell. (2004), 252–258.
- [7] H. Davenport, *On a principle of Lipschitz*, J. London Math. Soc. **s1-26** (1951), no. 3, 179–183.
- [8] C. Domokos, Z. Kato, and J. M. Francos, *Parametric estimation of affine deformations of binary images*, in: *Proc. of Int. Conf. on Acoustics, Speech, and Signal Process.* (Las Vegas, USA), IEEE, April 2008, 889–892.
- [9] L. Dorst and A. W. M. Smeulders, *Length estimators for digitized contours*, Comput. Vis. Graph. Image Process. **40** (1987), 311–333.
- [10] D. Dubois and M.-C. Jaulent, *A general approach to parameter evaluation in fuzzy digital pictures*, Pattern Recognit. Lett. **6** (1987), 251–259.
- [11] D. Eberly and J. Lancaster, *On gray scale image measurements: I. Arc length and area*, CVGIP: Graph. Models Image Process. **53** (1991), no. 6, 538–549.
- [12] C. Florea, C. Vertan, and L. Florea, *Logarithmic model-based dynamic range enhancement of hip X-ray images*, in: *Proc. of Advanced Concepts for Intell. Vision Systems* (Delft, Netherlands), LNCS, vol. 4678, Springer-Verlag, 2007, 587–596.
- [13] G. M. Foody, *Sharpening fuzzy classification output to refine the representation of sub-pixel land cover distribution*, Int. J. Remote Sensing **19** (1998), no. 13, 2593–2599.
- [14] H. Freeman, *On the encoding of arbitrary geometric configurations*, IRE Trans. Electron. Comput. **EC-10** (1961), 260–268.
- [15] L. Grady and M.-P. Jolly, *Weights and topology: A study of the effects of graph construction on 3D image segmentation*, in: *Proc. of Int. Conf. on Medical Image Computing and Comput. Assisted Intervention*, vol. 1, Springer-Verlag, 2008, 153–161.
- [16] L. Grady, *Space-variant machine vision—a graph theoretic approach*, Ph.D. thesis, Boston University, 2004.
- [17] R. Hagege and J. M. Francos, *Linear estimation of sequences of multi-dimensional affine transformations*, in: *Proc. of Int. Conf. on Acoustics, Speech and Signal Process.* (Toulouse, France), vol. 2, IEEE, 2006, 785–788.
- [18] M. K. Hu, *Visual pattern recognition by moment invariants*, IRE Trans. Inform. Theory **8** (1962), 179–187.
- [19] J. Jua, E. D. Kolaczykb, and S. Gopa, *Gaussian mixture discriminant analysis and sub-pixel land cover characterization in remote sensing*, Remote Sensing and Environment **84** (2003), no. 4, 550–560.
- [20] N. Kiryati and A. Bruckstein, *Grey levels can improve the performance of binary image digitizers*, CVGIP: Graph. Models Image Process. **53** (1991), no. 1, 31–39.
- [21] R. Klette, V. Kovalevsky, and B. Yip, *Length estimation of digital curves*, in: *Proc. of SPIE Vision Geometry VIII*, 1999, (expanded version available: Machine Graphics & Vision, vol. 9, pp. 673-703, 2000.), 117–129.
- [22] R. Klette and J. Žunić, *Multigrid convergence of calculated features in image analysis*, J. Math. Imaging Vis. **13** (2000), 173–191.
- [23] Z. Kulpa, *Area and perimeter measurement of blobs in discrete binary pictures*, Comput. Graphics and Image Process. **6** (1977), 434–454.
- [24] W. Van Leekwijck and E. E. Kerre, *Defuzzification: Criteria and classification*, Fuzzy Sets Syst. **108** (1999), 159–178.
- [25] K. Van Leemput, F. Maes, D. Vandermeulen, and P. Suetens, *A unifying framework for partial volume segmentation of brain MR images*, IEEE Trans. on Medical Imaging **22** (2003), no. 1, 105–119.

- [26] J. Lindblad, V. Ćurić, and N. Sladoje, *On set distances and their application to image registration*, in: *Proc. of Int. Symp. on Image and Signal Process. and Analysis* (Salzburg, Austria), IEEE, 2009, 449–454.
- [27] J. Lindblad, T. Lukić, and N. Sladoje, *Defuzzification by feature distance minimization based on DC programming*, in: *Proc. of Int. Symp. on Image and Signal Process. and Analysis* (Istanbul, Turkey), IEEE, 2007, 373–378.
- [28] J. Lindblad and N. Sladoje, *Feature based defuzzification at increased spatial resolution*, in: *Proc. of Int. Workshop on Combinatorial Image Analysis* (Berlin, Germany), LNCS, vol. 4040, Springer-Verlag, 2006, 131–143.
- [29] J. Lindblad, N. Sladoje, V. Ćurić, H. Sarve, C. B. Johansson, and G. Borgefors, *Improved quantification of bone remodelling by utilizing fuzzy based segmentation*, in: *Proc. of Scand. Conf. on Image Analysis* (Oslo, Norway), LNCS, vol. 5575, Springer-Verlag, 2009, 750–759.
- [30] J. Lindblad, N. Sladoje, and T. Lukić, *Feature based defuzzification in Z^2 and Z^3 using a scale space approach*, in: *Proc. Int. Conf. on Discrete Geometry for Comput. Imagery* (Szeged, Hungary), LNCS, vol. 4245, Springer-Verlag, 2006, 379–390.
- [31] S. Loncaric, *A survey of shape analysis techniques*, *Pattern Recognition* **31** (1998), no. 8, 983–1001.
- [32] T. Lukić, N. Sladoje, and J. Lindblad, *Deterministic defuzzification based on spectral projected gradient optimization*, in: *Proc. of Symp. of the German Association for Pattern Recognition* (Munich, Germany), LNCS, vol. 5096, Springer-Verlag, 2008, 476–485.
- [33] F. Malmberg, J. Lindblad, and I. Nyström, *Sub-pixel segmentation with the Image Foresting Transform*, in: *Proc. of Int. Workshop on Combinatorial Image Analysis* (Playa del Carmen, Mexico), LNCS, vol. 5852, Springer-Verlag, 2009, 201–211.
- [34] F. Malmberg, J. Lindblad, N. Sladoje, and I. Nyström, *A graph-based framework for sub-pixel image segmentation*, *Theor. Comput. Sci.* **412** (2011), no. 15, 1338–1349.
- [35] F. Malmberg, I. Nyström, A. Mehnert, C. Engstrom, and E. Bengtsson, *Relaxed Image Foresting Transforms for interactive volume image segmentation*, in: *Proc. of SPIE Medical Imaging*, vol. 7623, SPIE, 2010.
- [36] W. J. Niessen, K. L. Vincken, J. Weickert, B. M. ter Haar Romeny, and M. A. Viergever, *Multiscale segmentation of three-dimensional MR brain images*, *Int. J. Comput. Vision* **31** (1999), no. 2/3, 185–202.
- [37] A. Oprea and C. Vertan, *A quantitative evaluation of the hip prosthesis segmentation quality in X-ray images*, in: *Proc. of Int. Symp. on Signals, Circuits and Systems* (Iasi, Romania), vol. 1, IEEE, 2007, 1–4.
- [38] N. Otsu, *A threshold selection method from gray-level histograms*, *IEEE Trans. on System Man and Cybernetics* **9** (1979), no. 1, 62–66.
- [39] D. Proffit and D. Rosen, *Metrication errors and coding efficiency of chain-encoding schemes for the representation of lines and edges*, *Comput. Graphics and Image Process.* **10** (1979), 318–332.
- [40] L. Rondeau, R. Ruelas, L. Levrat, and M. Lamotte, *A defuzzification method respecting the fuzzification*, *Fuzzy Sets Syst.* **86** (1997), 311–320.
- [41] A. Rosenfeld, *The fuzzy geometry of image subsets*, *Pattern Recognit. Lett.* **2** (1984), 311–317.
- [42] A. Rosenfeld, *Fuzzy geometry: An updated overview*, *Information Sciences* **110** (1998), no. 3–4, 127–133.
- [43] A. Rosenfeld and S. Haber, *The perimeter of a fuzzy subset*, *Pattern Recognition* **18** (1985), 125–130.
- [44] E. Roventa and T. Spircu, *Averaging procedures in defuzzification processes*, *Fuzzy Sets Syst.* **136** (2003), 375–385.
- [45] E. H. Ruspini, *A new approach to clustering*, *Inf. Control* **15** (1969), 22–32.
- [46] P. Santago and H. D. Gage, *Quantification of MR brain images by mixture density and partial volume modeling*, *IEEE Trans. on Medical Imaging* **12** (1993), no. 3, 566–574.

- [47] H. Sarve, J. Lindblad, C. B. Johansson, G. Borgefors, and V. F. Stenport, *Quantification of bone remodeling in the proximity of implants*, in: *Proc. of Int. Conf. on Comput. Analysis of Images and Patterns* (Vienna, Austria), LNCS, vol. 4673, Springer-Verlag, 2007, 253–260.
- [48] N. Sladoje, *A straight line segment estimation by using discrete moments*, in: *Proc. of XIII Conf. on Applied Mathematics* (Igalo, Yugoslavia), 1998, 121–129.
- [49] N. Sladoje, *On analysis of discrete spatial fuzzy sets in 2 and 3 dimensions*, Ph.D. thesis, Swedish University of Agricultural Sciences, Uppsala, 2005.
- [50] N. Sladoje and J. Lindblad, *Estimation of moments of digitized objects with fuzzy borders*, in: *Proc. of Int. Conf. on Image Analysis and Process.* (Cagliari, Italy), LNCS, vol. 3617, Springer-Verlag, 2005, 188–195.
- [51] N. Sladoje and J. Lindblad, *Perimeter estimation based on grey level object representation*, Internal Report 33, Centre for Image Analysis, Uppsala, Sweden, 2008, Available from the authors.
- [52] N. Sladoje and J. Lindblad, *High-precision boundary length estimation by utilizing gray-level information*, *IEEE Trans. Pattern Anal. Mach. Intell.* **31** (2009), no. 2, 357–363.
- [53] N. Sladoje and J. Lindblad, *Pixel coverage segmentation for improved feature estimation*, in: *Proc. of Int. Conf. on Image Analysis and Process.* (Vietri sul Mare, Italy), LNCS, vol. 5716, Springer-Verlag, 2009, 923–938.
- [54] N. Sladoje, J. Lindblad, and I. Nyström, *Defuzzification of spatial fuzzy sets by feature distance minimization*, *Image Vis. Comput.* **29** (2011), no. 2-3, 127–141.
- [55] N. Sladoje, I. Nyström, and P. K. Saha, *Measurements of digitized objects with fuzzy borders in 2D and 3D*, *Image Vis. Comput.* **23** (2005), 123–132.
- [56] M. Soret, S. L. Bacharach, and I. Buvat, *Partial-volume effect in PET tumor imaging*, *The J. Nuclear Medicine* **48** (2007), no. 6, 223–235.
- [57] A. Souza, J. K. Udupa, and P. K. Saha, *Volume rendering in the presence of partial volume effects*, *IEEE Trans. on Medical Imaging* **24** (2005), no. 2, 223–235.
- [58] R. Strand, *Distance functions and image processing on point-lattices*, Ph.D. thesis, Uppsala University, 2008.
- [59] M. Tajine and A. Daurat, *On local definitions of length of digital curves*, in: *Proc. of Int. Conf. on Discrete Geometry for Comput. Imagery* (Naples, Italy), LNCS, vol. 2886, Springer-Verlag, Nov. 2003, 114–123.
- [60] A. Tanács, C. Domokos, N. Sladoje, J. Lindblad, and Z. Kato, *Recovering affine deformations of fuzzy shapes*, in: *Proc. of Scand. Conf. on Image Analysis* (Oslo, Norway), LNCS, vol. 5575, Springer-Verlag, 2009, 735–744.
- [61] A. Tanacs, J. Lindblad, N. Sladoje, and Z. Kato, *Estimation of linear deformations of 3D objects*, in: *Proc. of IEEE Int. Conf. on Image Process.* (Hong Kong), IEEE, 2010, 153–156.
- [62] J. K. Udupa and G. J. Grevera, *Go digital, go fuzzy*, *Pattern Recognit. Lett.* **23** (2002), 743–754.
- [63] P. W. Verbeek and L. J. van Vliet, *Estimators of 2D edge length and position, 3D surface area and position in sampled grey-valued images*, *Bioimaging* **1** (1993), 47–61.
- [64] B. Verwer, *Local distances for distance transformations in two and three dimensions*, *Pattern Recognit. Lett.* **12** (1991), 671–682.
- [65] K. L. Vincken, A. S. E. Koster, and M. A. Viergever, *Probabilistic segmentation of partial volume voxels*, *Pattern Recognit. Lett.* **15** (1994), 477–484.
- [66] A. M. Vossepoel and A. W. M. Smeulders, *Vector code probability and metrication error in the representation of straight lines of finite length*, *Comput. Graphics Image Process.* **20** (1982), 347–364.
- [67] J. Žunić and N. Sladoje, *Efficiency of characterizing ellipses and ellipsoids by discrete moments*, *IEEE Trans. Pattern Anal. Mach. Intell.* **22** (2000), no. 4, 407–414.
- [68] L. Zadeh, *Fuzzy sets*, *Inf. Control* **8** (1965), 338–353.
- [69] B. Zitová and J. Flusser, *Image registration methods: A survey*, *Image Vis. Comput.* **21** (2003), no. 11, 977–1000.