Semi-automated segment generation for geographic novelty detection using edge and area metrics

Christoff Fourie[1], Adriaan van Niekerk[2], Ladislav Mucina[3]

[1]Department of Geography and Environmental Studies, Stellenbosch University, Stellenbosch, South Africa
[2]Centre for Geographical Analysis, Stellenbosch University, Stellenbosch, South Africa, avn@sun.ac.za
[3]Department of Botany & Zoology, Stellenbosch University, Stellenbosch, South Africa & Department of Environment and Agriculture, Curtin Institute for Biodiversity and Climate, Curtin University, Perth, Australia

Abstract

An approach to generating accurate image segments for land-cover mapping applications is to model the process as an optimisation problem. Area-based empirical discrepancy metrics are used to evaluate instances of generated segments in the search process. An edge metric, called the pixel correspondence metric (PCM), is evaluated in this approach as a fitness function for segmentation algorithm free-parameter tuning. The edge metric is able to converge to user-provided reference segments in an earth observation mapping problem when adequate training data are available. Two common metaheuristic search functions were tested, namely particle swarm optimisation (PSO) and differential evolution (DE). The edge metric is compared with an area-based metric, regarding classification results of the land-cover elements of interests for an arbitrary problem. The results show the potential of using edge metrics, as opposed to area metrics, for evaluating segments in an optimisation-based segmentation algorithm parameter-tuning approach.

1 Introduction

Object-based image analysis (OBIA) involves the grouping of pixels of similar spectral and spatial properties into segments (or objects). These segments are then analysed and classified rather than the individual pixels (Navulur, 2006). Generating accurate image segments is, however, challenging and image segmentation is a research topic that has gained importance in the remote sensing community, mainly due to the advent of very-high resolution (VHR) satellite imagery and the popularisation of object-based image analysis (OBIA) approaches (Blaschke, 2010). For VHR imagery it has been shown that the spatial aggregation and subsequent attribution of pixel groups improves classification and interpretation (Baatz et al., 2008; Lang, 2008; Blaschke, 2010; Addink, Van Coillie & De Jong, 2012). However, a poor segmentation often has a negative impact on classification accuracy (Navulur, 2006).

Several automated and semi-automated image segmentation methods are available for generating image objects in a remote sensing context. For example, Hay et al.'s (2005) unsupervised process uses measures of empirical goodness to select optimal "scales" for generating multi-scale image

segments. In contrast, Zhang & Maxwell (2006) proposes a semi-automated, parameter-tuning system functioning with a region-merging segmentation. Their workflow starts by over-segmenting an image, then all the segments comprising a feature of interest are selected by the user and a fuzzy inference system is used to predict the appropriate set of parameters that result in the segmentation of the selected subsegments as a single segment.

Osman et al. (2009) developed a semi-automated segmentation-generating system that operates on local-only features. The approach performs segmentation on input data transformed by a maximum-margin classifier. The system asks the user to digitise a limited number of lines within the feature of interest, as well as a few lines on the outside of the feature. The system automatically transforms the input data, followed by unsupervised clustering of pixels in the input samples and a separation of clusters via support vector machine (SVM). Finally, a region-growing algorithm is applied, with seed pixels in the feature of interest, to generate the segment.

Another approach to image segmentation is to model the generation of segments as an optimisation or search problem (Bhanu et al., 1995; Pignalberi et al., 2003; Feitosa et al., 2006; Fredrich & Feitosa, 2008; Lübker & Schaab, 2009; Derivaux et al., 2010). The process begins with a reference set of desired optimal output or segments, after which a search method is invoked that traverses the parameter space of the segmentation algorithm, searching for results most closely resembling the user-provided reference. Because segmentation is typically computationally expensive, population-based stochastic search methods, such as genetic algorithms, are recommended to produce usable results in a relatively short time.

Different categories of methods (even methods within the same category) that measure segment quality emphasise different aspects when evaluating segments. The ability of a search heuristic to find a suitable parameter set for a given task depends on the ability of discrepancy measures to correctly – as perceived by the users of the segments – gauge the quality of the generated segments. A specific metric (e.g. edge, area, data clustering or hybrid) may not be optimal for all types of features (Weidner, 2008). The amount and quality of training information also influences the performance of the search heuristic (Feitosa et al., 2006). Furthermore, results depend on the intrinsic ability of the search heuristic to effectively traverse the search surface created by the fitness function.

Area metrics, such as the larger segments booster (LSB) (Fredrich & Feitosa, 2008) or object-level consistency error (Polak et al., 2009), measure the correspondence of segment areas. The use of area metrics to guide the search for optimal segmentation parameters is well documented (Derivaux et al., 2010; Feitosa et al., 2010), whereas edge-based metrics have only been used as methods to judge segment quality (Prieto & Allen, 2003; Weidner, 2008) and not as metrics to guide the search for optimal segmentation algorithm parameters. For a given area, edge metrics generate segment borders, irrespective of the features of interest. When using an edge metric, a user

is required to delineate the perceived optimal segments in a specific area and to have knowledge of the behaviour of the specific algorithm. An area metric only requires the accurate delineation of features of interest, an exercise that is not always objectively possible.

This paper reports on an edge metric used as a fitness function and which converges to synthetic (reachable) and user-provided reference segments. A case study illustrates that the edge metric generates segments more closely resemble the user-provided reference segments compared to an area metric that compensates for over- and under-segmentation.

2 Materials and methods

The next section overviews the general approaches to searching for segmentation algorithm parameters. The pixel correspondence metric (PCM) (Prieto & Allen, 2003) and an area metric (LSB) is compared. Section 2.2 describes the data used and the application-specific characteristics involved. The experimental methodologies for evaluating an edge metric as an adequate fitness function for segmentation algorithm parameter tuning are explained in Section 2.3.

2.1 Semi-automated image segmentation

2.1.1 Segmentation algorithm parameter tuning

The search methodology to find satisfactory segmentation algorithm parameters using segment metrics is presented in Figure 1.

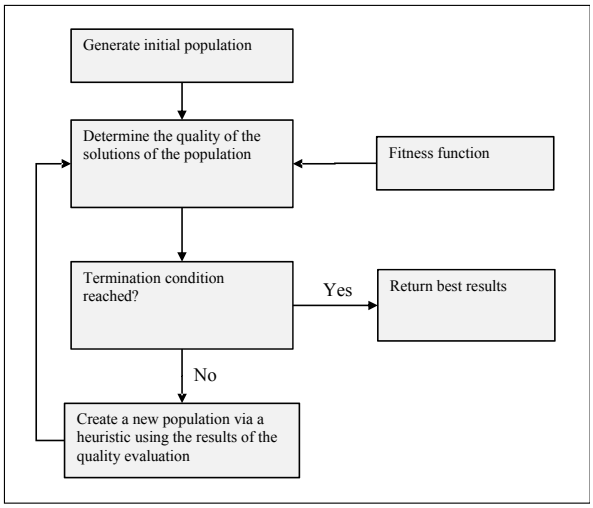

Figure 1. General search methodology for finding segmentation algorithm parameters

Although the methodology can be based on any search function, stochastic population-based search strategies are preferred for computational efficiency. Each agent or member of the search population contains variables or values representing the segmentation algorithm parameters. At each evaluation step of the search algorithm, an image is segmented using the parameters carried by the specific agent. The resulting segments are evaluated against the user-provided set of reference segments via a fitness function (described in Section 2.1.3). In this case the fitness function

comprises empirical discrepancy metrics (area and edge). The algorithm terminates after a predefined number of iterations. The best results of the search procedure are returned as output constituting a set of segmentation algorithm parameters that, when used with the segmentation algorithm (described in Section 2.1.2), most accurately approximates the user-provided reference segmentation set.

2.1.2 Region-merging segmentation

Multiresolution segmentation (MRS), a region-merging segmentation algorithm variant, was used in this work (Baatz & Schäpe, 2000). MRS has seen widespread implementation in earth observation image processing applications (Blaschke, 2010). The algorithm contains several subheuristics that influence the region-merging process. In addition, the behaviour of these components is controlled by user-adjustable parameters. The algorithm contains three parameters, namely:

- A *scale* parameter that is a unitless measure governing the relative segment sizes;
- A *shape-colour* parameter that controls the relative importance of spectral and segment shape properties for region merging;
- A *compactness-smoothness* parameter that controls the relative influence that these two shape properties have on region merging.

In addition to these parameters, the relative importance that input bands have on region-merging is encoded as additional *weight* parameters. Thus, for a three-band image, the algorithm will have six controlling parameters. (See Baatz & Schäpe (2000) for a full formulation of the algorithm.)

2.1.3 Edge and area metrics

The PCM (Prieto & Allen, 2003) evaluates the similarity of two edge images. Compared to other common edge metrics, the PCM incorporates a measure of pixel offset for edges that do not match precisely. The metric also allows for weighted or greyscale edge matching. In this paper a binary version of this metric is implemented.

The PCM matches pixels from the reference image with pixels from the generated image. Pixels that are not matched from either the reference or the generated images are counted as errors. If multiple pixels from an image match with one pixel from the other image, the pixel pairing that minimises the metric is selected. The PCM also takes the spatial offset of matched pixels into account.

The cost of a possible pixel match between a hypothetical reference and a generated pixel is defined as:

$$C\big((i,j),(k,l)\big) = E\big(\max(|k-i|,|l-j|)\big) \qquad [1]$$

where i and j are the row and column numbers of the reference image with (i,j) representing the

corresponding pixel. The same holds for *k* and *l*. The value returned by $\max(|k-i|,|l-j|)$ is the chessboard distance between the two given pixels. *E* is a scaling function in the range [0...1]. In this study pixels within a five-pixel chessboard distance of each other are considered for matching. *E* scales this value accordingly. A cost value of 0 designates an exact match of a pixel pair while a cost value of 1 is a match with a chessboard distance separation of 5 between pixels.

The PCM metric is defined as:

$$PCM(g,r) = 100\left(1 - \frac{C(M_{opt}(g,r))}{|g \cup r|}\right) \quad [2]$$

where *g* is the reference image and *r* the generated reference image. The element $|g \cup r|$ signifies the number of pixels in both images (that are not zero or blank). $C(M_{opt}(g,r))$ denotes the cost of the optimal matching of all pixels in the images *g* and *r*. A PCM value of 100 implies no matching between the imagery and a value of 0 implies a perfect match.

Finding the optimal matching (M_{opt}) that will minimise the PCM constitutes a bipartite graph matching problem (Prieto & Allen, 2003). The nodes of a bipartite graph represent pixels, while vertices indicate the cost of a possible match. In bipartite graph matching, a node may only be attached to one vertex. The aim of solving such a problem is to match all nodes using the least expensive vertices. This work uses the Hungarian algorithm (also called the Munkres algorithm; Munkres, 1957) to search for the optimal matching of pixels. Bipartite graph matching is computationally expensive so that less precise but faster matching strategies (approximation algorithms) may be needed for the operational use of such a metric in a search procedure.

The LSB is an area-based empirical discrepancy metric containing mechanisms to compensate for over- and under-segmentation. The LSB metric is written as:

$$F(S,P) = \frac{1}{n}\left[NS + \sum_{SO(P)_i \neq \emptyset} \frac{fp_i + fn_i + b_i}{\#(S_i)}\right], if\ NS < n \quad [3]$$

where *n* symbolises the number of reference segments delineated by the user. $\#(S_i)$ is the area (in pixels) of the *i*th reference segment. $SO(P)_i$ is the set of segments produced by the segmentation algorithm and at least half of its pixels intersect S_i. *P* is the parameter set used in the segmentation algorithm. Further, fp_i is the number of pixels in $SO(P)_i$ that do not belong to the *i*th reference segment, thus called false positives. And fn_i is the number of pixels in the *i*th reference segment that do not belong to $SO(P)_i$, thus called false negatives. The b_i variable is the number of border pixels in $SO(P)_i$ that intersects the *i*th reference segment area. *NS* is the number of empty $SO(P)_i$ (empty generated segments). Figure 2 illustrates some of the terminology of the LSB metric.

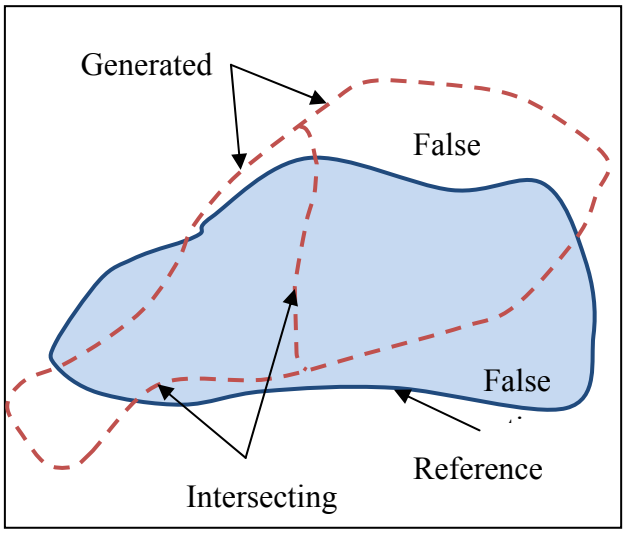

Figure 2. Larger segments booster (LSB) metric (after Fredrich & Feitosa, 2008)

This metric favours generated segments with a tight fit to the reference segment. The parameter *b* provides an offset to this notion by penalising a parameter set for over-segmentation by counting the pixels in the intersecting borders of the reference segment. A prime advantage of the LSB metric is that it does not assume that a single generated segment will be paired with a user created reference segment. Feitosa et al. (2009) offer a quantitative comparison of the abilities of the LSB and reference-bounded segments booster metrics for segmenting urban land-cover features.

2.2 Study area and data

The study area (Figure 3) consists of a single SPOT 5 scene (www.spotimage.com) (identifier 5 119-418 06/11/07 08:49:26 2 J) of an area east of Cape Town. The scene was captured on 7 November 2006 and covers an area of roughly 60 x 60km. The scene location was selected to cover a large portion of the Cape thicket habitat as described by Campbell (1985).

The SPOT 5 scene was geometrically corrected (orthorectified) using PCI Geomatica 10.2 software, while radiometric, atmospheric and topographic effects were corrected using ATCOR3. ATCOR3, as implemented in PCI Geomatica 10.2, was chosen due to the mountainous nature of the study area (Richter & Schläpfer, 2011). A 5m-resolution digital elevation model (DEM), interpolated from 1:10 000 contour and spot height data, was used as input to ATCOR3 and the orthorectification. Statistical image fusion of the panchromatic and multispectral bands was performed using PANSHARP software.

2.3 Methodology

The PCM edge metric is tested as a fitness function in an evolutionary search method for its ability to effectively guide the search procedure as reference for applicable problems. The test is conducted by providing the segmentation algorithm parameter search heuristic with a reference generated by the segmentation algorithm itself to ensure that the reference can be accurately replicated by the algorithm (Feitosa et al., 2006), as opposed to a user-delineated reference.

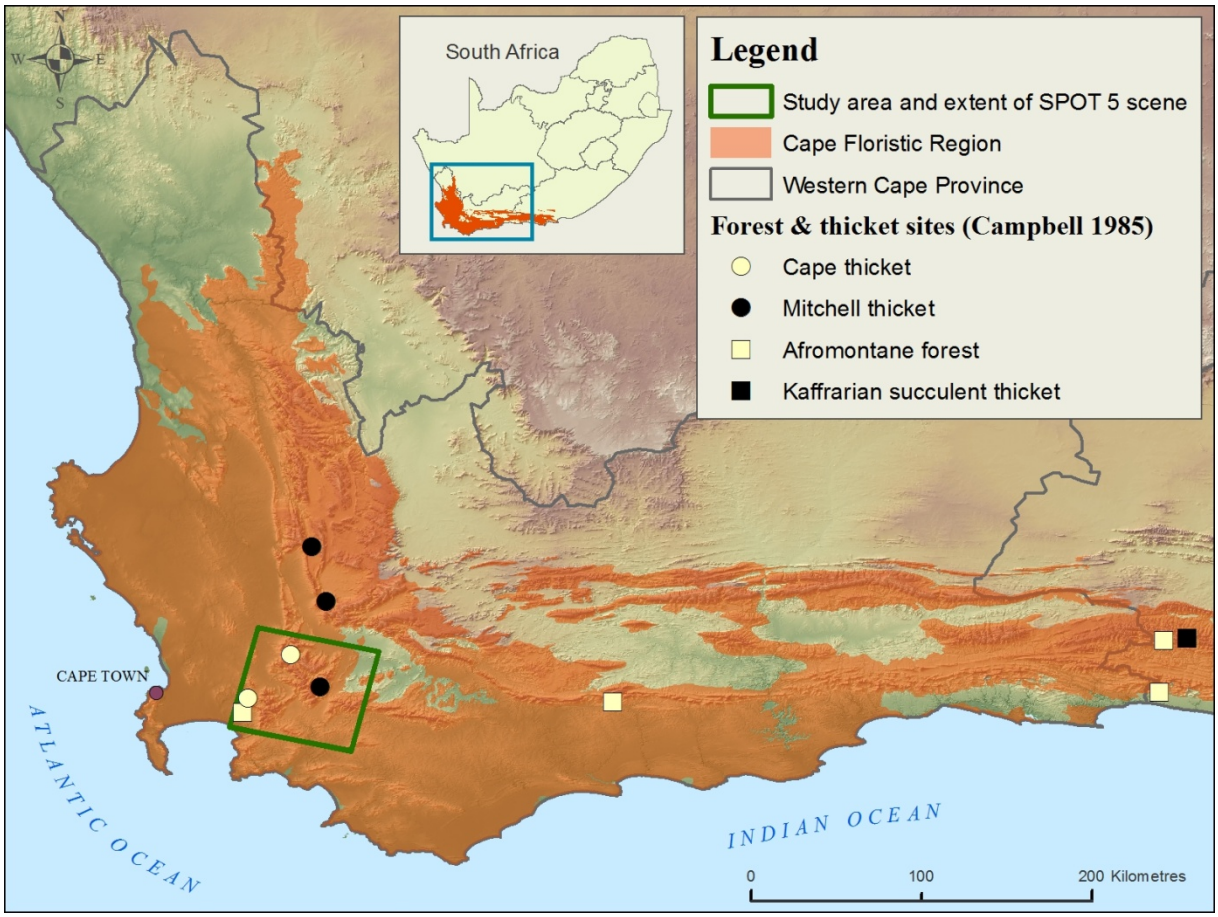

Figure 3. Study area showing the extent of the selected SPOT 5 scene

Two search heuristics, namely self-adapting differential evolution (DE; Brest et al., 2006) and particle swarm optimisation (PSO) (Kennedy & Eberhart, 1995) were implemented in this experiment. Both heuristics were given 60 agents and 15 generations to explore the parameter space so totalling 900 segmentation and evaluation runs. The metaparameters of DE (Brest et al., 2006) and PSO (Vesterstroem & Thomsen, 2004) were set to default values that the literature recommends for unseen problems.

Two regions are investigated, namely an area with agricultural-viticultural fields displaying distinct edges on most features, and a mountain slope with Cape thicket patches displaying less distinct edges on arbitrary land-cover elements of interest. Figure 4 illustrates the regions selected for this experiment.

Three different image subsets (with 300 x 300, 200 x 200 and 100 x 100 pixel dimensions) of these areas were used to investigate the effect of reference sample size on the performance of the search heuristic. The subsets were named fields_100, fields_200, fields_300, natural_100, natural_200 and natural_300 to correspond with the physical area under observation and the size of the image. The suffixes DE and PSO were added to the subset names to denote the type of search heuristic used. A lossy image-splitting function is used to reduce the computational load of the bipartite graph matching on the larger imagery.

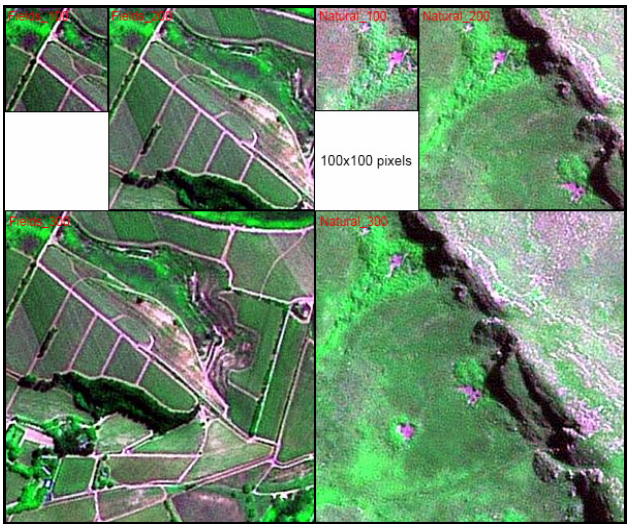

Figure 4. Fields and natural areas with differently sized subsets used in the PCM suitability test

Conducting graph matching with a theoretical upper limit of 90 000 (300 x 300 pixel images) vertices at each iteration of the search algorithm are computationally expensive and would probably not be done in practice. Agents were consequently reduced to thirty for the 300 x 300 imagery to reduce computational loads. The fitness trace of the fittest agent was recorded as the search proceeded in each experiment. Each experiment was conducted five times to get a measure-average value.

Algorithm-generated segments, instead of user-delineated features, were specified as reference. The segmentation algorithm's (Baatz & Shape, 2000) arbitrarily chosen parameters were scale (30), shape-colour (0.5), compactness-smoothness (0.5) and the original SPOT 5 image bands 1 to 3 with a weight of 0.7 for all three bands. The segments generated by this parameter set were taken as the optima for a hypothetical problem, irrespective of land-cover elements present in the images. The search bounds for the scale parameter were set to [20...50] and for the other five parameters to [0.2...0.9]. In total, 60 parameter search experiments were executed (10 experiments for each scene, using two different search algorithms).

The PCM and LSB metrics are also compared quantitatively based on classifier accuracies for the particular problem of generating segments over non-homogeneous features with weak edges displaying some variation in feature size. Cape and Riparian Thicket patches of varying geometric characteristics (region of interest sizes varying from 50 x 50 to 150 x 150 pixels) are selected as the objective of a one-class classification problem. A small region directly east of Stellenbosch, South Africa (3171 x 2265 pixel dimensions) is used for input data.

An automatic classifier parameter tuning and attribute subset selection tool (Fourie, Van Niekerk & Mucina, 2011), operating with both fixed and variable attribute sets, is used to obtain an indication of maximum achievable classifier accuracy, using the PCM and LSB generated parameter

sets/segments. A particle swarm search heuristic using ten-fold cross-validation is employed to tune a one-class SVM. Seven segment attributes judged to hold discriminative value for this arbitrary problem were selected for use in this experiment. The classification experiment was conducted 25 times. The PCM and LSB metrics are compared indirectly according to the classification performance of their resulting segment sets.

3 Results and discussion

The fitness values obtained in the 60 experiments are summarised in Table 1 (segmentation algorithm parameters are also provided). The values are the average of five experiments and the standard deviation over the five experiments are prefixed by '±'. The reference values of the segmentation algorithm parameters are listed below the fitness values. The row of emboldened values indicates the average parameter sets with the closest match to the reference. A fitness value of 0 corresponds to the reference.

Table 1. Mean fitness values and segmentation algorithm parameters by obtained running the segmentation algorithm search heuristic over the six experimental images

Experiment	Fitness	Scale	Shape-Colour	Compactness-smoothness	Band 1 weight	Band 2 weight	Band 3 weight
Fields_100_DE	14.85±1.96	25.64±2.51	0.57±0.15	0.51±0.14	0.62±0.24	0.48±0.21	0.56±0.16
Fields_100_PSO	16.18±1.39	21.17±1.73	0.38±0.23	0.35±0.12	0.57±0.22	0.63±0.26	0.70±0.25
Fields_200_DE	**15.12±5.90**	**29.97±2.37**	**0.56±0.07**	**0.50±0.08**	**0.77±0.15**	**0.71±0.20**	**0.76±0.17**
Fields_200_PSO	17.67±5.19	28.18±3.61	0.58±0.25	0.49±0.17	0.73±0.21	0.67±0.29	0.74±0.20
Fields_300_DE	17.11±6.63	32.24±2.42	0.68±0.16	0.65±0.07	0.70±0.12	0.59±0.20	0.65±0.14
Fields_300_PSO	19.55±1.99	27.74±4.21	0.53±0.25	0.59±0.14	0.55±0.30	0.63±0.29	0.59±0.28
Natural_100_DE	29.62±2.34	29.38±1.78	0.56±0.16	0.50±0.06	0.73±0.13	0.67±0.21	0.70±0.11
Natural_100_PSO	30.54±4.53	24.47±5.36	0.38±0.23	0.48±0.10	0.44±0.15	0.73±0.15	0.54±0.21
Natural_200_DE	30.70±1.91	29.06±2.20	0.62±0.19	0.58±0.07	0.62±0.13	0.61±0.15	0.55±0.03
Natural_200_PSO	36.46±3.43	26.16±3.42	0.58±0.22	0.55±0.13	0.65±0.27	0.63±0.18	0.65±0.23
Natural_300_DE	44.59±3.03	28.24±6.28	0.49±0.19	0.56±0.08	0.52±0.13	0.61±0.12	0.47±0.09
Natural_300_PSO	41.92±1.42	26.62±3.76	0.68±0.20	0.59±0.17	0.54±0.31	0.73±0.23	0.60±0.27
Reference	0.0	30	0.5	0.5	0.7	0.7	0.7

Figure 5 illustrates the average fitness traces of the fittest agent at each iteration of all the experiments conducted on the fields images. The dotted line in the graph signifies the fitness trace corresponding to the lowest fitness value obtained in all 60 experiments. Figure 6 shows the corresponding fitness traces for the six natural image segmentation experiments.

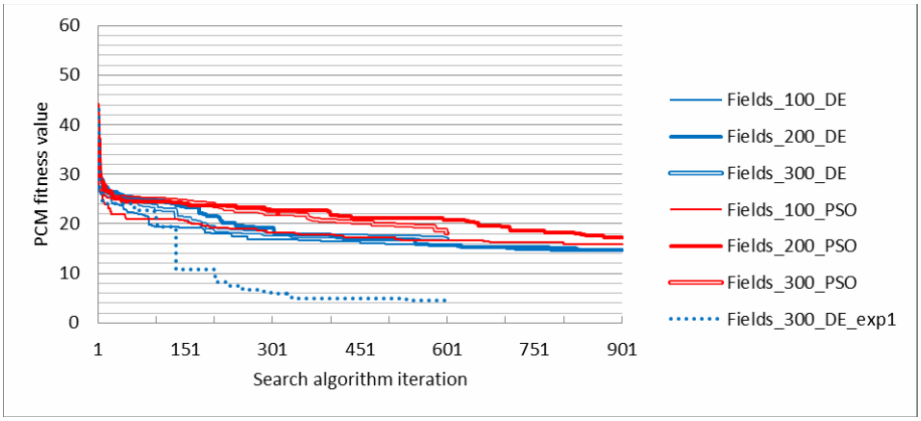

Figure 5. Average fitness traces of the fittest agents of PSO and DE using the three fields experimental images

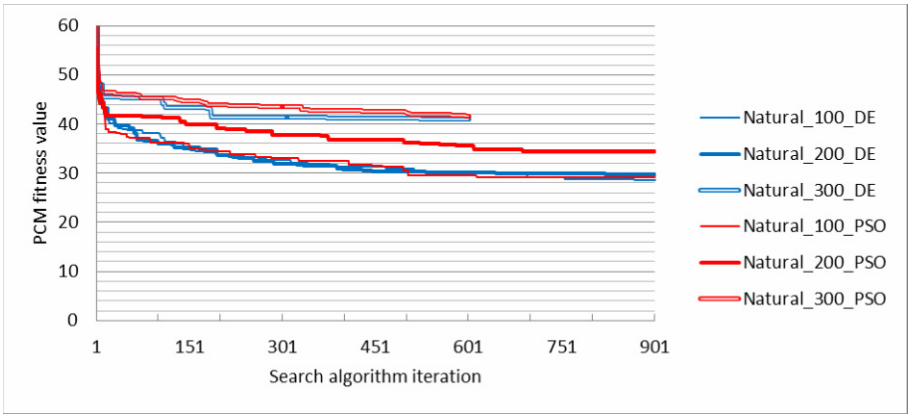

Figure 6. Average fitness traces of the fittest agents of PSO and DE using the three natural experimental images

Figure 7 gives an example of the segmentation results of the experiments and illustrates the Fields_300 image layered with the reference segments and segments generated with a parameter set that obtained a fitness of 16.2 with a scale parameter of 33.6 using DE as the search heuristic. The reference is represented by the green segments and the segmentation results by the red segments. The image (a) on the left is overlaid with the reference (green segments) while the image on the right is overlaid with the generated segments (red segments). Similarly, Figure 8 illustrates the Natural_300 image overlaid with the reference (green lines) and segments generated with a parameter set that obtained a fitness of 43.1 with a scale parameter of 29.8 for the PSO heuristic. The blue areas in Figure 8 are Cape thicket found within the reference and generated segments. The yellow areas are Cape thicket found in either the reference or the generated segments.

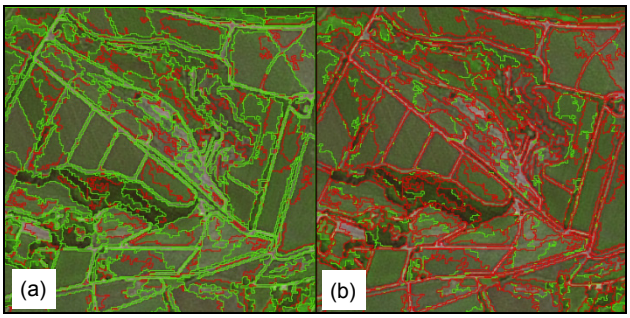

Figure 7. The Fields_300 image segmented with a parameter set that obtained a fitness of 16 with the PCM corresponding to the reference parameters. In (a) the reference segments are overlaid by the generated segments while in (b) the generated segments are overlaid with the reference

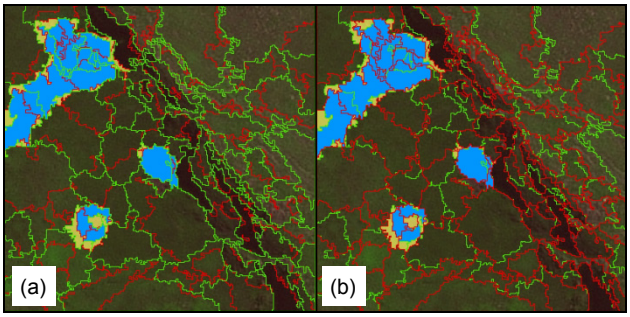

Figure 8. The Natural_300 image segmented with a parameter set that obtained a fitness of 16 with the PCM corresponding to the reference parameters. In (a) the reference segments are overlaid by the generated segments while in (b) the generated segments are overlaid with the reference

The two images subjected to the experiments display very different land-cover element characteristics. The fields image has more distinct edges while the only detectable edges in the natural image were those of the cliff face shadows and to a lesser extent the boundaries of the Cape thicket patches. As observed in Table 1, the fields imagery experiments obtained substantially lower (better) fitness values over all three image sizes compared to the natural images. As shown in Figure 5, the search heuristics were able to converge to a parameter set closely resembling the reference in the fields images. The PCM with the given search heuristics was not able to replicate segment lines cutting through fields. Trial-and-error experimentation with the fields images revealed that extremely small changes in parameter values (fourth-order decimals) resulted in vastly different segments in land-cover elements with weak or no edges.

The phenomenon of the search heuristic being unable to match edges in areas with no or weak edges is prominent in the natural images, resulting in lower PCM scores. The Baatz & Schäpe (2000) region-merging algorithm can generate a vastly different segment set with relatively similar parameter sets in images with weakly defined edges (see Figure 8). Despite this disparity between reference and generated segments, the few land-cover features with identifiable edges (thicket and shadows) allowed the search algorithm (using the PCM) to converge to a scale parameter closely

resembling the reference for these two types of features. Subjectively, the average achieved fitness value of 15-20 for the fields images and a fitness of 40 for the natural images correspond to good parameter sets which accurately delineate all features with identifiable boundaries. In Figure 8, the blue areas are Cape thicket found within the reference and generated segments. The yellow areas are Cape thicket found in either the reference or the generated segments.

The PCM was able to converge closely to the reference scale parameter (30) in most of the test images except for fields_100 and to some extent, natural_100. This is attributable to a small training area, with a larger range of different parameter sets able to replicate the reference. For example, the fields_100_pso image obtained a fitness of 16 and scale parameter of 21 over five consecutive experiments with very little standard deviation.

Note that although the critical scale parameter was consistently approximated, the remaining five parameters displayed less consistency in approximated values. The fields_200_DE image attained a parameter set closely resembling the reference (Table 1). All the other fields images attained similarly low fitness values and a scale parameter close to 30, but displayed greater variation in values of the other five parameters. This is probably attributable to the weak edge phenomenon where an extremely small numerical margin is needed to accurately replicate the reference. In most cases, the DE and PSO heuristics were unable to converge to the absolute reference in the allocated 900 algorithm iterations, generating parameters (excluding the scale parameter) with relatively high standard deviations.

One can deduce from Table 1 and Figures 7 and 8 that the DE search heuristic slightly outperforms the PSO search heuristic concerning convergence, accuracy and convergence speed. Over all six experimental images, the DE search heuristic converged to a scale parameter closer to the reference than the scale parameters generated by the PSO heuristic. The DE heuristic attained a lower mean fitness value compared to the PSO heuristic in all but one experiment (fields_300). On three occasions the DE heuristic achieved a considerably lower fitness value (as indicated by a high standard deviation in Table 1). Through the course of the experiments, the DE heuristic attained a fitness of 7 and 8 in the Fields_200 images and a fitness of 4 in the Fields_300 image. The fitness trace with a resulting fitness value of 4 is illustrated in Figure 5 (dotted line). The lowest fitness score achieved by the PSO heuristic was 10. Figures 7 and 8 suggest that for this domain-specific problem, despite both algorithms performing adequately, the DE search heuristic is preferred to the PSO search heuristic. Note that by performing metaparameter tuning (to the DE and PSO search heuristics) or using more recently proposed meta-heuristics, improved accuracy can be achieved.

As the size of the images increases, so do the average fitness values: a direct relationship attributable to more weak edges cutting fields or fynbos that could not be replicated. Prieto & Allen (2003) advance that this phenomenon relates to the robustness of a metric and they note that some other edge metrics do not display this behaviour. Bear in mind that the remarks by these authors are

not related to convergence experiments but that they are subjective observations suggesting that the PCM is a more accurate edge metric than other edge metrics.

Interestingly, Melo et al. (2008) observed, when using the LSB metric and the Baatz & Schäpe (2000) region-merging segmentation algorithm, that near-optimal fitness values are normally achieved after 400 iterations (10 agents) of the search heuristic. A similar number of iterations is necessary to achieve a near-optimal fitness when using the PCM (30-60 agents), as illustrated in Figures 7 and 8. This implies that 400 iterations is a safe lower-minimum default value in the practical use of such a system.

Table 2 lists the achieved accuracies of classifying Cape and Riparian Thicket using the LSB and PCM generated segments. The mean segment values of band 2, band 4 and band3/band4, the standard deviation of band3/band4, the GLCM homogeneity of NDVI and the GLCM contrast of NDVI and intensity attributes were presented to the classifier. The values indicated are the mean of 25 runs with the standard deviation signified by '±'. The attribute columns denote the number of times the specific attribute was selected by the classifier over the 25 runs. The last column on the right in Table 4 lists the average number of attributes selected by the classifier when operating with variable attribute sets. The differences between "LSB fixed" and "PCM fixed" and between "LSB variable" and "PCM variable" are statistically significant (paired t-test with <0.0005 confidence). The boldface values indicate the metric/segments and attribute selection strategy delivering the highest classification accuracy.

Table 2. Classification results of Cape thicket using both PCM and LSB generated segments with variable and fixed attribute sets

	Gamma	Nu	Accuracy	M B2	M B4	M B3/B4	S B3/B4	H NDVI	C NDVI	C Int	# Attr
LSB fixed	5.08 ±3.78	0.17 ±0.02	78.99 ±0.0027	25	25	25	25	25	25	25	7
LSB variable	15.88 ±17.01	0.17 ±0.03	79.58 ±0.0050	23	21	25	1	19	16	12	4.68 ±0.80
PCM fixed	3.29 ±1.59	0.27 ±0.11	80.34 ±0.0042	25	25	25	25	25	25	25	7
PCM variable	5.11 ±10.71	0.31 ±0.08	**83.06 ±0.0053**	22	13	0	3	24	19	11	3.68 ±1.07

Note: M = mean; S = standard deviation; H = GLCM homogeneity; C = GLCM contrast.

The PCM generated segments resulted in statistically significant higher classification accuracies than the LSB generated segments with a very low standard deviation. A difference of three per cent in geometric means accuracy separates the classifiers using the PCM and LSB segments with variable attribute subsets. Use of variable attribute subsets (Fourie, Van Niekerk & Mucina, 2011) also resulted in higher classifier accuracies compared to using all seven given attributes.

4 Conclusion

In this work an edge metric, called the pixel correspondence metric (PCM), was evaluated as a fitness function for segmentation algorithm free parameter tuning via optimisation. By using the PCM, the common PSO and DE search heuristics were able to converge to reference parameters in typical earth observation imagery when adequate training data were available. The search functions displayed some difficulty in detecting segment lines on weak edges.

The PCM was compared with an area-based metric (the LSB) for segment algorithm parameter generation on heterogeneous features with non-discrete boundaries. The PCM delivered segmentation algorithm parameters that corresponded closer to the user input. Moreover, for an arbitrary problem, a subsequent classification with the PCM-generated segments delivered superior classifier results compared to the LSB-generated segments.

These results show a need for further investigation into the utility of edge-based metrics in methods that attempt to automatically tune segmentation algorithm parameters. It is recommended that hybrid approaches to segment evaluation, such as presented by Weidner (2008) or Derivaux et al. (2010) be examined for balancing notions of edge preservation and accuracy of overlap areas.

5 References

Addink EA, Van Coillie FMB & De Jong SM 2012, 'Introduction to the GEOBIA 2010 special issue: From pixels to geographic objects in remote sensing image analysis', *International Journal of Applied Earth Observation and Geoinformation*, vol. 15, pp. 1-6.

Baatz M, Hofmann P & Willhauck G 2008, 'Progressing from object-based to object-oriented image analysis', in Blaschke T, Lang S & Hay GJ (eds.), *Object-based image analysis: Spatial concepts for knowledge-driven remote sensing applications*, Springer, Berlin.

Baatz M & Schäpe A 2000. 'Multiresolution segmentation: An optimization approach for high quality multi-scale image segmentation', *Angewandte Geographische Informationsverarbeitung*, vol. 12, pp. 12-23.

Bhanu B, Lee S & Ming J 1995, 'Adaptive image segmentation using a genetic algorithm', *IEEE Transactions Systems, Man, and Cybernetics*, vol. 25, pp. 1543-1567.

Blaschke T 2010, 'Object-based image analysis for remote sensing', *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 65, pp. 2-16.

Brest J, Greiner S, Boskovic B, Mernik M & Zumer V 2006, 'Self-adapting control parameters in differential evolution: A comparative study on numerical benchmark problems', *IEEE Transactions on Evolutionary Computation*, vol. 10, pp. 646-657.

Campbell BM 1985, 'A classification of the mountain vegetation of the Fynbos Biome', *Memoirs of the Botanical Survey of South Africa*, vol. 50, pp. 1-121.

Derivaux, S, Forestier, G, Wemmert, C, & Lefèvre, S, 2010. 'Supervised image segmentation using watershed transform, fuzzy classification and evolutionary computation', *Pattern Recognition Letters*, vol. 31, no. 15, pp. 2364-2374.

Feitosa RQ, Costa G, Fredrich CMB, Camargo FF & de Almeida CM 2009, 'Uma avaliacão de metodos

geneticos para ajuste de parametros de segmentacão', Paper presented at the Brazillian symposium of remote sensing held 25-30 April 2009, Natal, Brazil.

Feitosa RQ, Costa GAOP, Cazes TB & Feijo B 2006, 'A genetic approach for the automatic adaptation of segmentation parameters', paper presented at the 1st international conference on object-based image analysis held 4-5 July 2006, Salzburg, Austria.

Feitosa RQ, Ferreira RS, Almeida CM, Camargo FF & Costa G 2010, 'Similarity metrics for genetic adaptation of segmentation parameters', paper presented at the Geographic Object-Based Image Analysis (GEOBIA) 2010 conference held 29 June to 2 July 2010, Ghent, Belgium.

Fourie CE, Van Niekerk A & Mucina L 2011, 'Optimising a one-class SVM for geographic object-based novelty detection', in *Proceedings of the first AfricaGeo conference*, Cape Town, 31 May to 2 June 2011.

Fredrich CMB & Feitosa RQ 2008, 'Automatic adaptation of segmentation parameters applied to inhomogeneous objects detection', paper presented at the Conference on geographic object-based image analysis held 6-7 August 2008, Calgary, Canada.

Hay GJ, Castilla G, Wulder MA & Ruiz JR 2005, 'An automated object-based approach for the multiscale image segmentation of forest scenes', *International Journal of Applied Earth Observation and Geoinformation*, vol. 7, pp. 339-359.

Kennedy J & Eberhart RC 1995, 'Particle swarm optimization', paper presented at the IEEE international conference on neural networks held 26-28 June 1995, Cambridge, United Kingdom.

Lang S 2008, 'Object-based image analysis for remote sensing applications: Modeling reality – dealing with complexity', in Blaschke T, Lang S & Hay GJ (eds.) *Object-based image analysis: Spatial concepts for knowledge-driven remote sensing applications*, Springer, Berlin.

Lübker T & Schaab G 2009, 'Optimization of parameter settings for multilevel image segmentation in GEOBIA', paper presented at the High resolution earth imaging for geospatial information workshop held 2-5 June 2009, Hanover, Germany.

Melo LM, Costa GAOP, Feitosa RQ & Abs da Cruz AV 2008, 'Quantum-inspired evolutionary algorithm and differential evolution used in the adaptation of segmentation parameters', paper presented at the Conference on geographic object-based image analysis held 6-7 August 2008, Calgary, Canada.

Munkres J 1957, 'Algorithms for the assignment and transportation problems', *Journal of the Society for* Industrial and Applied Mathematics, vol. 5, pp. 32-38.

Navulur K 2006, *Multispectral Image Analysis Using the Object-Oriented Paradigm*, CRC Press, New York.

Osman J, Inglada J & Christophe E 2009, 'Interactive object segmentation in high resolution satellite images', paper presented at the IEEE international geoscience & remote sensing symposium held 12-17 July 2009, Cape Town, South Africa.

Pignalberi G, Cucchiara R, Cinque L & Levialdi S 2003, 'Tuning range image segmentation by genetic algorithm', *EURASIP Journal on Applied Signal Processing*, vol. 8, pp. 780-790.

Polak M, Zhang H & Pi M 2009, 'An evaluation metric for image segmentation of multiple objects', *Image and Vision Computing*, vol. 27, pp. 1223-1227.

Prieto MS & Allen AR 2003, 'A similarity metric for edge images', *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 25, pp. 1265-1273.

Richter and Schläpfer 2011. Atmospheric / topographic correction for satellite imagery. ReSe Applications

Schläpfer, Wil.

Vesterstroem J & Thomsen R 2004. A comparative study of differential evolution, particle swarm optimization, and evolutionary algorithms on numerical benchmark problems. Paper presented at the IEEE conference on evolutionary computation held 20-23 June 2004. Portland, Oregon.

Weidner U 2008, 'Contribution to the assessment of segmentation quality for remote sensing applications', paper presented at the International Society for Photogrammetry and Remote Sensing annual conference: Silk road for information from imagery held 3-11 July 2008, Beijing, China.

Zhang Y & Maxwell T 2006, 'A fuzzy logic approach to supervised segmentation for object-oriented classification', paper presented at the ASPRS 2006 annual conference held 1-5 May 2006, Reno, Nevada.