INT. J. REMOTE SENSING, 20 NOVEMBER, 2003,
VOL. 24, NO. 22, 4251–4266

Taylor & Francis
Taylor & Francis Group

Information extraction from very high resolution satellite imagery over Lukole refugee camp, Tanzania

S. GIADA, T. DE GROEVE, D. EHRLICH

Institute for the Protection and Security of the Citizen, Joint Research Centre of the European Commission, TP 267, 21020 Ispra (VA), Italy; ?e-mail: silvia.giada@jrc.it, tom.de-groeve@jrc.it, daniele.ehrlich@jrc.it

and P. SOILLE

Institute for the Environment and Sustainability, Joint Research Centre of the European Commission, TP 262, 21020 Ispra (VA), Italy; ?e-mail: pierre.soille@jrc.it

(*Received 21 February 2002; in final form 15 August 2002*)

Abstract. This paper addresses information extraction from IKONOS imagery over the Lukole refugee camp in Tanzania. More specific, it describes automatic image analysis procedures for a rapid and reliable identification of refugee tents as well as their spatial extent. From the identified tents, the number of refugees can be derived and a map of the camp can be generated, which can be used for improving refugee camp management. Four information extraction methods have been tested and compared: supervised classification, unsupervised classification, multi-resolution segmentation and mathematical morphology analysis. The latter two procedures based on object-oriented classifiers perform best with a spatial accuracy above 85% and a statistical accuracy above 97%. These methods could be used for refugee camp information extraction in other geographical settings and on imagery with different spatial and spectral resolutions.

1. Introduction

The United Nations High Commissioner for Refugees (UNHCR) has reported 12 million refugees for 2001 (UNHCR 2001). An almost equivalent number of internally displaced people share the plight of refugees. Refugees usually gather in camps that provide absolutely essential facilities for survival. Yet, life conditions remain extremely difficult for the high concentration of people and the relatively simple infrastructure. Refugee camps are made of tents that lack the amenities, security, and stability of dwellings. Often refugees enter into competition with the local population for the use of local resources, which they inevitably deplete. Despite these difficulties, refugee camps last for years and are maintained through the continued support of the donor community.

The Lukole refugee camp addressed in this study was established in 1994 in the north-western Ngara province of Tanzania in the aftermath of the Rwanda crisis. Since then it has hosted different waves of refugees (Anonymous 1997). At the time the satellite image was collected in September 2000, the European Commission

International Journal of Remote Sensing
ISSN 0143-1161 print/ISSN 1366-5901 online © 2003 Taylor & Francis Ltd
http://www.tandf.co.uk/journals
DOI: 10.1080/0143116021000035021

Humanitarian Aid Office (ECHO) of the European Union, which financially supports the camp, reported over 130 000 refugees living in the camp (personal communication). A significant number of similar refugee camps worldwide are set up and maintained by donor institutions. Statistics on refugees are required by donors to estimate and justify the resources that they provide. They are also needed by host governments for security purposes and to anticipate the social and economic impact of refugees. Provision of reliable statistics including the number of refugees in refugee camps is a priority task for aid organizations (Crisp 1999).

Satellite derived information can greatly complement the information that is traditionally collected by field observations (UNHCR 2000, Bjorgo 2000a). In fact, imagery provides a unique overview of the refugee camp. Size, density of tents, and geographical setting visible on the image allow for an intuitive assessment of the magnitude of the phenomenon that only a picture can capture (see also the cover of this issue). Repeated acquisitions are also used to assess changes in the population of camps as well as pressure exerted on the environment by human activities such as fuel gathering (Jacobsen 1997, Lodhi *et al*. 1998, ENVIREF 2001). In addition, image analysis makes it possible to derive other important information (Bjorgo 2000b, Dare and Fraser 2001). Accurate estimates of refugees can be inferred from the number of the tents. The production of maps with every single tent contributes to better camp management especially when the digital map is interfaced with a database. Satellite imagery and the derived geographical information can be disseminated through the internet facilitating a rapid and simultaneous communication of the information to all parties that have at stake the maintenance and running of the camp.

Information extraction from satellite imagery is often carried out by visual inspection. Visual inspection of high resolution imagery is accurate but very time consuming. Computer assisted procedures offer opportunities to provide rapid and robust information. This study discusses four computer assisted procedures to count and locate tents within the camp in order to generate a camp map. The procedures aim to benefit from the newly available satellite imagery at sub-meter resolution now becoming available or aerial photography.

2. Background

Information extraction related to human activities from very high resolution imagery poses new challenges. Very high resolution satellite imagery (with sub-meter resolution) depicts man-made structures such as houses, trucks, and cars (Tanaka and Sugimura 2001). In this study, tents from refugee camps are clearly identifiable as objects containing multiple pixels. Information extraction in this case addresses the identification of objects as opposed to the classification of pixels into area classes, which is typical for environmental applications. Image analysis techniques therefore have to be adapted to the new tasks.

Traditional parametric classifiers are based on per pixel analysis to be used for area mapping. These parametric classifiers are usually based on maximum likelihood criteria or neural networks for supervised procedures and self-organizing clustering algorithms for unsupervised procedures. These algorithms are commonly found in commercially available image processing systems but fail to incorporate image characteristics that are not based on pixel values, which are available in object-oriented methods.

Unlike traditional methods, an object-oriented method for image analysis treats the image as a set of meaningful objects rather than pixels. Objects are defined on

the basis of the internal homogeneity of their spectral values, the contrast with neighbouring objects, their own spatial or spectral characteristics, or a combination of these three properties. While some methods immediately extract objects of interest from the background, other methods do not differentiate between foreground objects and background. Rather, they segment the whole image in objects and use an additional classification step to extract objects of interest. For the classification, more information than just the spectral information of a pixel is available. For example, the object size, shape (elongated or compact), texture (variation of pixel values within the object) and relationship to neighbouring objects are typically used. This semantic information makes complex class definitions possible such as: 'Tents are bright, square objects of about 20 m^2. Background soil consists of dark, textured, large, fractal-like objects.'

Two object-oriented approaches were tested. A first approach is a multi-resolution segmentation (Baatz and Schaepe 2000) followed by a supervised fuzzy logic classification as implemented in the commercial software package eCognition. The second approach is an object extraction algorithm based on mathematical morphology (Soille 1999) and developed by the Joint Research Centre (JRC) of the European Commission, Ispra, Italy. We will briefly describe background notions related to both object-oriented approaches.

2.1. *Multi-resolution segmentation*

In the eCognition software, objects are extracted from the image in a number of hierarchical segmentation levels. Each subsequent level yields image objects of a larger average size by combining objects from a level below. This entails the representation of image information on different scales simultaneously. Objects (or pixels at the first level) are grouped into a larger object based on spectral similarity, contrast with neighbouring objects, and shape characteristics of the resulting object. These three characteristics are grouped into a single image parameter called *heterogeneity*.

Throughout a single segmentation step, the underlying optimization procedure minimizes the heterogeneity of resulting image objects weighted by their size. A segmentation step is finished when every original object (or pixel in the first step) is assigned to the optimal higher level object. To achieve adjacent image objects of similar size and thus of comparable quality, the procedure simulates an even and simultaneous growth of objects over a scene in each step and also for the final result. Thus, the procedure starts at any point in the image with one-pixel objects. A treatment sequence with a stochastic, historical element based on a binary counter guarantees a regular spatial distribution of treated objects.

For the description of spectral or colour heterogeneity, eCognition uses the sum of the standard deviations of spectral values in each layer σ_c weighted with the weights for each layer w_c: $\Sigma w_c.\sigma_c$. This approach enables the simultaneous use of multiple bands and even multiple co-registered images of different resolutions. To avoid branched and fractal objects, the criterion for spectral heterogeneity is mixed with a criterion for spatial heterogeneity reducing the deviation from a compact or smooth shape. Heterogeneity as a deviation from a compact shape is described by the ratio of the border length l and the square root of the area A of the image object: $l/\sqrt{A}$. Smoothness of the boundary is described as the ratio of the border length l and the shortest possible border length b given the size of the image object: l/b. The segmentation result depends on the requested level of heterogeneity h^* as

well as the respective weights of the three criteria, which can be set by the user. Equation (1) shows the optimization function.

$$w_1 \sum_c w_c \sigma_c + w_2 \frac{l}{\sqrt{A}} + w_3 \frac{l}{b} \leqslant h^* \quad (1)$$

Since for a higher level of heterogeneity relatively contrasted regions are merged the size of the resulting objects increases with increasing h^*.

In order to extract objects of interest from the segmented image, eCognition offers a fuzzy logic supervised classification tool where hierarchical classes can be defined using both spectral signatures of samples and a list of 57 object related variables including pixel value, shape, texture, hierarchical, and neighbourhood properties. For each parameter, one can define a class membership function describing the probability of an object with a given parameter value of belonging to a given class. All parameters are combined with fuzzy logic operators, which are an extension of the Boolean 'and' and 'or' logical operators (Zadeh 1965).

The eCognition object-oriented approach has advantages over traditional pixel-based classification approaches. Segmentation drastically reduces the number of units to be handled for classification, reducing the treatment time. Homogeneous objects also provide a significantly increased signal-to-noise ratio compared to single pixels within one class: therefore, a single dark pixel within a bright region will not generate a hole in the object, yielding a result closer to a human interpretation.

2.2. *Mathematical morphology*

A second object-oriented approach is based on a theory for the analysis of spatial structures called *mathematical morphology* (Serra 1982). It is called morphology because it aims at analysing the shape and form of objects. It is mathematical in the sense that the analysis is based on the set theory, integral geometry, and lattice algebra. Mathematical morphology is not only a theory, but also a powerful image analysis technique.

In mathematical morphology, two-dimensional grey tone images are seen as three-dimensional sets by associating each image pixel with an elevation proportional to its intensity level. Another set of known shape and size, called the *structuring element*, is then used to investigate the morphology of the input set. This is achieved by positioning the origin of the structuring element to every possible position of the space and testing, for each position, whether the structuring element either is included or has a non-empty intersection with the studied set. By keeping those points for which the inclusion test is positive, we obtain an *erosion* of the input set while the non-empty intersection test leads to a *dilation* of this set. Erosions (respectively dilations) by flat (i.e. no variation in grey tone) structuring elements come down to calculating, for each pixel, the minimum (respectively maximum) image value within the neighbourhood defined by the structuring element. Erosions and dilations are the letters of the morphological alphabets since all other transformations are based on these two elementary transformations. The shape and size of the structuring element must be selected according to the morphology of the targeted image structures. For instance, line segments are suited to the analysis of thin elongated structures such as road and railway networks while disks are more appropriate for more isotropic structures such as crop fields and lakes.

The erosion of a set not only removes all components that cannot contain the structuring element but it also shrinks all other components. The search for an operator recovering as much as possible the information lost by the erosion leads to the definition of the morphological *opening*. This is achieved by dilating the eroded image. In general, only part of the information is recovered. For example, connected components of pixels completely destroyed by the erosion are not recovered at all. This behaviour is at the very basis of the non-linear filtering properties of the opening operator: foreground image structures are selectively filtered out, the selection depending on the shape and size of the structuring element.

The geometric formulation of the morphological opening of a set in terms of structuring element fits is as follows: '*Does the structuring element fit the set?*' Each time the answer to this question is affirmative, the *whole* structuring element must be kept (for erosion, it is the origin of the structuring element that is kept). For example, the opening of a rectangle by a disk rounds off its corners.

Beyond these basic concepts, a thorough presentation of the principles and applications of morphological image analysis is given by Soille (1999). A more mathematical exposition is proposed by Heijmans (1994) while the key seminal reference book is by Serra (1982). Soille and Pesaresi (2002) studied the application of morphology to remote sensing and geoscience in detail.

3. Data

The satellite image used in this study was acquired on the 24 September 2000 by the Optical Sensor Assembly (OSA) scanner telescope, mounted on board the IKONOS satellite. IKONOS orbits at an altitude of 681 km, at an inclination of 98.1° and it is sun-synchronous. It provides simultaneous imaging, collecting 1 m panchromatic (0.45–0.90 μm) and 4 m multi-spectral (0.45–0.52, 0.52–0.60, 0.63–0.69, 0.76–0.90 μm) bands. The image covers an area of 11 km by 11 km centred on longitude 30°52′55.96″ East and latitude 2°37′03.23″ South. It was geometrically corrected to the UTM Zone 36S (WGS 84) map projection by the data provider.

The PAN and multi-spectral bands were merged to obtain new images with the spatial resolution of PAN while maintaining the spectral resolution of the multi-spectral bands. The merging procedure consisted of computing a principal component analysis (PCA) on the spectral bands. The method remaps the high-resolution image into the data range of the first principal component (PC-1) and substitutes it for PC-1 then applies an inverse principal components transformation. The original bit range and number of bands are maintained and the original spectral values are preserved to a large extent, while each band is sharpened.

4. Methodology

This article focuses on extraction of two types of information from the very high resolution IKONOS image: the number of tents and the location of every individual tent. While for both types of information all tents in the camp play a role, they do not need to be considered individually to estimate the total number of tents. Statistics on tent density (i.e. number of tents per area unit) and the total area of the camp can lead to acceptable estimates of the total number of tents. However, for creating a map of the refugee camp, all tents need to be considered and the location of each measured. Since all tents need to be considered one by one for the latter type of information, methods based on interpretation by a human operator

are practically not feasible with a large number of tents. Computer based methods are the only alternative.

We will describe five different methodologies for extracting information on refugee camps: visual interpretation by a human operator, supervised classification based on a maximum likelihood classifier, unsupervised classification using a self-organizing clustering algorithm, and the two object-oriented methods named earlier in the text: multi-resolution segmentation and mathematical morphology.

4.1. *Visual interpretation*

Because of the huge number of tents in the Lukole refugee camp, manual counting by a human operator – while the most straightforward procedure – is practically not feasible. Therefore, we applied statistical extrapolation techniques to estimate the most probable number of tents based on sample areas.

We performed a visual interpretation on the panchromatic band and the merged image using 18 sample areas of 300×300 pixels (about 9 ha). The sample areas are distributed in a systematic way by locating them on every third row and every third column of a grid with cells of 300×300 pixels (figure 1). Within each sample area, all tents were counted. Tents that are only partially in the sample area were included if their centroid lies in the sample area. While built-up structures were not counted, big tents, which are probably not used as shelter but for common activities, were counted too.

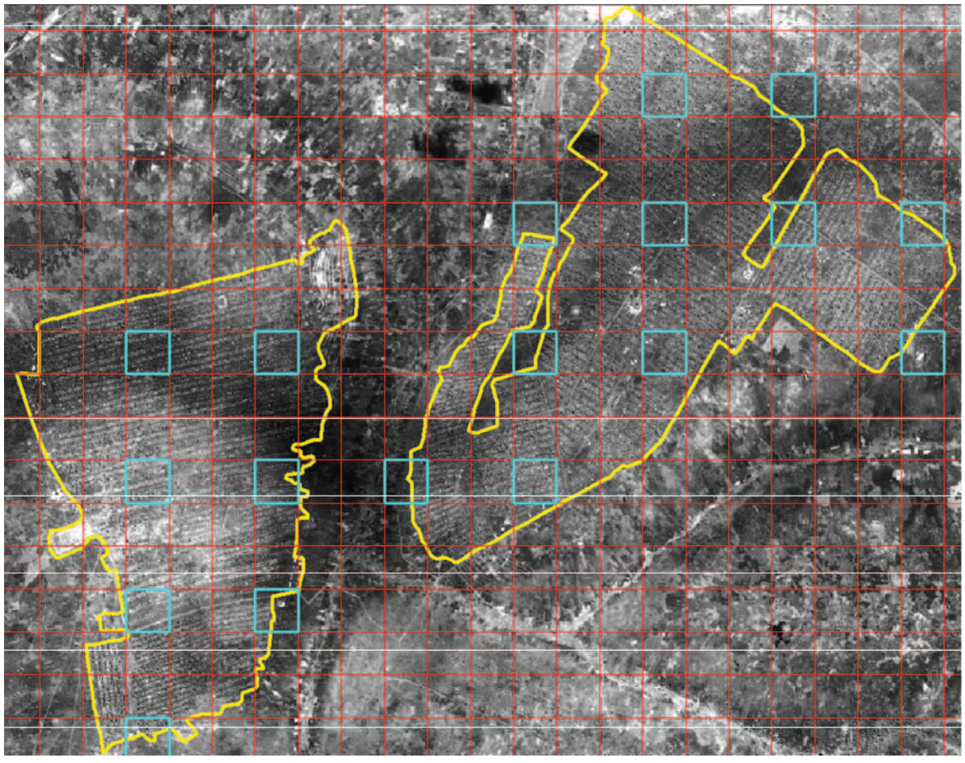


Figure 1. The Lukole refugee camp is outlined in yellow, the grid used for the systematic stratification is visible in red. The light blue squares were used as sample areas. Each red grid cell covers an area of $300 \times 300\,m^2$.

Within each sample area, we calculated the density of tents by dividing the number of tents by the area of the sample area lying within the camp's borders. The average density for the whole camp is then the average of all density values. We obtained the total number of tents in the whole camp by multiplying the average density by the total area of the camp. The statistical error associated with the number of tents in the entire camp can be calculated by multiplying the error of the average tent density by the camp area. Under assumptions of normality of tent distribution and an error interval of 1.96 times the standard deviation the error interval individuates a 95% confidence interval.

For creating a map of the whole camp the location of every single tent is required, which has not been recorded by visual interpretation. However, the location of the tents within the sample areas was stored in a digital file as points. Therefore, the true location of all tents is known within sample areas. This information is used as ground data for accuracy assessment of the computer assisted methods.

4.2. *Supervised and unsupervised classification*

To perform the supervised classification, spectral signatures were collected for five classes from the merged image: tents, trees, roads, bare soil, and buildings. These signatures were then used within a maximum likelihood classifier. The resulting classification was then converted into a vector file where only objects belonging to the 'tents' class were exported. In the resulting vector file, every polygon represents an object classified as a tent.

Unsupervised classification was based on the iterative self-organizing data analysis (ISODATA) clustering algorithm (Ball and Hall 1967). Classifications with 5, 10 and 15 classes were run. The class representing best the tents in the classification was identified, and only objects belonging to this class were exported to a vector file in a similar way as the supervised classification.

Because both the supervised and the unsupervised procedure are pixel-based, their application on very high resolution imagery yields a speckle-like classification result where the background contains a large amount of small objects classified as tents, although they are not. In a post-classification processing step, these small objects were removed by defining an area threshold for a tent. All objects classified as tent and having an area smaller than the area threshold were deleted.

The area threshold was defined by computing omission and commission error using the visually interpreted ground data from the 18 sample sites. Omission and commission errors were calculated based on the number of objects and not in the traditional terms of number of pixels (Congalton *et al*. 1983, Congalton 1991). We made the assumption that the threshold should be set to a size of objects that equals the omission (E_o) and commission (E_c) error. This guarantees that the number of objects classified as tents is the same as the number of tents in the ground data. Although the value of omission and commission can only be calculated for the sample areas, the individuated threshold is applicable for the whole camp.

To determine the size threshold A for which $E_o(A) = E_c(A)$, we studied the area distribution of objects resulting from the classification $C(A)$. First, we determined whether an object is truly a tent by comparing its position with each tent position recorded in the visual interpretation. Figure 2 shows the area distribution (number of objects at given area A) of all objects $C(A)$ and of objects correctly classified as tents $C_c(A)$. The area below the C_c curve contains all tent objects and the area

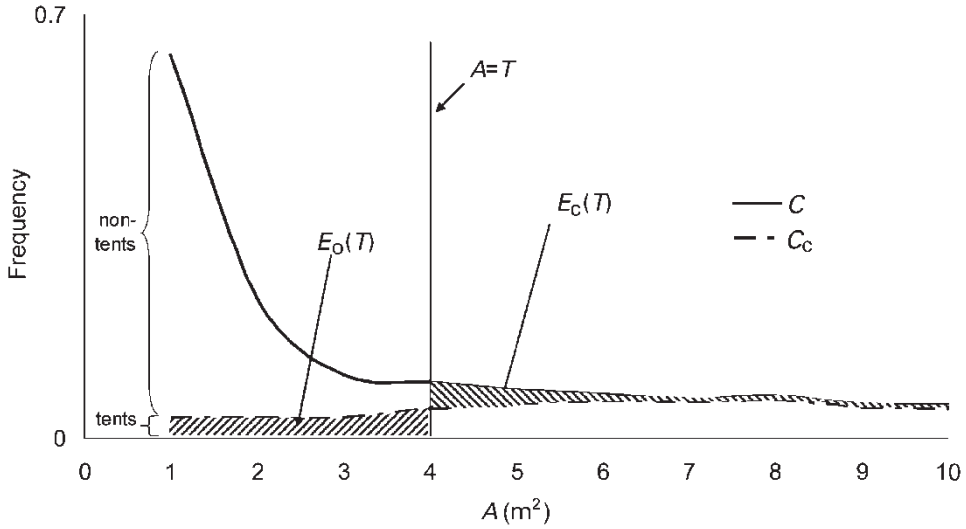


Figure 2. The size distribution of C and C_c. Curves start at 1, the minimum resolution of the imagery. For a given threshold T, $E_o(T)$ and $E_c(T)$ are shown. The example is taken from the sample areas for unsupervised classification with 10 classes.

above all non-tent objects. For a given area threshold T, the commission error is the sum of all non-tent objects with area $A \geqslant T$: $E_c(T) = \sum_{A \geqslant T} (C(A) - C_c(A))$. This is equivalent to the area above the C_c curve and right from the threshold T. The omission error, however, is composed of two parts. It equals the sum of all tent objects (the area below the C_c curve) left from the threshold in addition to the tents that have not been recognized at all by the method and do not have an area associated with them. We only know the latter number from the ground data. We assign this number to the omission error at area 0: $E_o(0)$. The omission at threshold T is shown in equation (2).

$$E_o(T) = E_o(0) + \sum_{A<T} C_c(A) \tag{2}$$

When plotting both $E_o(T)$ and $E_c(T)$, it is possible to locate where the omission error equals the commission error (figure 3). Since the size of tents is discrete (pixels have an area of $1\,m^2$), the threshold value must be a discrete integer number. Therefore, omission and commission will not be exactly equal. The resulting size threshold for both the supervised and the unsupervised procedure is $T = 4\,m^2$. The corresponding tent size gives the value of the threshold to use for the whole camp.

For the unsupervised classification, there is another parameter to set: the number of classes. We select the number of classes in such a way that omission (and commission) errors are at a minimum. In doing so, the maximum amount of objects is located at the proper geographical location. For the classification with 10 classes omission and commission were lowest.

4.3. *Multi-resolution segmentation*

The multi-resolution segmentation approach is done using eCognition. Since the tent boundaries are much better defined in the 1 m resolution panchromatic band than in the 4 m resolution multi-spectral bands, we assigned a higher weight to this band for all multi-resolution segmentation steps. However, for the classification of

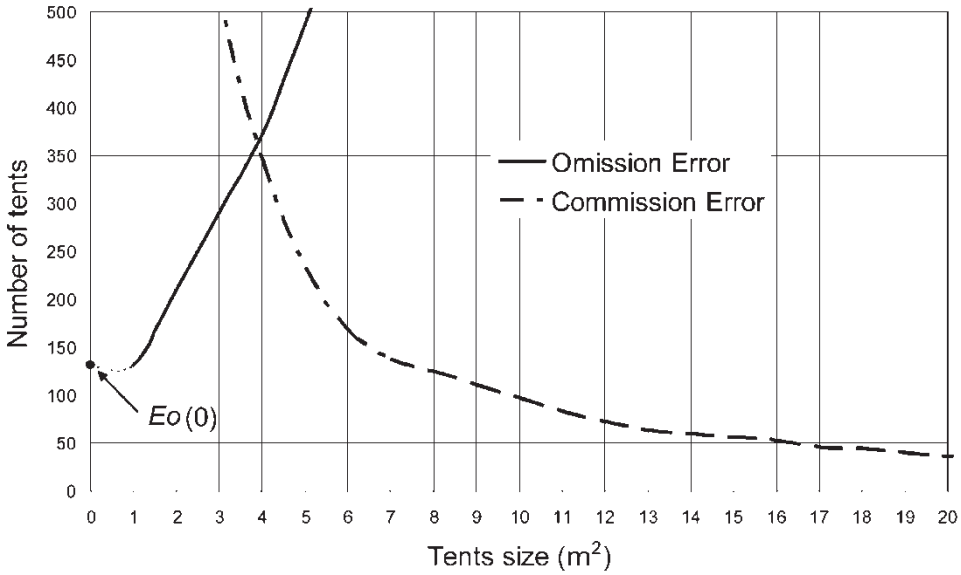


Figure 3. Omission and commission errors are plotted against size. The point where they meet identifies the size threshold to be applied to the whole camp.

segmented objects, we used the four spectral bands together with the panchromatic band. We created two segmentation levels to extract the tent objects from the image. For the first segmentation, we applied the heterogeneity level of 10 (no units), which typically increases the segmented object size from 1 pixel to 5 pixels. Since the white tents are spectrally distinct from the dark background, we emphasized the spectral information in the first level by assigning a weight of 0.8 to the spectral criterion. Further, we favoured compact objects over smooth objects by attributing 90% of the remaining weight to the compactness criterion. For the second level, we applied a heterogeneity level of 20, yielding an average segmented object size of 60 pixels. Since at this level the tent objects are formed from the small objects from level 1, the shape criteria are more important. Therefore, we attributed 50% to the spectral criterion, and 50% to the shape ones; again the compactness is more important than smoothness since tents are compact and square. The resulting weights for the segmentation are summarized in table 1.

Once we extracted the objects, we classified them in order to distinguish tents from non-tents. Therefore, we used the fuzzy logic classification tools of eCognition. These tools allow us to produce two-dimensional scatter plots of all parameters available for classification. Based on supervised sampling and subsequent examination of the scatter plots, fuzzy class boundaries can be set for the most significant parameters. A fuzzy boundary of a parameter is described by a probability function varying between 0 and 1. To describe the fuzziness of a boundary in the following class descriptions, we will give the value with probability 0.5 and between brackets the values of probability 0 and 1, in that order. The final classes were defined as follows (see table 2):

– Tents differ from other objects in the image by their size, compactness, shape index and brightness. They are between 6 m^2 (3.5–8.5) and 60 m^2 (80–40) in

Table 1. Parameters for multi-resolution segmentation of eCognition.

	Pixel level	Level 1	Level 2
Parameters for segmentation			
Heterogeneity		10	20
Spectral criterion w_1		0.8	0.5
Compactness criterion w_2		0.18	0.45
Smoothness criterion w_3		0.02	0.05
MS blue band w_{b1}		0	0
MS green band w_{b2}		0	0.15
MS red band w_{b3}		0	0
MS infrared band w_{b4}		0	0.35
PAN band 1 w_{b5}		1	0.5
Results of segmentation			
Average object size	1	5	60
Number of objects	3.8×10^7	5.7×10^6	4.5×10^5

Table 2. Parameters for fuzzy logic classification of eCognition.

Parameter	Formula	Range	Tents	Buildings	Trees	Terrain
Spectral signature	–	–	From samples	From samples	From samples	From samples
Size	A (m^2)	0–∞	>6 and <60	>75	<300	>100
Shape index (fractal-like)	$l/4\sqrt{A}$	1–∞	Small (<2)			Large (>2)
Compactness	A/r r = object radius	0–1	Large (>0.4)		Small (<0.5)	

size. They equally are very compact, which translates in a compactness larger than about 0.4 (0.35–0.45). Their boundaries are generally smooth yielding a shape index lower than about 2 (2.5–1.5).

– Buildings are larger than tents, typically over about 75 m^2 (60–90). They also have a very bright spectral signature.
– Trees have a similar size to tents and are spectrally bright compared to the background, which leads inevitably to confusion between tents and trees. However, trees typically have a smaller compactness than tents, so smaller than about 0.5 (0.55–0.45), and have a different spectral signature.
– Terrain is the background of the image, which is covered by cloud shadows in some parts of the image. The objects of this class are typically very large (larger than about 100 m^2 (80–120)) and have fractal boundaries (shape index larger than 2 (1.5–2.5)).

4.4. *Mathematical morphology*

If searched image objects are characterized by a clear shape, size, and contrast, it is likely that mathematical morphology will turn out to be successful. In this study, we are looking for objects having a specific shape (tents are more or less rectangular objects), size (tents are between about 6 and 60 m^2), and contrast (in panchromatic images, tents always appear brighter than their surroundings). The

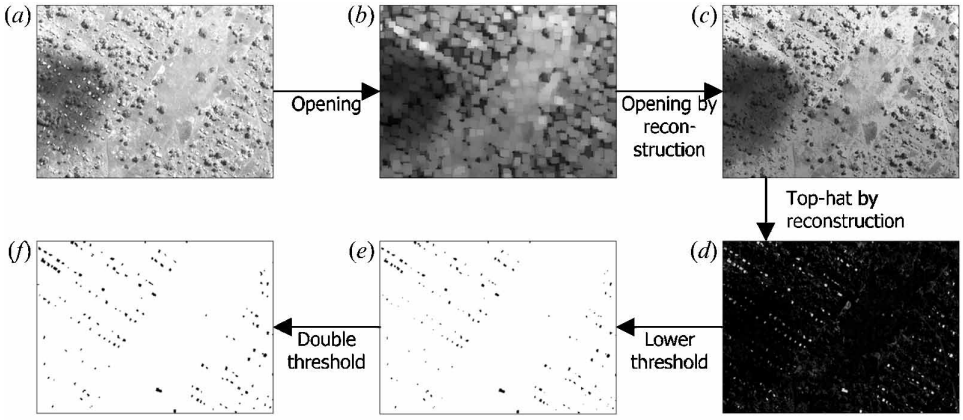


Figure 4. Morphological detection of tents: summary of the methodology on an image sample.

application of the method is a morphological chain of operators. Each step is illustrated on a small sample of a full IKONOS scene in figure 4.

First, we filter the input image using *a priori* knowledge about the size of the searched image objects: we assume that every bright image object that can contain a square structuring element of at least 9×9 is *not* a tent. That is, by performing a morphological opening with a 9×9 square-structuring element, all tents are removed as shown in figure 4(*b*). However, a side effect of the opening by a structuring element is that boundaries of other bright image structures that cannot be covered by the selected structuring element are also filtered out. This is why we clearly recognize the shape of the structuring element in the opened image. A solution to this problem is to reconstruct the image using the opened image as a marker. This operation is known in morphology as an *opening by reconstruction* (Soille 1999: 176–177). By doing so, the tents are not reconstructed because they were *totally* removed by the initial opening. However, all objects partially modified by the opening are reconstructed. The output of the opening by reconstruction is displayed in Figure 4(*c*). Since we have succeeded in removing the tents from the image, we can highlight them by computing the difference between the input image and the opening by reconstruction. This operation is called a *top-hat by reconstruction* (Soille 1999: 177). Figure 4(*d*) shows the output of the top-hat by reconstruction. Note that so far, only the size of the square used for the initial opening needs to be selected. The top-hat by reconstruction provides an image with tents whose contrast is roughly equal to that occurring in the input image but, now, the background of the tents is more or less constant across the image (i.e. very low values, see figure 4(*d*)). This enables the application of a global threshold for the extraction of a mask of the tents. Still, a unique threshold would produce only part of the tent outline (if the threshold is too low). This problem can be solved by the double threshold technique known in morphology (Soille 1999: 166). This procedure consists of thresholding the input image for two ranges of grey scale values, one being included in the other. The threshold for the narrow range is then used as a seed for the reconstruction of the threshold for the wide range. For the processed scene, we have selected 40% for the upper and 25% for the lower intensity values. The lower threshold as well as the output of the reconstruction process are displayed in figure 4(*e*)–(*f*).

Table 3. Statistics from visual interpretation and four computer assisted procedures.

	Sample areas						Camp	
Classification procedure	Objects classified as tents (C)	Objects correctly classified as tents (C_c)	Omission error (E_o)	%	Commission error (E_c)	%	Number of tents	Error (%)
Visual interpretation	2467	n/a	n/a	n/a	n/a	n/a	23 834 ± 4633[1]	n/a
Supervised	2516	2141	326	13.21	375	14.90	27 134	13.85
Unsupervised ISODATA 10	2443	2096	371	15.04	347	14.20	26 332	10.48
Multi-resolution segmentation	2396	2090	377	15.28	306	12.77	23 994	0.67
Mathematical morphology	2414	2155	312	12.65	259	10.73	24 387	2.32

[1]95% confidence interval error margin.

5. Results and discussion

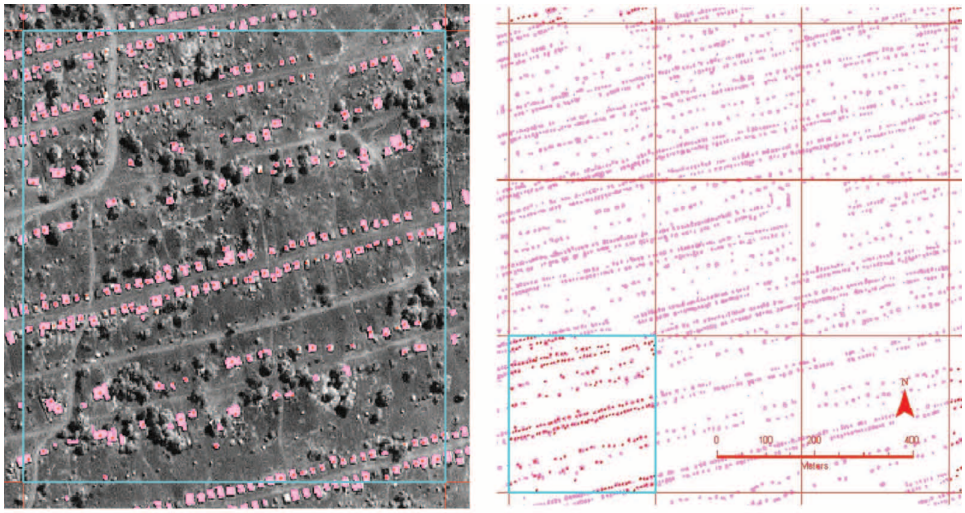
Results of the visual interpretation and of the four automated classification procedures are summarized in table 3.

In the visual interpretation, a total of 2467 tents were identified in the 18 sample areas totalling 109.8 ha. Extrapolation of the number of tents to the whole camp area (1119.4 ha) yields a total of 23 834 tents. The estimate of the average tent density has a 95% confidence interval of 4.1. From this interval, we derived that the true number of tents lies with 95% confidence between 19 201 and 28 467 (i.e. 23 834 ± 4633). This number is the point of reference for the automated methods since it is based only on human counting and statistical methods.

The number of refugees is then derived from the occupancy statistics. From ECHO (personal communication), we obtained estimates of occupancy rates between 5 and 6 people per tent, depending on coming and going waves of refugees. Depending on the method, these figures suggest a total number of refugees between 119 970 and 162 804. From the same source, we know that in December 2000 (three months after the image was taken), there were 129 840 refugees.

All automated classification methods result in a number of tents that lie within the statistical interval. These positive results were to be expected because of the 'clear' image: the tents are clearly different from the background and the sparse vegetation by their brightness, spectral signature, and shape. While the pixel-based methods yield acceptable results (with errors between 10% and 15%), the object-oriented methods have values closer to the statistical best estimate (errors $< 3\%$). The error values are calculated as the ratio of the difference between the statistical best estimate and the number of tents and the statistical best estimate.

To measure the accuracy of the computer assisted image analysis techniques we calculated omission and commission errors for the ground data of the 18 sample areas. While this approach is valid for both object-oriented methods, it is somewhat biased for the pixel-based methods since we used the same sample areas to establish the area threshold of the post-classification filtering step. Therefore, the omission and commission errors can be artificially low for the pixel-based methods. However, this is not necessarily the case. The error estimates tend to be stable over the image because of the regular distribution of tents within the camp, the fixed size



(*a*) (*b*)

Figure 5. (*a*) Satellite image of a fraction of the camp with superimposed visual interpretation results (red dots) and classification results (pink areas). (*b*) Results of the classification for an area 9 times larger than (*a*) (which falls within the blue rectangle on the bottom left).

range of tents and the discrete nature of the derived size threshold (the threshold never optimizes the omission and commission errors since it has to be a discrete number of pixels). We assume therefore that the omission and commission errors of all methods can be compared.

The accuracy of the automated methods is similar, each having classification errors between 10% and 15%. Upon visual inspection, most committed objects are bright objects with a position that is not on the parallel tent rows (figure 5). Omitted tents are in general dark or partially covered by tree canopies. Such omission and commission errors are inevitable, but can, if necessary, be corrected quite easily by a human interpreter in order to obtain a high quality camp map.

It is also clear that the supervised and unsupervised classification methods, which use a global brightness threshold to identify tents, have a disadvantage in parts of the camp covered with cloud shadows. The changing contrast between tent and background causes more frequent classification errors in these parts of the image (figure 1). The multi-resolution segmentation method and the mathematical morphology method are based on a combination of global and local image characteristics, which makes them more robust to contrast variations.

While all methods can be adapted for higher resolution imagery and different tent sizes, the object-oriented methods have an advantage over the pixel-based ones. They start with a segmentation or object extraction step in which objects are defined with the number of pixels they cover. Once these objects are defined, their classification is more robust since all pixels of the object are necessarily classified in the same class. This is not the case in pixel-based methods where a tent can be composed of multiple objects that are not connected. Of course, for all methods, the size related parameters (area threshold for the pixel-based methods, fuzzy size

of tents for eCognition, and maximum tent size for morphology) must be estimated accordingly.

Software and hardware requirements can be important in an operational use. While all methods need specialized software, the pixel-based approaches are more commonly available in image processing packages. The multi-resolution segmentation method is commercially available in the eCognition package, while the morphology method is available as the SDC Morphology Toolbox for MATLAB. Hardware requirements are similar for all methods. With large images such as the Lukole image (5458×6969 pixels) the analysis time varied between 15 minutes for the morphology method, one hour for the pixel-based methods using ERDAS, and eight hours for eCognition on a 1.6 GHz Intel processor in the Windows environment and 1 GB RAM. Splitting the image into three sub-images keeps the performance similar for the pixel-based and morphology method, but improves it drastically for eCognition (reducing the time to 20 minutes per sub-image). The first methods have a linear relation of treatment time to number of pixels, while eCognition has an exponential relation. While splitting the image is an obvious advantage for eCognition, it requires extra post-processing to merge the results. Another important factor for choosing between the methods can be the training requirements. Traditional methods are more commonly taught than more recent object-oriented approaches, which increases the initial costs of implementing the latter approaches.

6. Conclusion

Gathering statistics on refugees, including enumeration of refugees, and providing tools to help in the management of large refugee camps is an important task for the relief community. Decision makers from donor institutions need to know the magnitude of the need while implementers need to know precisely the resources required. This paper investigates the use of information technology and specifically the use of satellite imagery and image processing techniques to derive qualitative and quantitative information over refugee camps.

Very high resolution satellite imagery can provide the number of tents and the extent of a refugee camp. The image by itself provides an intuitive assessment of the plight of refugees. The power of the satellite image – even in its qualitative form as a picture from space – should not be underestimated especially when presented to an often non-technical donor community on which the refugees are dependent for survival. For example, images such as the one shown on the cover of this issue presented to the decision makers of donor institutions can be intuitive evidence of the need for aid. However, the real value of such imagery data lies in the quantitative information that can be derived from it.

This work has explored rapid ways to count tents from which the number of refugees can be estimated and to produce a camp map using computer assisted procedures applied to the satellite imagery. Visual interpretation of satellite imagery conducted over samples of tents in the camp provided by far the fastest method of estimating the total number of tents. Considering that no tents would be left unoccupied and that the relief community managing the camp has information on the average occupancy this method can provide a way to quickly compute the number of refugees. Accuracy of statistics would depend on the sampling size and the occupancy statistics.

Mapping tents within a refugee camp the size of Lukole can also be performed through visual interpretation. However, mapping involves identification as well as

outline of the tents. This task would be extremely time-consuming and error-prone if performed through visual analysis. We have therefore tested four computer assisted image analysis procedures that can provide alternatives to visual interpretation for camp mapping. All four procedures provided acceptable results from an operational point of view, with accuracies between 85% and 90%. However, in all cases additional visual interpretation is required to improve the overall accuracy for the actual use.

The plight of refugees and internally displaced people is, unfortunately, an issue civil society will continue to address. Information technology such as earth observation described in this paper can contribute to improve information for decision makers. Some of the information extraction procedures described in this paper could be used on imagery of other refugee camps to improve and standardize the process of refugee enumeration.

Acknowledgments

We would like to thank Mr F. J. Gallego (JRC) for helpful discussions in the early stage of the work and Mr B. Eckhardt (JRC) for the initial processing of the images.

References

ANONYMOUS, 1997, *Tanzania: refugee situation report*. UN DHA Integrated Regional Information Network (IRIN). http://www.reliefweb.int/w/rwb.nsf/s/C181C357EA-347169C125647D002F5B68.

BAATZ, M., and SCHAEPE, A., 2000, Multiresolution segmentation – an optimization approach for high quality multi-scale image segmentation. *Angewandte Geographische Informationsverarbeitung XII. Beitraege zum AGIT Symposium, Salzburg 2000*, edited by J. Stroble (Karlsruhe: Herbert Wichmann Verlag), pp. 12–23.

BALL, G. H., and HALL, D. J., 1967, A clustering technique for summarizing multivariate data. *Behavioural Sciences*, **12**, 153–155.

BJORGO, E., 2000a, Refugee camp mapping using very high spatial resolution satellite sensor images. *Geocarto International*, **15**, 77–86.

BJORGO, E., 2000b, Using very high spatial resolution multispectral satellite sensor imagery to monitor refugee camps. *International Journal for Remote Sensing*, **21**, 611–616.

CONGALTON, R. G., ODERWALD, R.G., and MEAD, R.A., 1983, Assessing Landsat classification accuracy using discrete multivariate analysis statistical techniques. *Photogrammetric Engineering and Remote Sensing*, **49**, 1671–1678.

CONGALTON, R. G., 1991, A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment*, **37**, 35–46.

CRISP, J., 1999, Who has counted the refugees? UNHCR and the politics of numbers. New issues in refugee research. Working paper No. 12. Centre for documentation and research, UNHCR, Geneva, Switzerland. http://www.unhcr.ch/cgi-bin/texis/vtx/home/+YwwBmeqmJ69wwwwqwwwwwwwhFqo20I0E2gltFqoGn5nwGqrAFqo20I0E2glcFqybrdarw5aqd1DBnmaBrnaGnh1tnn5eYX3qmxwwwwwww1FqmRbZ/opendoc.pdf.

DARE, P. M., and FRASER, C. S., 2001, Mapping informal settlements using high resolution satellite imagery. *International Journal of Remote Sensing*, **22**, 1399–1401.

ENVIREF, 2001, Environmental Monitoring of Refugee Camps using High-Resolution Satellite Images. http://www.enviref.org.

HEIJMANS, H., 1994, *Morphological Image Operators. Advances in Electronics and Electron Physics Series* (Boston: Academic Press).

JACOBSEN, K., 1997, Refugees' environmental impact: the effect of patterns of settlement. *Journal of Refugee Studies*, **10**, 19–36.

LODHI, M. A., ECHAVARRIA, F. R., and KEITHLEY, C., 1998, Using remote sensing data to monitor land cover changes near Afghan refugee camps in northern Pakistan. *Geocarto International*, **13**, 33–39.

SERRA, J., 1982, *Image Analysis and Mathematical Morphology* (London: Academic Press).

SOILLE, P., 1999, *Morphological Image Analysis* (Berlin, New York: Springer-Verlag).

SOILLE, P., and PESARESI, M., 2002, Advances in mathematical morphology applied to geoscience and remote sensing. *IEEE Transactions on Geosciences and Remote Sensing*, **40**, 2042–2055.

TANAKA, S., and SUGIMURA, T., 2001, A new frontier of remote sensing from IKONOS images. *International Journal of Remote Sensing*, **22**, 1–5.

UNHCR, 2000, Population estimation and registration. In *Handbook for Emergencies,* 2nd edn (Geneva: UNHCR Publications), pp. 118–131.

UNHCR, 2001, *Refugees by Numbers,* 2001 edn. www.unhcr.ch/cgi-bin/texis/vtx/home.

ZADEH, L.A., 1965, Fuzzy sets. *Information and Control*, **8**, 338–353.