GENERATING IMPACT MAPS FROM AUTOMATICALLY DETECTED BOMB CRATERS IN AERIAL WARTIME IMAGES USING MARKED POINT PROCESSES

C. Kruse [1], F. Rottensteiner [1], T. Hoberg [2], M. Ziems [2], J. Rebke [2], C. Heipke [1]

[1] Institute of Photogrammetry and GeoInformation, Leibniz Universität Hannover, 30167 Hanover, Germany - (kruse, rottensteiner, heipke)@ipi.uni-hannover.de
[2] State Office for Geoinformation and Surveying of Lower Saxony, 30659 Hanover, Germany - (thorsten.hoberg, marcel.ziems, julia.rebke)@lgln.niedersachsen.de

Commission III, ICWG II/III

KEY WORDS: Aerial Wartime Images, Bomb Craters, Marked Point Processes, RJMCMC, Simulated Annealing

ABSTRACT:

The aftermath of wartime attacks is often felt long after the war ended, as numerous unexploded bombs may still exist in the ground. Typically, such areas are documented in so-called impact maps which are based on the detection of bomb craters. This paper proposes a method for the automatic detection of bomb craters in aerial wartime images that were taken during the Second World War. The object model for the bomb craters is represented by ellipses. A probabilistic approach based on marked point processes determines the most likely configuration of objects within the scene. Adding and removing new objects to and from the current configuration, respectively, changing their positions and modifying the ellipse parameters randomly creates new object configurations. Each configuration is evaluated using an energy function. High gradient magnitudes along the border of the ellipse are favored and overlapping ellipses are penalized. Reversible Jump Markov Chain Monte Carlo sampling in combination with simulated annealing provides the global energy optimum, which describes the conformance with a predefined model. For generating the impact map a probability map is defined which is created from the automatic detections via kernel density estimation. By setting a threshold, areas around the detections are classified as contaminated or uncontaminated sites, respectively. Our results show the general potential of the method for the automatic detection of bomb craters and its automated generation of an impact map in a heterogeneous image stock.

1. INTRODUCTION

The state of Lower Saxony was the target of numerous air strikes by the allied forces during the Second World War. The aftermath of these attacks is still present today. Numerous unexploded bombs still exist in the ground. Especially for construction works in the earth the danger of a detonation of duds is given; there have also been incidents without external influence. A central task of Lower Saxony's explosive ordnance disposal service is the manual evaluation of aerial wartime images to identify suspicious locations concerning unexploded bombs. Despite the restriction to particularly endangered or otherwise relevant areas, this results in a high processing effort. For many questions, it is sufficient to have comprehensive information on the basic occurrence of warlike impacts in the form of "impact maps". An impact map indicates whether areas are likely to be contaminated or not. In this context, contaminated areas are expected to contain one or more duds with a high degree of certainty, whereas uncontaminated areas should not contain any dud. For the cost-efficient creation of such a map, an automatic recognition of indications of war loads in aerial wartime images, especially of bomb craters (Fig. 1), is indispensable.
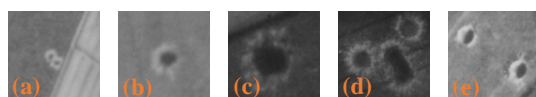

Figure 1: Bomb craters appear differently in aerial images due to varying scales and sizes (a-c) or different appearance (d, e).

The problem we want to solve is detecting bomb craters in aerial wartime images. The bomb craters are used as indicators of the areas in which unexploded bombs may be located. Thus, we want to deduce a probability for the occurrence of duds based on the

detections. This probability can be used to identify contaminated areas. In the paper, two scenarios are of interest. On the one hand, we want to detect areas that have a very high likelihood of containing a dud so that it makes sense to send a team of experts to that scene to probe it using geophysical detectors. As this is expensive, the focus is on avoiding false detections for this scenario. In the other scenario, we want to be able to exclude areas in advance in which there is a high probability that no dud exists. Here, bomb craters must not be missed, because an area falsely classified as uncontaminated could contain a dud, which might lead to detonation, e.g. in a building project.

In image analysis, probabilistic models are increasingly being used for the detection of objects. Examples for approaches that can consider contextual knowledge in a probabilistic framework are Conditional Random Fields (Kumar and Hebert, 2006), typically favoring similar classes for neighboring pixels. However, it is difficult to integrate more generic knowledge about the objects, for example models of object shape. This is where marked point processes (Descombes and Zerubia, 2002; Daley and Vere-Jones, 2003) come into play. A marked point process is a random variable whose realizations are configurations of geometrical objects. This stochastic method uses a strong object model while sampling provides the globally optimal configuration for objects of a certain type. Thus, knowledge about the objects is expressed in a more holistic way and characteristics of objects can be integrated beyond pixel-based relations. Marked point processes have shown to achieve good results in various object detection problems (e.g. Lafarge et al., 2010; Tournaire et al., 2010; Börcs and Benedek, 2015; Benedek, 2017; Schmidt et al., 2017). In addition to determining the globally optimal object configuration, they also offer a considerable flexibility in integrating knowledge about the

objects and their relationship to other objects. Furthermore, the number of objects in the scene during sampling is variable.

The individual detections of bomb craters could be relatively uncertain. This is why we would like to deduce area-based statements about the occurrence of warlike impacts. A useful tool to represent such data (detected bomb craters) is statistical modeling since it can provide area-wide information. Here, the existence of duds is represented by a probability density function (p.d.f.). In general, a p.d.f. can be modelled by parametric approaches, where an analytical model for the p.d.f. is assumed. This function depends on parameters, which are determined from training data. On the other hand, there are nonparametric approaches, which estimate the p.d.f. directly from the data. This approach avoids having to select a model and estimating its distribution parameters (Bishop, 2006). Kernel density estimation (Parzen, 1962), as a nonparametric density estimation technique, is quite popular (e.g. Scott, 2015).

In this paper, we want to use the advantages of marked point processes to detect bomb craters in aerial wartime images. The model we use are ellipses, and we find an optimal configuration of such objects by sampling. In the sampling process, high magnitudes of the gray value gradients at the ellipse borders are favored and an overlap of ellipses is penalized. In connection with kernel density estimation, we create a probability map based on the automatically detected bomb craters. Since our goal is to differentiate between potentially contaminated and uncontaminated areas, respectively, a threshold is set to classify the sites accordingly, which results in an impact map.

1.1 Related Work

Here, we focus on applications of marked point processes and on methods for detecting craters. In this context, we also consider papers dealing with the detection of planetary craters, because they have a similar appearance as bomb craters.

In connection with marked point processes, model knowledge can be integrated in different ways. Simple geometric primitives, which can be described by a small number of parameters, are typically used to represent the objects. Rectangles are frequently used to extract buildings or other man-made objects in the scene (Tournaire et al., 2010; Chai et al., 2012; Brédif et al., 2013). In these papers, marked point processes are applied to digital surface models. For a rectangle to be included in the object configuration, high gradient magnitudes of the heights at the rectangle border must be present and overlaps between rectangles are penalized. Rectangles have also been used to interpret facades of buildings based on rectified images (Wenzel and Förstner, 2016). Here, in contrast to the methods mentioned earlier, statistics of typical configurations of facade objects (windows, entrances) are learned from training data. In addition to rectangles, it is possible to use ellipses, e.g. for the detection of flamingos (Descamps et al., 2008) and seed products (Dubosclard et al., 2014), and several geometric primitives were used for object detection. Recently, Benedek (2017) proposed a method for extracting complex hierarchical object structures from digital images using different types of objects, namely ellipses, rectangles and isosceles triangles. In this embedded marked point process framework, object-subobject ensembles in parent-child relationship are admitted and corresponding objects may form coherent object groups. Bomb craters have no object-subobject relationships, so that complex hierarchical object structures are not necessary in our case. Marked point processes are also applied to biomedical imagery (e.g. Descombes, 2017). Due to the flexibility of the approach, biological variabilities can be handled. We have another application domain and the structure of the images is different.

To the best of our knowledge there are only two contributions dealing with marked point processes in the context of planetary crater detection (Troglio et al., 2010; Solarna et al., 2017). As the optimization of marked point processes is computationally expensive, Solarna et al. (2017) create a birth map from the contour map via Generalized Hough Transform and Gaussian filtering. In this way, only a part of the image has to be considered for sampling. Other approaches, not dealing with marked point processes, have also been proposed for planetary crater detection. For example, unsupervised (e.g. Meng et al., 2009) and supervised (e.g. Urbach and Stepinski, 2009) algorithms were applied. Unsupervised methods are frequently based on a Hough transform for detecting circles. However, this only works for a small dimensionality of the parameter space. In connection with supervised methods, machine learning models like boosting (Bandeira et al., 2010) or Convolutional Neural Networks (Cohen et al., 2016) were used. To the best of our knowledge, there are only two contributions dealing with the detection of bomb craters in aerial wartime images. Jensen et al. (2010) use a two-step approach. First, candidates are searched via cross correlation with representative crater-templates. Then, the candidates are classified by linear discriminant analysis. Merler et al. (2005) use different boosting approaches for the classification of image sections, which requires a high computational effort. Their result is a map of the spatial density of craters, an indicator for the risk of duds. However, algorithms based on machine learning need training data, which we want to avoid.

The listed articles show the potential of stochastic methods based on marked point processes in different application fields of image analysis. In addition to its applicability for heterogeneous image contents (e.g. different lighting situations, varying crater sizes), the discussion of the related work shows that marked point processes allow a flexible integration of knowledge about the objects, too. That is why we suppose the procedure of marked point processes is suitable for our application scenario. Hence, in this paper we propose a special type of marked point process based on ellipses for the detection of bomb craters in aerial wartime images. In contrast to images of planets, in our case not only craters but also objects such as trees or houses make correct detection more challenging. The detection of bomb craters using marked point processes and the subsequent inclusion of the results in an impact map are not yet described in the literature.

2. MATHEMATICAL BASICS

2.1 Marked point processes

Spatial point processes (Daley and Vere-Jones, 2003; Descombes, 2013) are stochastic processes and allow the mathematical description of random events depending on the position in space in remote sensing data sets. They offer the possibility to describe a scene by an unordered set of objects of a certain type within a limited region $F \subset R^2$. An object u_i is characterized by its position $p_i = (x_i, y_i)$ and a vector of additional parameters m_i. The object u_i is fixed by the position p_i and the *marks* m_i, which can be different from point to point, contain further information about the object. By adding marks to each point, a *marked point process* is created, which can be understood as a stochastic model of configurations of an unknown number of objects of type $u_i = (p_i, m_i)$ in F. The model of the homogeneous Poisson point process assumes a purely random distribution of objects in space that are not related

to each other. In a Poisson point process the probability $P_\lambda(n)$ for the number of objects n follows a discrete Poisson distribution

$$P_\lambda(n) = \frac{\lambda(F)^n}{n!} \cdot e^{-\lambda(F)}. \qquad (1)$$

The parameter λ describes the expected number of objects within F and is often referred to as intensity parameter. The probability for the positions of the objects follows a uniform distribution. In practice, the assumption of complete randomness often does not apply, because there are dependencies between the objects. Models that are more complex are required to measure the quality of the object configuration. To achieve this goal, a probability density $h(.)$ of the marked point process can be formulated with respect to a reference point process, which is usually defined as the Poisson point process. We define $h(.)$ by a Gibbs energy $U(.)$ in the form of $h \propto exp - U(.)$. The Gibbs energy consists of two parts, a data energy $U_D(.)$ and a prior energy $U_P(.)$. The relative influence is modeled by a weighting parameter $\beta \in [0, 1]$.

$$U(.) = \beta \cdot U_D(.) + (1 - \beta) \cdot U_P(.). \qquad (2)$$

$U_D(.)$ measures the conformity of the object configuration with the input data. Interactions between the objects are taken into account by $U_P(.)$. The optimal object configuration $u^* = \{u_1, \ldots, u_n\}$ can be determined by maximizing the probability density $h(.)$, i. e. $u^* = arg\, max\, h(.)$ or, equivalently, by minimizing the Gibbs energy $U(.)$, i. e. $u^* = arg\, min\, U(.)$. The probability density $h(.)$ is usually multi-modal and is defined in a configuration space with a variable dimension, because the number of objects can change. For this reason, a Reversible Jump Markov Chain Monte Carlo (RJMCMC) sampler in combination with simulated annealing is often used to estimate the global minimum of $U(.)$ to find an approximation of u^*.

2.2 Reversible Jump Markov Chain Monte Carlo sampling

RJMCMC methods allow the modeling of scenes with an unknown number of objects. This was first proposed by Green (1995) and is achieved by defining a set of changes (jumps) of the current configuration that are reversible. Reversibility means that it is possible to undo any change, so that one can always return to a previous state. In each iteration t, the sampler proposes a change of the current object configuration from the predefined set of jumps. For each type of change, there is a density function Q_m, which is also called kernel. This kernel Q_m leads from an object configuration X_t to a new configuration X_{t+1} according to a probability $Q_m(X_t \rightarrow X_{t+1})$. Following Metropolis et al. (1953) and Hastings (1970), the new configuration is accepted with a certain acceptance probability α depending on the energy difference of states X_t and X_{t+1}

$$\alpha = min\left(1, \frac{Q_m(X_{t+1} \rightarrow X_t)}{Q_m(X_t \rightarrow X_{t+1})} \cdot exp - \left(\frac{U(X_{t+1}) - U(X_t)}{T_t}\right)\right). \qquad (3)$$

In (3) the kernel ratio $Q_m(X_{t+1} \rightarrow X_t)/Q_m(X_t \rightarrow X_{t+1})$ describes the ratio of probabilities for changing the configuration from X_{t+1} to X_t and vice versa. $U(X_{t+1})$ and $U(X_t)$ are the Gibbs energy (2) of the new and current object configuration, respectively, and T_t is the temperature of the simulated annealing at iteration t. In order to find the optimum of the energy, the RJMCMC method is combined with simulated annealing (Metropolis et al., 1953; Kirkpatrick et al., 1983). The sequence of temperatures T_t tends towards zero while $t \rightarrow \infty$. A logarithmic cooling schedule guarantees convergence to the global optimum for any initial configuration X_0. However, this leads to high computation times, so that a faster cooling scheme based on a geometric sequence is typically used. This provides

an approximate result, which is usually close to the optimum (Van Laarhoven and Aarts, 1987).

2.3 Kernel Density Estimation

Kernel density estimation (e.g. Parzen, 1962) allows estimating the probability density function (pdf) of a random variable in a nonparametric way. Given a sample $(x_1, x_2, \ldots, x_n)$ drawn from a distribution with an unknown density p, an estimate $\hat{p}$ of this density can be calculated via

$$\hat{p}(x) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{x - x_i}{h}\right). \qquad (4)$$

Here, K is a kernel function and h is called the bandwidth parameter. Note that this definition of kernel is not to be confused with the kernels in connection with RJMCMC sampling from Section 2.2. The kernel function $K(k)$ has to be a non-negative function $(K(k) \geq 0)$ that integrates to one $(\int K(k)\, dt = 1)$. Equation (4) can be thought of as an estimate of the pdf by averaging the effect of a set of kernel functions centered on each data point. For this purpose, various kernel functions may be used. Often, the Gaussian kernel is considered, but other functions (e.g. uniform, triangular, Epanechnikov, etc.) can also be used. For further details, the reader is referred to Scott (2015).

3. METHODOLOGY

In this work, bomb craters are modeled as ellipses. A marked point process is used to find the optimal configuration of such objects in the scene. Minimization is based on RJMCMC sampling in combination with simulated annealing. In the optimization process, the object configuration changes continuously and is evaluated based on its conformity with a predefined model. Our model favors high gradient magnitudes along the border of the ellipse and penalizes overlapping ellipses. On the basis of the final object configuration, a probability map is created by kernel density estimation which allows to differentiate between potentially contaminated and uncontaminated areas.

3.1 Object model

The model used for the detection of the bomb craters are ellipses. As in (Perrin et al., 2006), each ellipse is described by its two-dimensional position (x, y) and a 3D mark (a, b, θ). (x, y) are the coordinates of the ellipse center in the image. The semi-major and the semi-minor axes are denoted by $a \in [a_m, a_M]$ and $b \in [b_m, b_M]$, respectively, where a_m and b_m are the minimum values while a_M and b_M are the maximum values. The orientation of the ellipses is described by the angle $\theta \in [0, \pi[$, which is defined clockwise relative to the positive x-axis. We require $a \geq b$ as well as $b > a/1.5$, because bomb craters are almost circular.

3.2 Changes in the object configuration

In the RJMCMC process, four types of changes to the current configuration are possible; modifications are implemented using the corresponding kernels. *Birth* and *death kernels*, Q_B or Q_D, with the associated proposition probabilities p_{Q_B} and p_{Q_D}, enable to add or remove an object to and from the current object configuration, respectively. In the case of a birth, the position of a new ellipse is sampled from likely positions for bomb craters detected in the way described in Section 3.4. This technique also provides information about the extent of the associated crater and, thus, it is used for the initialization of the two semi-axes.

The orientation is drawn randomly from a uniform distribution. For the death event, a randomly selected ellipse is removed from the existing object configuration. The kernel ratio, see equation (3), takes into account the probability of changing the configuration from X_t to X_{t+1} and vice versa. Similar to Schmidt et al. (2016), we model the kernel ratio of the birth event by

$$\frac{Q_B(X_{t+1} \to X_t)}{Q_D(X_t \to X_{t+1})} = \frac{p_{Q_D}}{p_{Q_B}} \cdot \frac{\lambda}{n}. \qquad (5)$$

The current number of ellipses within the object configuration is represented by n and the Poisson parameter λ describes the expected number of objects in the image. In the case of the death event, the kernel ratio corresponds to the inverse birth rate.

In addition, the parameters of an object configuration can be modified. The *translation kernel* Q_T moves a randomly chosen ellipse from its current position. For this purpose, a displacement vector is randomly generated in the local neighborhood based on a uniform distribution in a given interval. The three marks of the ellipse can be changed using the *mark-variation kernel* Q_M. An ellipse of the current configuration is randomly selected, from which the new semi-minor and semi-major axis as well as the orientation are drawn from a uniform distribution within predefined intervals. The kernel ratios of the *translation* and *mark-variation kernels* are set to one. The proposition probabilities p_{Q_T} and p_{Q_M} belong to these two types of changes.

3.3 Energy function

In each iteration step, the object configuration is compared with the new one based on the Gibbs energy (2) describing the consistency of the configuration with our bomb crater model.

3.3.1 Data energy: The data energy $U_D(X_t)$ from equation (2) describes the consistency of the object configuration with the input data. Bomb craters are predominantly characterized by locally darker gray values in comparison to the surrounding area. This is mainly due to the shadow cast by the sun. Its shape is mostly elliptical within the bomb craters (Fig. 1). A newly created or modified ellipse leads to a reduction of the data energy if high gradient magnitudes occur along the edges of the ellipse, i. e. the shape of the ellipse fits to the border of the shadow. We determine the gradients along the edge of the ellipse by

$$U_D(X_t) = \sum_{e_j \in X_t} \left(c - \frac{1}{n_b} \sum_{j=1}^{n_b} \nabla_{IMG_{p_j}}^{\flat} \right). \qquad (6)$$

In (6), $\nabla_{IMG_{p_j}}^{\flat}$ is the component of the gray value gradient at the border pixel p_j in the direction of the normal vector of the ellipse e_j pointing outside (Fig. 2). To calculate the sum of the gradients along the border of the ellipse, n_b pixels p_j are used. The edge of the ellipse is approximated by a polygon with a constant number of n_p vertices. We set $n_p = 32$. The constant $c \geq 0$ ensures that the energy only decreases if the sum in (6) is larger than c.

3.3.2 Prior energy: The prior energy penalizes configurations in which ellipses overlap. Usually, bomb craters do not overlap, or they overlap only to a small extent. That is why we allow a certain amount of overlap in the formulation of the prior term, while a strong overlap is avoided. As in (Perrin et al., 2005), all possible combinations of overlapping ellipse pairs e_i, e_j are considered and the overlapping areas A_{ij} of the ellipses e_i and e_j as well as the respective relative overlapping areas $A_{ij}/A_i(e_i)$ and $A_{ij}/A_j(e_j)$ are calculated in each case. Here, $A_i(e_i)$ and

$A_j(e_j)$ are the areas of the ellipses e_i and e_j, respectively. The prior energy (the strength of the penalization is weighted by f) is

$$U_P(X_t) = f \cdot \sum_{e_i, e_j} \max\left(\frac{A_{ij}}{A_i(e_i)}, \frac{A_{ij}}{A_j(e_j)} \right). \qquad (7)$$
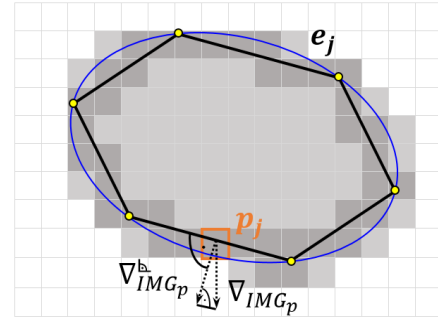

Figure 2: For each pixel p_j of the ellipse border, here approximated by six vertices (yellow dots), we consider the sum of the components of the gradients $\nabla_{IMG_{p_j}}^{\flat}$ in the direction of the normal vector (dark gray). ∇_{IMG_p} is the gradient in pixel p_j. Gray pixels indicate the expected shadow within a bomb crater.

3.4 Limitation of the search space

The procedure of the marked point processes is preceded by a method for restricting the search space in the image to reduce the computational effort for sampling (see Section 4). We use the blob detector, described in Mallick (2015) and implemented in OpenCV, for that purpose. A blob is a group of connected pixels that share common properties (e.g. similar gray values).

First, the image is converted into several binary images by applying thresholds. Beginning with a minimum threshold B_{T_min} this threshold is increased by means of a parameter for the step size B_{T_step} until a maximum threshold B_{T_max} is reached. Connected components are extracted from each binary image and their centers are calculated. The centers are grouped, with neighboring centers forming a group that corresponds to one blob. In this context, the procedure requires another parameter B_D, which ensures that blobs located closer than B_D are merged. The final centers and radii of the blobs are determined from the groups. The process also provides numerous filter options. Blobs can be filtered according to their circularity B_{Circ}, convexity B_{Conv}, the inertia ratio B_{InRat} or size B_N. The method also allows to detect only dark blobs, bright blobs or both types of blobs.

The result of the blob-detector is used to restrict the search space in the image for the optimization. This is realized by only allowing the birth of an ellipse at a blob location. In order not to miss any bomb craters, the blob detector is tuned so that as many craters as possible are detected (Section 4.1.2). This is important because other image locations are no longer taken into account. The initial configuration consists of an empty set of objects.

3.5 Impact map

Based on the automatically detected bomb craters we want to deduce a probability for a location that there are unexploded bombs nearby. The associated probability map is generated from the centers of the detected craters by kernel density estimation. As a kernel function we use $K(k) = (1 - |k|)$. The bandwidth h in equation (4) is related to the size of the radius for which a bomb crater increases the likelihood of a dud being within this

range. Consequently, if one wants to be sure that there is no dud in the neighborhood, a larger radius has to be chosen to minimize the risk of missing a dud. Using the probability map, our aim is to classify the entire scene in potentially contaminated and uncontaminated areas. For this purpose, a threshold p is applied to the probabilities resulting in an impact map. A pixel with a probability larger than p is classified as contaminated, which implies that a dud is likely to be present. On the other hand, the probability that a dud exists in an uncontaminated area is low. In this paper, two scenarios with different implications are of interest for the parameters h and p that are actually used.

In Scenario 1, we are interested in detecting areas that have a very high likelihood of containing a dud so that it makes sense to send a team of experts to that area to probe it, i.e. to visit the potentially contaminated area and take measurements using geophysical detectors. In general, probing is recommended for the areas of bomb craters as well as clusters of craters. Thus, the focus is on avoiding false detections, because they would cause high costs. Hence, the correctness of the results is important. i.e. that areas classified as contaminated are in fact contaminated. The main benefit for the employees of the explosive ordnance disposal service is that in these areas the aerial wartime images would then no longer have to be examined manually. In the impact map, potential probing areas appear in red while areas, which probably do not have to be probed, show green (Fig. 3b).

In Scenario 2, we want to exclude areas with a high probability of not containing a dud. Here, bomb craters must not be missed, because this could lead to a disaster. The results can be used to avoid the manual analysis of aerial images in uncontaminated areas. Aerial images only have to be inspected in the red areas of the impact map, but not in the green ones (Fig. 3c).
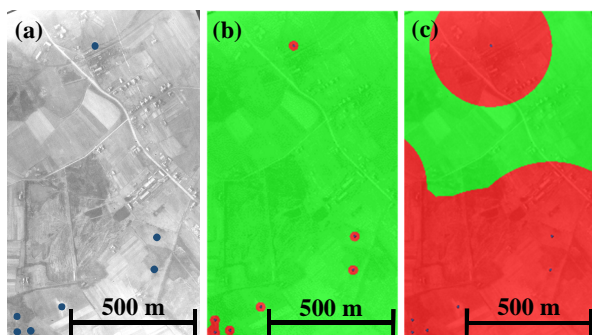

Figure 3: Subset of an aerial wartime image with bomb craters in blue (a) and two impact maps based on different bandwidths h for the Scenarios 1 and 2 of the subset with contaminated and uncontaminated areas shown in red and green, respectively (b, c).

4. EXPERIMENTS

4.1 Test data and test setup

4.1.1 Data: The evaluation of our method is based on panchromatic aerial wartime images scanned with 1200 dpi. They have a radiometric resolution of 8-bit and were captured by the Allies during World War II. Experts generated the area-based reference by manual annotation. In total, eight aerial wartime images are investigated (Table 1). The data sets (DS) have image scales of 1:5900 to 1:10600, resulting in ground sampling distances (GSD) of 0.13 m to 0.22 m. The areas (A) and the number of pixels (NP) of the images varies between 1.7 km² to 5.8 km² corresponding to $8.4 \cdot 10^7$ and $1.2 \cdot 10^8$ pixels. Furthermore, the degree of impact, i.e. the number of bomb

craters (NC) in each image varies between zero and 500. The investigated images are representative for certain cases, e.g. different lightning situations exist and the image content varies. The focus of the investigations and developments is on rural areas, because bomb craters are covered by the rubble of destroyed buildings in densely built-up areas and, thus, they cannot be seen in the images.

DS	Scale	A	GSD	NP	NC
I	1:5900	1.7 km²	0.13 m	$1.1 \cdot 10^8$	0
II	1:10000	3.9 km²	0.21 m	$8.6 \cdot 10^7$	24
III	1:8800	2.9 km²	0.19 m	$8.4 \cdot 10^7$	26
IV	1:8200	2.6 km²	0.17 m	$8.8 \cdot 10^7$	53
V	1:8200	2.8 km²	0.17 m	$9.2 \cdot 10^7$	140
VI	1:10600	5.8 km²	0.22 m	$1.2 \cdot 10^8$	345
VII	1:8400	2.7 km²	0.18 m	$8.5 \cdot 10^7$	425
VIII	1:9000	3.5 km²	0.19 m	$9.8 \cdot 10^7$	495

Table 1: Characteristics of the data sets used for evaluation. For an explanation of the abbreviations, see main text.

4.1.2 Parameter settings: The parameter selection for our experiments was done empirically. If not specifically indicated, the parameters for the eight images were set to identical values. The parameters of the blob detector (Section 3.4) are selected as $B_{T\_min} = 10$, $B_{T\_max} = 245$, $B_{T\_step} = 2$ and $B_D = 5$. This selection of B_{T_min} means that image errors, which usually appear very dark in the image, are mostly not detected as blobs. As bomb craters are generally darker than their surroundings, the procedure should only detect dark blobs. The parameters for filtering are set to $B_{Circ} \in \{0.1, 1\}$, $B_{Conv} \in \{0.4, 1\}$ and $B_{InRat} \in \{0.1, 1\}$, which allows craters to deviate from a circle. Selecting such loose restrictions on the parameters results in many false detections, but experiments have shown that a more restrictive choice excludes the detection of many bomb craters in advance. Similarly, the selection of B_N in the interval $[B_{N\_m}, B_{N\_M}]$ makes it possible to detect bomb craters with a different number of pixels and thus different sizes (B_{N_m} and B_{N_M} are the minimum and maximum values for the number of pixels). The variation of B_{N_m} for the eight data sets can be seen in Table 2, while B_{N_M} is set to 4900 and is kept constant. This variation is necessary because the size of the bomb craters varies depending on different factors like the image scale and the explosive force of the bomb.

The data and prior energy are weighted equally, i.e. β from equation (2) is set to $\beta = 0.5$. Simulated annealing uses a geometric cooling scheme, using $T_t = T_0 \cdot f_T^t$. Here, we set the start temperature $T_0 = 100$ and set to $f_T = 0.999994$. The lower and upper limits of the semi-major and semi-minor axes (Section 3.1) are derived from the minimum and maximum blob radius $B_r \in \{B_{r\_min}, B_{r\_max}\}$ occurring in the image after the blob detection. This results in $B_{r\_min} = a_m = b_m$ and $B_{r\_max} = a_M = b_M$, respectively. The proposition probabilities of the kernels from Section 3.2 are set to $p_{Q_B} = 0.00025$, $p_{Q_D} = 0.00475$ and $p_{Q_T} = p_{Q_M} = 0.4975$. The probability of selecting a birth event needs to be significantly lower than the probability of selecting a death event, because many of the detected blobs do not actually correspond to bomb craters. The Poisson parameter λ from equations (1) and (5) depends on the expected number of objects in the image. In order to avoid manual intervention, λ is set to $\lambda = \#blobs/20$, where $\#blobs$ is the number of blobs. Moreover, the parameter c from equation (6) varies between 150 and 250 (see Table 2), because the gray value gradients show different degrees of intensity in the images. Here, the selection of c was performed empirically based on visual inspection of image contrast. Based on the formulations in Section 3.3.2, we select

the factor $f = 1000$ (equation 7). Experiments have shown that this value allows for minor overlaps of objects.

	Data sets							
	I	II	III	IV	V	VI	VII	VIII
B_{N_m}	1156	400	576	576	676	576	400	324
c	250	240	200	150	250	225	200	250

Table 2: Parameter selection for B_{N_m} and c.

In connection with the kernel density estimation, impact maps with a radius of 20 m (Scenario 1) and 250 m (Scenario 2) are derived. For this purpose, the bandwidth h from equation (4) is varied based on the image scale and an appropriate threshold p for the probabilities is set in a way that the area around the center of a detected bomb crater is always classified as contaminated within a radius of 20 m respectively 250 m for single detections. As areas of bomb craters and their immediate surroundings are likely to contain duds, the radius of 20 m is set relatively small in order to detect those areas that probably need to be probed by experts. On the other hand, Scenario 2 uses a relatively large radius to minimize the risk that a dud remains in the area classified as potentially uncontaminated.

4.1.3 Evaluation criteria: The evaluation of the results is carried out both object-based and pixel-based. For the object-based evaluation, we consider the bomb craters detected by the procedure of the marked point processes and the area based reference data for the evaluation. For each reference crater we know its position and radius. If the distance from the center of the ellipse to the corresponding reference center is smaller than the radius of this crater, the ellipse is defined as correct. The correctness is the percentage of ellipse centers which fulfill this criterion. The completeness of the object configuration is the percentage of bomb craters found by our method. The pixel-based evaluation is based on the impact maps (Section 3.5). For both scenarios, the reference impact maps are derived from the reference centers of the craters (the size of the radius corresponds to the one used for the generation of impact maps from the ellipse centers). The corresponding impact maps from the reference and the automatic detection are compared and each pixel is classified as being either a True Positive (TP), False Negative (FN), False Positive (FP) or True Negative (TN). A TP is a pixel that has been correctly classified as contaminated in both the reference and automatic detection. A TN pixel was correctly classified as uncontaminated in both cases. FN pixels were falsely classified as uncontaminated by the automatic detection. Finally, FP pixels have been falsely classified as contaminated. The completeness is the percentage of the actually contaminated area found by our method, i.e. TP / (TP + FN). The correctness is the percentage of areas from the automatic detection which lie in the actually contaminated area, i.e. TP / (TP + FP). The quality considers both types of errors, FN and FP (Heipke et al., 1997).

4.2 Results

Figure 4a shows the resultant optimal object configuration after the sampling procedure for a part of data set VII. In this example, there is one false detection (Fig. 4a, cyan ellipse) and comparably few bomb crater are not detected (Fig. 4a, red arrows). This observation is reflected in the superposition of the two impact maps for the reference centers and the centers of the automatic results (Fig. 4a) for Scenario 1 (Fig. 4b). Only a few areas were falsely not detected as contaminated (Fig. 4b, red). The areas that were wrongly classified as contaminated are negligible (pale blue areas in Fig. 4b). As correctness is important for this scenario, the procedure provides very good results for this example. In other

words, this implies that almost all areas which have been classified as contaminated (Fig. 4b, dark green and pale blue) actually have to be probed. The results for Scenario 2 are shown in Figure 4c, here for the complete aerial image V. It can be noticed that the area falsely classified as uncontaminated (Fig. 4c, red), which is of importance for this scenario, is small compared to the complete area of the aerial image. Thus, it is very likely that the majority of the areas classified as uncontaminated (Fig. 4c, pale green and red) are actually free of duds. We can therefore also speak of a good result here.
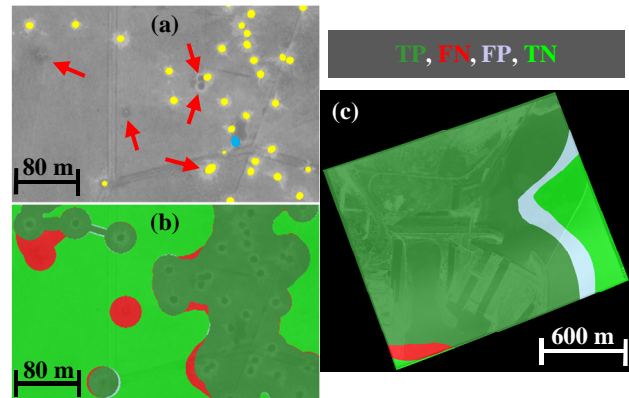

Figure 4: Subset of the aerial wartime image VII with the final object configuration for the marked point process procedure (ellipses) with correct detections in yellow and false detections in cyan (a), the derived impact map for Scenario 1, showing the superposition of the impact maps for the reference and the automatic results with TP-areas in dark green, FN-areas in red, FP-areas in pale blue and TN-areas in lime green (b) and the impact map superimposed on data set V for Scenario 2 (c).

The numerical values of completeness (CP), correctness (CR) and quality (Q) for the object-based evaluation can be found in Table 3. Besides the quality measures for the investigated data sets (DS), it again shows the number of bomb craters (NC, same values as in Table 1). Additionally, the number of detected ellipses (NE) is given. The results for the evaluation based on the bomb craters show that an average completeness and correctness of 57 % and 49 %, respectively, can be achieved (Total NC-NE, Table 3). These quality measures for the individual detections are too small to be useful for our purposes. In connection with the pixel-based evaluation, the impact maps for both scenarios are considered (Table 4; the abbreviations are the same as in Table 3). The results for Scenario 1 show that an average completeness and correctness of 63 % and 53 %, respectively, can be achieved by considering the total area. For Scenario 2, these values are significantly higher, at 98 % and 81 %, respectively. Thus, the results for Scenario 2 are useful, as the completeness, which is important for this scenario, is close to 100 % (Table 4, underlined). Nevertheless, there is still a small residual risk that a dud is present within the areas falsely classified as uncontaminated. In general, the completeness and correctness of Scenario 1 would be to too low to integrate the results into the workflow of the explosive ordnance disposal service. With a correctness of 53 % (Table 4, underlined), too many areas would be falsely probed, resulting in enormous costs. Furthermore, the results vary a lot. In some cases, such as for data sets VII and VIII, the completeness of 89 % and 63 %, respectively, is acceptable. However, the impact maps for other aerial images (e.g. II, III and IV) are far too uncertain and, thus, unreliable. It should be noted that if there is a small number of craters, the evaluation results should be considered with caution, as a small number of errors already have a significant impact on the results.

DS	NC	NE	CP [%]	CR [%]	Q [%]
I	0	34	N/D	0	0
II	24	38	21	13	9
III	26	114	23	5	4
IV	53	110	68	33	28
V	140	260	47	25	20
VI	345	403	56	48	35
VII	425	336	68	86	61
VIII	495	453	52	57	37
Total NC-NE			**57**	**49**	**35**

Table 3: Evaluation based on bomb craters (N/D: not defined).

DS	NC	Scenario 1			Scenario 2		
		CP [%]	CR [%]	Q [%]	CP [%]	CR [%]	Q [%]
I	0	N/D	0	0	N/D	0	0
II	24	23	14	10	88	61	56
III	26	24	6	5	100	72	72
IV	53	66	32	28	100	84	84
V	140	55	33	26	98	94	92
VI	345	60	51	38	100	93	93
VII	425	75	89	69	100	100	100
VIII	495	61	63	45	99	100	99
Total area		**63**	**53**	**40**	**98**	**81**	**80**

Table 4: Evaluation based on impact maps (N/D: not defined).

In contrast to the example shown in Figure 4, there are often objects that have a similar appearance as bomb craters, which leads to numerous false detections. This is in particular due to shadows cast by houses and trees (Fig. 5). If stereoscopic imagery is available, one way of counteracting these limitations is to integrate height information. To check whether the integration of 3D information is useful, we simulated this information: For this purpose, inside six of the examined aerial images, high objects (trees, houses, others) including their shadows were manually masked out, i.e. pixels of masked areas were not considered for the detection procedure and not for its evaluation, either. The other two images did not have high objects. In the case of the object-based evaluation, the integration of 3D resulted in an increase of approx. 12 % in correctness while maintaining the same level of completeness. The same could be determined for Scenario 1. As in this scenario, correctness is important in order to avoid unnecessary probing, the integration of height information seems very useful. For Scenario 2, no significant changes occurred with regard to the quality measures. Thus, the integration of 3D does not seem necessary.
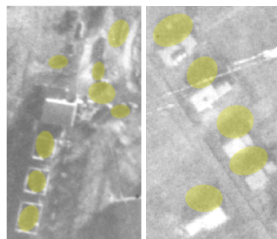

Figure 5: Objects that appear in a way that is similar to bomb craters in the image (e.g. houses or shadows of trees and houses) lead to numerous false detections.

We point out two more general aspects related to the quality of the results. Sometimes bomb craters are located in low-contrast regions or appear in a way that is not representative for the model used (see e.g. the two craters marked by red arrows to the left in Fig. 4a, craters that have small gradients or were filled up and

thus appear bright in the image). These craters are currently not detected which has a negative effect on the completeness of the results. On the other hand, the spatial representation by means of impact maps, especially for Scenario 2, has the following advantage: not every crater needs to be found, as the detection of other craters means that their surroundings are already classified as contaminated (see e.g. the three undetected craters marked by red arrows in the right part of Fig. 4a and the corresponding impact map for Scenario 1 in Fig. 4b).

5. CONCLUSION AND OUTLOOK

We have presented a procedure for the detection of bomb craters in aerial wartime images. The results for the individual detections with a completeness and correctness of 57 % and 49 %, respectively, are quite low. In addition, the representation of the individual bomb craters is not descriptive with regard to the occurrence of warlike impacts. Deriving an impact map from the detected bomb craters provides a quick overview of contaminated and uncontaminated areas for the respective application. In Scenario 1, we are interested in detecting areas that have a very high likelihood of containing a dud so that it makes sense to send a team of experts to that area to probe it. In Scenario 2, we want to exclude areas in advance in which there is a high probability that no dud exists. Our results show the potential, but also the limitations of the method. In connection with the impact maps, our method gives useful results in relation to Scenario 2, with a completeness of 98 % and thus only a small remaining risk. For Scenario 1, the correctness of 53 % would be to too low for an integration into the workflow of the explosive ordnance disposal service, because too many areas would be probed unnecessarily, resulting in enormous costs. It seems, however, that additional height information will increase the quality of these results.

A problem in connection with false detections arises from objects that appear like bomb craters in the image, e.g. shadows of houses or trees. In order to overcome this problem, in the future the data term is to be extended to include a band of brighter pixels around the border of the ellipse (e.g. Descamps et al., 2011). This pheno-menon is often observed in the images due to the earth ejected by the detonation. Furthermore, the object model will be modified in a way that it is also possible to detect filled craters, which sometimes appear in the images and are usually brighter than their surroundings. For this purpose, the direction of the consi-dered gradients would have to be reversed (Section 3.3.1). Also, the integration of height information will be considered, although we first need to assess the image quality with regard to computing height. In addition, in future investigations, the number of points representing the ellipse border will be selected automatically depending on the image scale and the expected bomb crater sizes in the scene, and ellipses could be approximated by more points in areas of high curvature than in areas of low curvature. Finally, the parameter c from equation (6) will be made dependent on the standard deviation of the gray values. For this purpose, the image will be divided into a grid and for each grid cell, the parameter c will be set as a function of the local standard deviation, avoiding the need to examine the image before processing.

REFERENCES

Bandeira, L., Ding, W., Stepinski, T.F., 2012. Detection of sub-kilometer craters in high resolution planetary images using shape and texture features. *Advances in Space Research*, 49 (1): 64–74.

Benedek, C., 2017. An Embedded Marked Point Process Framework for Three-Level Object Population Analysis. *IEEE Transactions on Image Processing: a publication of the IEEE Signal Processing Society*, 26 (9), pp. 4430–4445.

Bishop, C. 2006. *Pattern Recognition and Machine Learning* (1st edition). Springer, New York, USA.

Börcs, A., Benedek, C., 2015. Extraction of vehicle groups in airborne LiDAR point clouds with two-level point processes. *IEEE Transactions on Geoscience and Remote Sensing*, 53 (3), pp. 1475–1489.

Brédif, M., Tournaire, O., Vallet, B., Champion, N., 2013. Extracting polygonal building footprints from digital surface models: A fully-automatic global optimization framework. *ISPRS Journal of Photogrammetry and Remote Sensing*, 77, pp. 57–65.

Chai, D., Förstner, W., Yang, M.Y., 2012. Combine Markov random fields and marked point processes to extract building from remotely sensed images. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, I-3, pp. 365–370.

Cohen, J.P., Lo, H.Z., Lu, T., Ding, W., 2016. Crater Detection via Convolutional Neural Networks. http://arxiv.org/abs/1601.00978 (10 September 2017).

Daley, D.J., Vere-Jones, D., 2003. An Introduction to the Theory of Point Processes: Volume 1: Elementary Theory and Methods (2nd edition). Springer, New York.

Descamps, S., Descombes, X., Béchet, A., Zerubia, J., 2008. Automatic flamingo detection using a multiple birth and death process. *IEEE International Conference on Acoustics, Speech and Signal Processing* (ICASSP), pp. 1113-1116.

Descamps, S., Béchet, A., Descombes, X., Arnaud, A., Zerubia, J., 2011. An automatic counter for aerial images of aggregations of large birds. *Bird study*, 58 (3), pp. 302–308.

Descombes, X., Zerubia, J., 2002. Marked point process in image analysis. *IEEE Signal Processing Magazine*, 19 (5), pp. 77–84.

Descombes, X., 2013. *Marked Point Processes for Object Detection*. In: *Stochastic Geometry for Image Analysis*. John Wiley & Sons, Inc, Hoboken, NJ, USA, pp. 11–27.

Descombes, X., 2017. Multiple objects detection in biological images using a marked point process framework. *Methods*, 115, pp. 2–8.

Dubosclard, P., Larnier, S., Konik, H., Herbulot, A., Devy, M., 2014. Automatic method for visual grading of seed food products. *International Conference on Image Analysis and Recognition* (ICIAR), 1, pp. 485-495.

Green, P.J., 1995. Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika*, 82 (4), pp. 711–732.

Hastings, W.K., 1970. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57 (1), pp. 97–109.

Heipke, C., Mayer, H., Wiedemann, C., Jamet, O., 1997. Evaluation of automatic road extraction. *International Archives of Photogrammetry and Remote Sensing*, XXXII/3–4W2, pp. 151–160.

Jensen, L., Drauschke, M., Förstner, W., 2010. Automatische Detektion von Bombentrichtern in digitalisierten Luftbildern des 2. Weltkriegs. In: *DGPF Tagungsband 19/2010*, pp. 228-235.

Kirkpatrick, S., Gelatt, C.D., Vecchi, M.P., 1983. Optimization by simulated annealing. *Science*, 220 (4598), pp. 671–680.

Kumar, S., Hebert, M., 2006. Discriminative random fields. *International Journal of Computer Vision*, 68 (2), pp. 179–201.

Lafarge, F., Gimel'farb, G., Descombes, X., 2010. Geometric feature extraction by a multimarked point process. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32 (9), pp. 1597–1609.

Mallick, S., 2015. Blob Detection Using OpenCV (Python, C++). https://www.learnopencv.com/blob-detection-using-opencv-python-c/ (12 October 2017).

Meng, D., Yunfeng, C., Qingxian, W., 2009. Method of Passive Image Based Crater Autonomous Detection. *Chinese Journal of Aeronautics*, 22 (3), pp. 301–306.

Merler, S., Furlanello, C., Jurman, G., 2005. *Machine Learning on Historic Air Photographs for Mapping Risk of Unexploded Bombs*. In: Roli, F., Vitulano, S. (Eds.) *Image Analysis and Processing – ICIAP 2005*. Lecture Notes in Computer Science (LNCS), Springer, Berlin, Heidelberg, 3617, pp. 735-742.

Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., Teller, E., 1953. Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21 (6), pp. 1087–1092.

Parzen, E., 1962. On estimation of a probability density function and mode. *The Annals of Mathematical Statistics*, 33 (3), pp. 1065–1076.

Perrin, G., Descombes, X., Zerubia, J., 2005. Adaptive simulated annealing for energy minimization problem in a marked point process application. In: Rangarajan, A., Vemuri, B., Yuille, A.L. (Eds.) Energy Minimization Methods in Computer Vision and Pattern Recognition. Lecture Notes in Computer Science (LNCS), Springer, Heidelberg, 3757, pp. 3-17.

Perrin, G., Descombes, X., Zerubia, J., 2006. 2D and 3D Vegetation Resource Parameters Assessment using Marked Point Processes. *18th International Conference on Pattern Recognition* (ICPR'06), pp. 1–4.

Schmidt, A., Kruse, C., Rottensteiner, F., Sörgel, U., Heipke, C., 2016. Network detection in raster data using marked point processes. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLI-B3, pp. 701–708.

Schmidt, A., Lafarge, F., Brenner, C., Rottensteiner, F., Heipke, C., 2017. Forest point processes for the automatic extraction of networks in raster data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 126, pp. 38–55.

Scott, D.W., 2015. *Kernel Density Estimators*. In: *Multivariate Density Estimation*. John Wiley & Sons, Inc, pp. 137–216.

Solarna, D., Moser, G., Le Moigne, J., Serpico, S.B., 2017. Planetary Crater Detection and Registration Using Marked Point Processes, Multiple Birth and Death Algorithms, and Region-Based Analysis. *IEEE International Geoscience and Remote Sensing Symposium* (IGARSS), pp. 2337–2340.

Tournaire, O., Brédif, M., Boldo, D., Durupt, M., 2010. An efficient stochastic approach for building footprint extraction from digital elevation models. *ISPRS Journal of Photogrammetry and Remote Sensing*, 65 (4), pp. 317–327.

Troglio, G., Benediktsson, J., Moser, G., Serpico, S., 2010. Crater Detection Based on Marked Point Processes. *International Geoscience and Remote Sensing Symposium* (IGARSS), pp. 1378-1381.

Urbach, E.R., Stepinski, T.F., 2009. Automatic detection of sub-km craters in high resolution planetary images. *Planetary and Space Science*, 57 (7), pp. 880–887.

Van Laarhoven, P.J., Aarts, E.H., 1987. *Simulated Annealing: Theory and Applications*. Springer Netherlands, Dordrecht.

Wenzel, S., Förstner, W., 2016. Facade Interpretation Using a Marked Point Process. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, III-3, pp. 363–370.