2014 IEEE International Conference on Acoustic, Speech and Signal Processing (ICASSP)

AUTOMATIC INITIALIZATION FOR NAVAL APPLICATION OF GRAPH SEGMENTATION TECHNIQUES: A COMPARATIVE STUDY

Irene Camino, Udo Zölzer

Helmut Schmidt University
Hamburg, Germany

ABSTRACT

Nowadays, many different image processing applications are of high interest to maritime authorities because of security reasons. Depending on the application, different kinds of images are employed. The extraction of ship silhouettes requires high resolution images in order to obtain accurate results. However, when the characteristics of the naval environment are visible the background complexity increases greatly and automatic approaches fail. In order to overcome these difficulties we propose an automatic initialization for graph segmentation techniques. A comparative study of earlier suggested initializations for different graph segmentation techniques is also presented. It shows that, under such unfavorable image conditions, finding the proper initialization in an automatic way is not trivial. Yet, the precision and recall achieved by our initialization are considerable higher regardless the graph segmentation. Furthermore, the performance is highly increased since the best results are obtained after only the first iteration.

Index Terms— naval images, automatic initialization, silhouette extraction, graph segmentation

1. INTRODUCTION

Due to security reasons, there is a wide range of image processing applications that are of great interest to maritime authorities. The images employed differ according to the specific requirements of these activities. For example, remote sensing imagery has been very often employed in coastal surveillance due to its long operating distance and wide monitoring range. Infrared images are widely used as well. However, this paper concentrates on high resolution images, which are more suitable for extracting precise information, as ship silhouettes. Ship silhouettes are valuable features, also to build 3D models which can in turn be employed as reference for further ship classification.

The characteristics of infrared images allow to simply search for the darkest or brightest region [1] in the image. The same fact makes the segmentation performed in [2] a task of little difficulty, according to them. In remote sensing images, ships usually have opposite gray values of the sea region. Furthermore, the gray distribution of the background is barely spread. This may facilitate the ship localization as noted by [3]. Simple shape analysis [4], histogram based segmentation [5], filtering by local gradient analysis [6] or the application of an appropriate threshold [7] are enough to extract the ship candidate. Sometimes user interaction is requested [8] or indicated if the gray distribution of the sea region is very disperse [9].

As the image resolution increases, the background simplicity decreases and the previous techniques are no longer successful. Due to the reflection effects the gray distribution varies highly, which might also lead to the existence of regions in the water. The incidence of light modifies not only the intensity of the background but also the color. Besides, the number of edges produced by waves is particularly high, which rarely occurs in other scenarios. The waves might even have stronger edges than the ship, specially in wakes in the direction of the ship.

A course localization, which is obtained under these conditions, is suggested in [10]. However, when it comes to the precision necessary to build 3D models, a further improvement is required as stated there. In order to extract the ship silhouettes, we propose the use of graph segmentation techniques, which are proved in this paper to overcome these difficulties if they are properly initialized.

Despite of their high performance, the initialization is precisely one of the drawbacks of graph techniques. Furthermore, for an accurate segmentation user input is required unless an automatic procedure is specifically developed. Automatic initializations have been proposed in different contexts, as the detection of arbitrarily shaped buildings [11], human segmentation [12] or clothing extraction [13]. However, due to the complexity of the naval backgrounds in these images an automatic initialization that leads to precise results is not straightforward. Thus, the proposed initialization is compared to a ten percent frame initialization and a previous approach presented in [14], when the GrabCut [15], MorphCut [16] and an extended MorphCut [14] are applied.

This paper is organized as follows. The process of extracting the ship silhouette, which is composed of two parts, is shown in Section 2. The different initialization procedures to be compared, including the proposed initialization, are de-

978-1-4799-2893-4/14/$31.00 ©2014 IEEE

scribed in subsection 2.1. The graph segmentation techniques that are applied to them are summarized in subsection 2.2. Final results are evaluated in Section 3 and conclusions and further considerations are drawn in Section 4.

2. SILHOUETTE EXTRACTION

2.1. Automatic Initializations

The image conditions faced (see Fig. 1) cause that automatic initializations of graph segmentations are not trivial. This is proved by applying them to several different initializations, which are presented in the following.

Assigning a frame of the image as background, as in [13], is a simple automatic initialization. The portion of image used as frame can vary. For this study, a 10% portion of the image is indicated as background (see Fig. 2a).

A second initialization, which is developed also for this kind of images, is taken under consideration [14]. A grid of seed points is produced in the image and random pixels are assigned to them. Furthermore, a mesh point is calculated for each seed point and all of them are triangulated. A path resistance between the knots of the mesh is calculated based on the similarity between seed points and the silhouette is created. The result of Fig. 1 is shown in Fig. 2f.

The proposed initialization is a shape based approach and intends to take advantage of the adverse conditions of this environment. Despite of the variety of edges in these images, those that contribute to give the semantic sense of ship are, in fact, straight lines. Based on this simple but effective fact, the next steps are developed.

- The ship centerline, defined by [3] as the line connecting the point of the bow through the center of the stern, is estimated first. The Standard Hough Transform (SHT) [17] is applied to detect only the longest lines in the image. Next, the retrieved lines are clustered according to their orientation θ. Each group j is composed of l_j lines, where $j = 0, \ \ldots, n-1$ and n is the number of groups. The line i of the group j fulfills that $\theta_{i,j} \in [\theta_{0,j}, \theta_{0,j} + \varphi]$ with $i = 0, \ldots, l_j - 1$ and $\theta_{0,j} + \varphi < \theta_{0,j+1}$. The orientation of the centerline is then given by $\theta_{cl} = \frac{\sum_{i=0}^{l_{\hat{j}}-1} \theta_{i,\hat{j}}}{l_{\hat{j}}}$, where $\hat{j} = \arg\max_j l_j$ is the selected group.

- A set of feature angles, θ_d, θ_k and θ_m, is defined for the enhancement of meaningful lines, i.e. those that determine the ship. They are assumed to lie within certain ranges and fitted with respect to the centerline, which makes the algorithm independent of bearing and elevation changes. The voting process of the SHT is modified in order to ensure their enhancement.

Fig. 1: Original image

- In order to eliminate waves or disturbances that might belong to the feature angles model, a segment based processing is performed in two steps. One of the properties of the SHT is the accumulation of points as part of lines even though there might be gaps in between. Thus, it can be used to remove them. Subsequently, the last spurious segments are pruned setting an adaptive threshold that takes in consideration the eventual symmetry and elongated shape of the ships.

The Graham's scan algorithm [18] is applied to the endpoints of the remaining segments, achieving the convex initialization of Fig. 1 that can be seen in Fig. 2k.

2.2. Graph Segmentation Techniques

The GrabCut algorithm is an iterative segmentation for RGB color space. The color information is introduced by Gaussian Mixture Models (GMMs). A full-covariance Gaussian mixture with K components is employed for foreground pixels and another for background pixels,

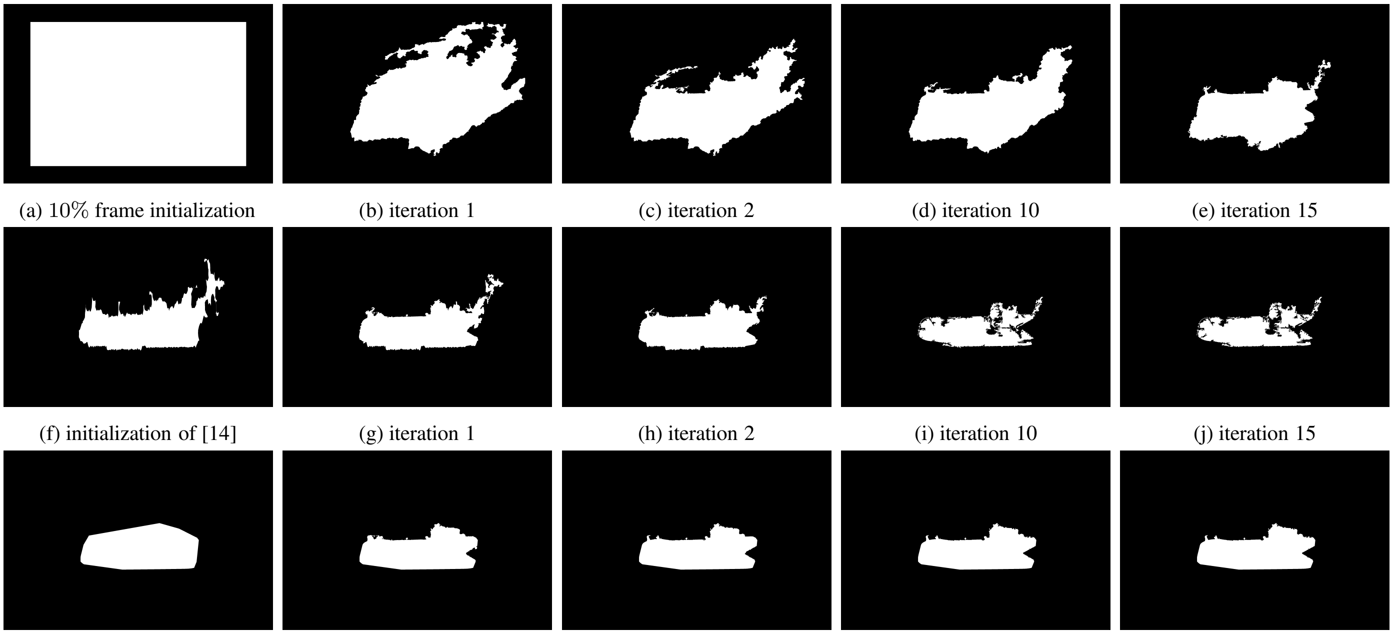
$$\boldsymbol{\theta} = \{\pi(\alpha,k), \mu(\alpha,k), \Sigma(\alpha,k), \alpha \in \{0,1\}, k = 0, \ldots K\}, \tag{1}$$

where α is the segmentation array, π are the weights, μ the means and Σ the covariances. Additionally, a vector $\boldsymbol{k} = \{k_1, \ldots, k_N\}$, where $k_i \in \{1, \ldots, K\}$, indicates to which GMM component each pixel belongs. After an initial trimap T is indicated, the energy minimization process begins. The energy for segmentation is given by

$$\boldsymbol{E}(\boldsymbol{\alpha}, \boldsymbol{k}, \boldsymbol{\theta}, \boldsymbol{z}) = U(\boldsymbol{\alpha}, \boldsymbol{k}, \boldsymbol{\theta}, \boldsymbol{z}) + V(\boldsymbol{\alpha}, \boldsymbol{z}), \tag{2}$$

where $\boldsymbol{z}$ represents the image array, $U(\boldsymbol{\alpha}, \boldsymbol{k}, \boldsymbol{\theta}, \boldsymbol{z})$ is the data term that considers the color GMM models and $V(\boldsymbol{\alpha}, \boldsymbol{z})$ is the smoothness term.

The second segmentation, MorphCut, is an extension of the GrabCut algorithm. It inserts an intermediate step that allows to reconsider the labels of pixels around the segmentation border in the next iteration. Thus, convergence is slower and furthermore, the energy no longer decreases monotonically. Nevertheless, as noted by the authors, the GMMs fit better to the color distribution of these images and therefore a stable result is also achieved. An example can be observed in Fig. 2.



(a) 10% frame initialization (b) iteration 1 (c) iteration 2 (d) iteration 10 (e) iteration 15

(f) initialization of [14] (g) iteration 1 (h) iteration 2 (i) iteration 10 (j) iteration 15

(k) proposed initialization (l) iteration 1 (m) iteration 2 (n) iteration 10 (o) iteration 15

Fig. 2: The results of applying the MorphCut to the three different initializations of Fig. 1 are shown in each line. Black pixels represent the background and white pixels the foreground.

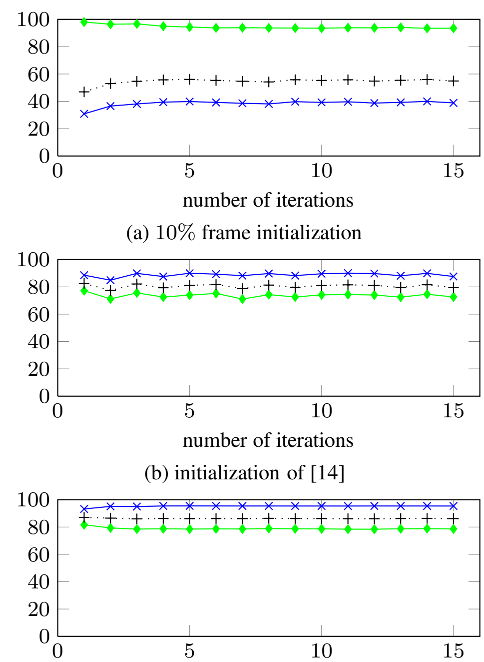
A pyramid structure to scale the input image of MorphCut is suggested by [19]. The levels of the pyramid, or the scaling factors, are based on the number of iterations selected. The first initialization of the GMMs is employed for the rest of iterations assuming they remain practically constant. However, in order to obtain a stable result, the initialization of the GMMs before every new iteration is suggested by [14], which is the extension of MorphCut finally employed.

3. EVALUATION

The evaluation of the different silhouette extraction procedures is performed on a dataset of 51 images. These images contain different models of ships under different illumination conditions. Furthermore, there are no constraints regarding their elevation or bearing. No simulated data is employed.

In order to assess the quality of the results, the ship silhouette is manually extracted from each image. The handmade ship silhouettes are compared to the results obtained for each initialization and segmentation. Thus, precision and recall are calculated. The average precision of the proposed initialization, i.e. before applying any segmentation technique, is 74.74%. And the recall is on average 85.95%.

The results of GrabCut along 15 iterations can be observed in Fig. 3. When it is initialized by the 10% frame, results reach stability after very few iterations. However, its quality is considerably poor compared to the other initializations. The second initialization leads to worse results than the



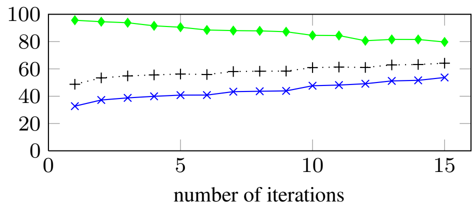
(a) 10% frame initialization

(b) initialization of [14]

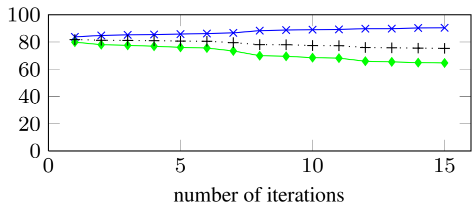
(c) proposed initialization

Fig. 3: Average precision ('×'), average recall ('◇') and F-measure ('+') of the image set when GrabCut is used.

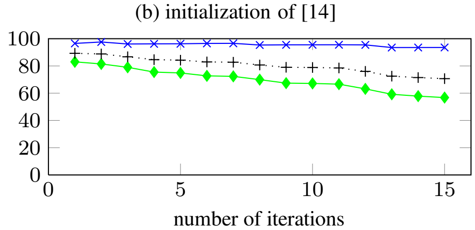
proposed initialization. In terms of F- measure, 82.48% and 87.05% after the first iteration respectively. Furthermore, it



(a) 10% frame initialization

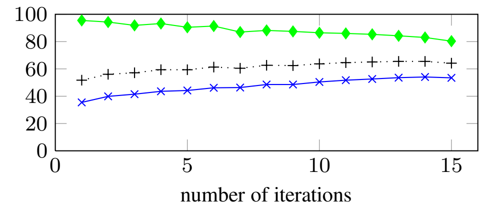


(b) initialization of [14]

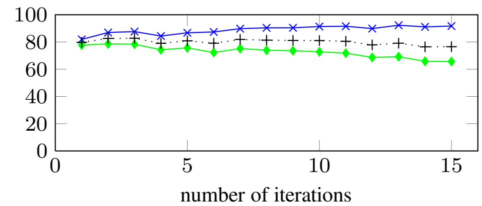


(c) proposed initialization

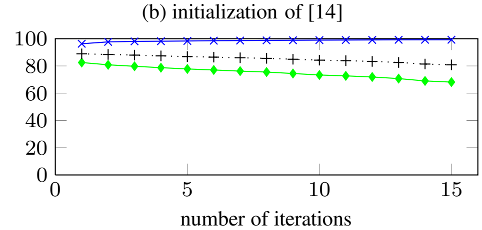
Fig. 4: Average precision ('$\times$'), average recall ('$\diamondsuit$') and F-measure ('$+$') of the image set when MorphCut is used.



(a) 10% frame initialization



(b) initialization of [14]



(c) proposed initialization

Fig. 5: Average precision ('$\times$'), average recall ('$\diamondsuit$') and F-measure ('$+$') of the image set when the extended MorphCut is used.

shows an irregular behavior while the proposed initialization corroborates the fast convergence of GrabCut.

The behavior of both, precision and recall, when MorphCut is applied (see Fig. 4), is very similar for all initializations. Precision tends to increase while recall decreases. Therefore a proper initialization can make the difference if both values are very high after the first iterations. This is the case of the proposed initialization. It achieves the best result after the first iteration, 83.03% according to the F-measure, and deteriorates along the rest. A similar development is shown in Fig. 4b, whose degradation is slower. However, precision and recall are significantly worse than the values obtained by the proposed initialization during the first iterations. As expected, the frame initialization leads to the worst results again. Nevertheless, MorphCut allows it to achieve better results than GrabCut after some iterations.

The quality development along iterations between GrabCut and the extended MorphCut are very similar, according to Fig. 5. Slightly better results are obtained for all initializations in this case though. For example, the proposed initialization achieves the best F- measure, 88.87%, after the first iteration. The recall degradation is very slow and the precision is specially high along iterations. On the other hand, the irregular development of Fig. 5b no longer leads to the best result after the first iteration.

4. CONCLUSIONS

An automatic initialization for graph segmentation techniques employed in high resolution naval images is presented. The desired refinement of our initialization is confirmed and the quality of the final results is substantially high, specially for an automatic approach in such unfavorable conditions. Thus, this shows that the suggested sequence of techniques is a suitable combination. A comparative study between the proposed initialization and earlier approaches is performed as well. This study proves that the precision and recall of our results are considerable higher when they are initialized as we propose regardless the graph segmentation. Furthermore, it reveals that for every graph segmentation the best results of the proposed initialization are obtained after only the first iteration. This not only facilitates the selection of a stopping criteria, but also increases the performance significantly.

5. REFERENCES

[1] H. Li and X. Wang, "Automatic recognition of ship types from infrared images using support vector machines," in *Computer Science and Software Engineering, 2008 International Conference on*, 2008, vol. 6, pp. 483–486.

[2] P. J. Withagen, K. Schutte, A. M. Vossepoel, and M. G. J. Breuers, "Automatic classification of ships from infrared (flir) images," 1999, vol. 3720, pp. 180–187.

[3] S. Musman, D. Kerr, and C. Bachmann, "Automatic recognition of ISAR ship images ," *Transactions on Aerospace and Electronic Systems*, vol. 32, no. 4, pp. 1392 – 1404, 1996.

[4] C. Zhu, H. Zhou, R. Wang, and J. Guo, "A novel hierarchical method of ship detection from spaceborne optical image based on shape and texture features," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 48, no. 9, pp. 3446–3456, 2010.

[5] J. Xu, K. Fu, and X. Sun, "An invariant Generalized Hough Transform based method of inshore ships detection," IEEE, 2011.

[6] L. Ma, J. Guo, Y. Wang, Y. Tian, and Y. Yang, "Ship detection by salient convex boundaries," in *3rd International Congress on Image and Signal Processing (CISP2010)*. 2010, vol. 1, pp. 202–205, IEEE.

[7] M. M. Menon, E. R. Boudreau, and P. J. Kolodzy, "An automatic ship classification system for isar imagery," in *Proc. of Applications and Science of Artificial Neural Networks, SPIE*, 1995, pp. 373–388.

[8] A. O. Knapskog, "Classification of ships in terrasar-x images based on 3d models and silhouette matching," in *Synthetic Aperture Radar (EUSAR), 2010 8th European Conference on*, 2010, pp. 1–4.

[9] C. Kao, S. Hsieh, and C. Peng, "Study of feature-based image capturing and recognition algorithm," in *International Conference on Control, Automation and Systems (ICCAS)*. 2010, pp. 1855 – 1861, IEEE.

[10] I. Camino Garcia and U. Zölzer, "A fully automatic ship localization algorithm for complex backgrounds," in *International Conference on Image Processing, Computer Vision and Pattern Recognition (IPCV)*, 2013.

[11] A.O. Ok, C. Senaras, and B. Yuksel, "Automated detection of arbitrarily shaped buildings in complex environments from monocular vhr optical satellite imagery," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 51, no. 3, pp. 1701–1717, 2013.

[12] A. Hernandez, M. Reyes, S. Escalera, and P. Radeva, "Spatio-temporal grabcut human segmentation for face and pose recovery," in *Computer Vision and Pattern Recognition Workshops (CVPRW), 2010 IEEE Computer Society Conference on*, 2010, pp. 33–40.

[13] M. Wang, L. Shen, and Y. Yuan, "Automatic foreground extraction of clothing images based on grabcut in massive images," in *Information Science and Technology (ICIST), 2012 International Conference on*, 2012, pp. 238–242.

[14] B. Holm, *Autonome Konturdetektion von Schiffen in Maritimem Bildmaterial*, Ph.D. thesis, Helmut-Schmidt-Universität, 2013.

[15] C. Rother, V. Kolmogorov, and A. Blake, ""grabcut – interactive foreground extraction using iterated graph cuts," *ACM TRANSACTIONS ON GRAPHICS*, vol. 23, pp. 309–314, 2004.

[16] O. Rusch, C. Ruwwe, and U. Zölzer, "Image Segmentation in Naval Ship Images," in *11. Workshop Farbbildverarbeitung der GFaI*, 2005.

[17] R. O. Duda and P. E. Hart, "Use of the Hough transformation to detect lines and curves in pictures," *Comm. ACM*, vol. 15, pp. 11–15, 1972.

[18] R. L. Graham, "An efficient algorithm for determining the convex hull of a finite planar set," *Inf. Process. Lett.*, vol. 1, no. 4, pp. 132–133, 1972.

[19] C. Ruwwe, *Volumetrische 3D-Modellierung aus unkalibrierten Freihandaufnahmen*, Ph.D. thesis, Helmut-Schmidt-Universität, 2008.