Analysis of Saliency Object Detection Algorithms for Search and Rescue Operations

Conference Paper · September 2016

DOI: 10.1109/SOFTCOM.2016.7772118

3 authors:


Sven Gotovac
University of Split, Faculty of electrical engineering, mechanical engineering and …
247 PUBLICATIONS **13,954** CITATIONS


Vladan Papić
University of Split
72 PUBLICATIONS **375** CITATIONS


Zeljko Marusic
University of Mostar
10 PUBLICATIONS **21** CITATIONS

Analysis of Saliency Object Detection Algorithms for Search and Rescue Operations

Sven Gotovac
University of Split
Split, Croatia
gotovac@fesb.hr

Vladan Papić
University of Split
Split, Croatia
vpapic@fesb.hr

Željko Marušić
University of Mostar
Mostar, BiH
zeljkomarusic01@gmail.com

Abstract – **The aim of saliency object detection algorithms is to find objects in image or video which draw attention of humans at the first sight. This very popular topic in robotics and computer vision research is useful in various areas and applications like object segmentation, adaptive compression, object recognition, visual surveillance and so on. In this paper, we will explore the possibilities of using these algorithms on the problem of detection of objects for Search and Rescue operations (SAR) in UAV images.**

Keywords - Saliency detection, saliency map, visual attention, UAV, Search and Rescue, CUDA, GPU.

I. INTRODUCTION

Visual saliency is an intriguing phenomenon observed in biological neural systems which received extensive attention by both psychologists and computer vision researchers [1] [2] [3]. It is based on the visual attention, human's ability to quickly locate the most important parts of the scene. Visual saliency is the perceptual quality that makes an object, person, or pixel stand out relative to its neighbors and thus captures our attention [4]. Therefore, the saliency object detection algorithms attempt to locate dominant, prominent or interesting objects in an image, objects on which humans may also pay more attention.

Generally speaking, there are two different processes that influence visual saliency, one is top-down visual attention model, which uses high-level semantic features and knowledge-driven to compute visual saliency. The other is a bottom-up visual attention model, which is data driven and it relies on image features.

Automatic extraction of high level information is hard and sometimes impossible because it doesn't exist in the particular image at all. Opposite to that, low level features like contrast, color, orientation are always available so in this research we focus on bottom-up approach.

Search and rescue operations can be very complex, expensive and can include large number of people. Additionally, modern search and rescue operations use sniffer dogs and various other equipment such as helicopters and drones equipped with thermal and optical cameras. The search is often carried out on the wide open spaces, mountain regions, oceans, desert areas. For example, such search operations are loss of Steve Fossett's [5] whose small plane crashed into a mountain range of the Rocky Mountains in the US (2007), and search for the remains of Air France Flight 447 (2009) [6]. For the latter search, 5.000 volunteers who analyzed thousands of digital satellite images covering several hundred square kilometers of ground surface were used. These two cases represent classic scenery for automatic analysis of images collected from satellites and aerial photography through the different aircrafts.

In addition to the search of the open space, terrain covered with small plants, shrubs, underbrush, rocky ground or slightly forested area are suitable for aerial recording using optical camera as well. This description represents typical vegetation and terrain in Mediterranean and sub-Mediterranean areas as in Dalmatia (Croatia) and lower Herzegovina (Bosnia and Herzegovina).

The main contribution of this paper is the analysis of salient object detection algorithms on images captured from the UAV (Unmanned Aerial Vehicles) aircraft for the purpose of Search and Rescue operations. We investigate the behavior of salient object detection algorithms for images in natural conditions on relative uniform grassy terrain and on complex terrain. For this kind of application, objects of interest are not only lost persons but cars and abandoned things (backpack, jacket, etc) as well – everything that may serve as clues.

Furthermore, we investigated the possibilities for parallelization of the most promising of analyzed algorithms (based on Wavelet Transform) on GPU with CUDA.

The paper is organized as follows: Section 2 is the description and review of related work. In section 3 we present the script for the saliency detection and in section 4 we present comparative analysis of selected saliency object detection algorithms. In section 5, we discuss performance considerations. Concluding remarks and future work are discussed in section 6.

II. RELATED WORK

One of the first biological inspiring models for simulating visual attention was created by Itti, Koch and Niebur [2]. It was designed to simulate the mechanisms of human visual attention and has its roots in psychological theories about human attention. The algorithm obtains the saliency map based on the intensity, color, and orientation conspicuity maps. These conspicuity maps are attained by a cross-scale addition of feature maps, while the feature maps capture the center-surround differences between various Gaussian pyramid and oriented pyramid scale.

Since then, this field of computer vision has seen tremendous interest in the scientific community and many proposed models and systems have been formed with many referent bases for evaluation. Most comprehensive review paper [7] for detection of salient objects and closely related fields (visual attention, fixation prediction, saliency segmentation and object proposal) indicated more than 80 different academic papers from year 1999 to 2014 which confirms the great interest of the academic community in this important area [8-12].

In paper [13] Zhai and Shah developed model for detection of salient objects in video sequences. Their model is composed of temporal and spatial mode. We are interested only in spatial attention model that is used on still images. Color statistics of the images are used to reveal the color contrast information in the scene. Based on color contrast they managed to construct a hierarchical representation of the saliency at the pixel level. Given the pixel-level saliency map, attended points are detected by finding the pixels with the local maxima saliency values. The region-level attention is constructed based upon the attended points. This spatial model is one of the fastest salient detection/ attention models in field.

In paper [14] Hou and Zhang proposed a method whose principle is based on the spectral domain and information theory. Efficient coding decomposes information from the image into innovation part and the redundant part which is already known. This known part is necessary to suppress by the coding system. Innovation part of the image is called spectral residual and calculated as a difference between logarithmised Fourier transformation of image and generalized shape Fourier. With Inverse Fourier Transform the spectral residual is converted to spatial domain, where it is used to construct saliency map.

Achanta, Hemami, Estrada and Susstrunk in paper [4] managed to achieve the following requirements: emphasizing largest salient objects; establishing well-defined boundaries of salient objects; disregarding high frequencies arising from texture, noise and blocking artifacts; efficiently output full resolution saliency maps. These requirements are achieved in frequency domain by filtering low, and high frequency values. They used DoG as a simple band-pass filter.

In paper [15], Imamoglu, Fang used the multi-scale wavelet transformation to create features and feature maps which represent the contrast or center-surround difference, taking both local and global factors into account. The wavelet decomposition has the advantage in extracting oriented details (horizontal, vertical and diagonal) in the multi-scale perspective, and enables high spatial resolution with higher frequency components and low spatial resolution with lower frequency components without information loss in details during the decomposition process. Proposed model creates local saliency map s_L using pixel-level combination of feature maps generated with inverse multilevel wavelet transformation. Additionally global saliency computation s_G is generated using probability density function (PDF) with a normal distribution.

Surprisingly, there are no numerous articles or literature about salient detection of objects on UAV images or detecting objects for Search and Rescue missions. Most of the researched articles are related with the Unmanned Aerial Vehicles (UAV), their control and terrain mapping.

In paper [16] Breckon proposed multi-stage salient detection system which combines low-level contrast features, mean-shift segmentation with additional histogram information and multichannel edge features gathered over several feature maps. Their work is primarily based on the work of Liu, Gleicher [17] and Itti, Koch [2]. Due to small number of test images they did not provide any quantitative analysis of model. A qualitative result suggests very good accuracy of proposed model. After time performance analysis of proposed model we concluded that some of the phases are quite time-consuming and unpractical on large size images.

In paper [18] Turić, Dujmić and Papić used two-stage mean shift segmentation in detection of artificial objects in aerial images for SAR operations. Although this novel method didn't use saliency model approach it is closely related. In first stage high-resolution image (2560x1920 pixels) is divided into smaller sub-images and then clustered using mean shift algorithm. Information about obtained cluster centers is transferred to the second stage where this information about cluster centers from all sub-images is used with same clustering method. In a decision-making module all clusters are evaluated and image segments that had high possibility of presenting the artificial material or the object are selected. With this approach on 22 aerial images they managed to achieve high recall (86%) while keeping the number of false alarms (precision) on acceptable level (56%). Average processing time was 270 seconds per image on Pentium Dual Core 1.86 GHz.

III. THE SCRIPT FOR THE SALIENCY OBJECT DETECTION IN SAR

Search and Rescue (SAR) valiantly strives to locate the subject alive in the shortest possible amount of time. It begins with a call for help and ends with the lost person being found, hopefully alive but possibly dead. Between the initial report and the result there is the search [19].

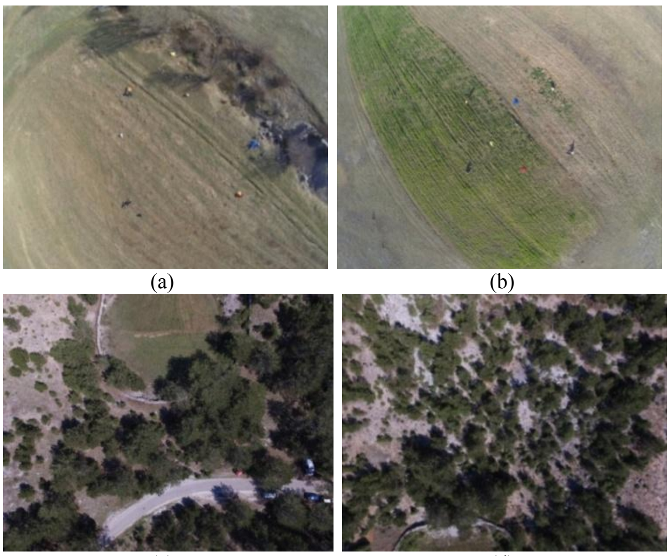
Prior search and rescue operation planners must take into account different kind of variables that can lead to a successful search. Some of these variables are: behavioral research of missing persons; availability of modern tools (helicopters, UAV, thermal cameras); a variety of information about a missing person (age, gender, clothing); information about where the person was last seen; psychological profile of the missing person; the number of available people for search operation ; the kind and type of terrain; the probability of finding a person on certain sections of the terrain; method of the deployment of teams over the area; meteorological and weather conditions; and much other information.

In the areas that are eligible for the use of drones, exact description of the environment allows us to precisely determine the type of images in which detection is performed for search and rescue UAV missions. For instance the

knowledge that the environment is generally uniformed, defines the background image as a largely invariant with respect to the SAR mission (operations on sea and oceans or grassy terrain on land). More realistic situations are when images are not uniform and contain many details in environment (general land missions).
The proposed algorithms should be able to detect salient and prominent objects of various sizes within the scene. Nevertheless, models should be robust and invariant on the size of prominent objects in the scene since the UAV platform can be placed at different altitudes during the flights. Altitude of UAV is very important parameter because we can assume dimensions and size of the target objects with this parameter.

We are not aware of the existence and availability of the relevant public database of color aerial images from UAV for the purposes of detection of objects for search and rescue operations.

In cooperation with Croatian Mountain Rescue Service (HGSS) search and rescue operation simulation has been organized on island of Brač in Croatia. UAV (DJI Phantom 3) flew over search area three times each time on different flight altitude. A similar simulation was carried out on the site, which is located near Mostar, Bosnia and Herzegovina (Fig.1).



(a) (b) (c) (d)

Fig. 1. Examples of images obtained from (UAV) aerial shooting. Relatively uniform pictures taken at location near Mostar (a,b), Relatively wooded terrain taken at location Brač (c,d)

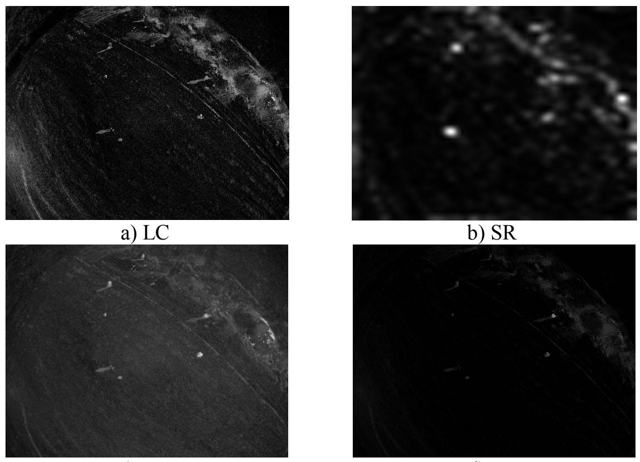
IV. COMPARITIVE ANALYSIS OF MODELS

The largest part of early computer vision image saliency detection work considers high quality and high resolution color images with minimal noise where the salient object is the main subject of the image [16]. These kinds of images are provided in MSRA-A and MSRA-B datasets [20]. Since more models have been proposed in the literature, more datasets have been introduced to further challenge saliency detection models. Recent attempts have been made to collect datasets with multiple objects and complex/cluttered backgrounds as in ImgSal [21]. Resolution of images in these databases are relatively modest (from 400x300 to 1024x768 pixels) in comparison to images that we get from aerial UAV recordings for search and rescue operations (4000x3000 pixels). Higher resolution of the images acquired by digital camera provides significant benefits for positive detection rate of the relatively small objects and allows larger area (terrain) to be captured in a single image (photos taken from higher altitude). Due to the robustness and versatility, GoPro camera is a popular choice on UAV aircraft. On custom built UAV we used the GoPro 3 Black Edition camera that can capture wide FOV (Field of View) still color images with resolution of 4000 pixels of width and 3000 pixels of height (12Mpx). Also we used Sony Exmor 12Mpx sensor in camera embedded in DJI Phantom 3 Professional Quadcopter.

For the analysis we chose 4 popular saliency object detection algorithms: FT (Frequency-tuned Salient Region Detection) [4] LC (Visual attention detection in video sequences using spatiotemporal cues) [13] SR (Saliency detection: A spectral residual approach) [14] WT (A Saliency Detection Model Using Low-Level Features Based on Wavelet Transform) [15]

A. Qualitative analysis

We evaluated the performance of saliency detection algorithms qualitatively, in particular with respect to accuracy and recall. We essentially compared the saliency map to the original image by using a simple algorithm to determine an object map based on the saliency map (Fig. 2). Our analysis was first carried out on the 10 images which contain target objects on a modest uniformed environment. After analysis on this type of images we carried out analysis on 10 images which are more complex. On most images there are multiple objects of interest.



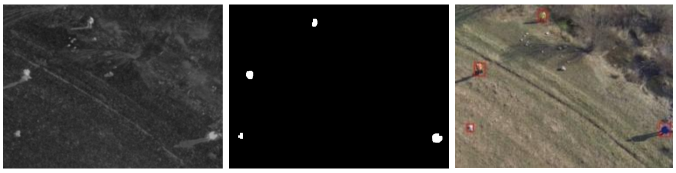
a) LC b) SR c) WT d) FT

Fig. 2. Saliency maps (a), (b), (c), (d) for relatively uniform color natural image in Fig. 1a

After executing each algorithm, we implemented the following phases on produced saliency map to extract and segment objects and to generate object map (Fig. 3):

1. Segmentation with simple threshold

2. Filtering in regards to size using connected component analyses
3. Reading contours and marking salient objects on original images
4. Evaluating against ground truth binary map



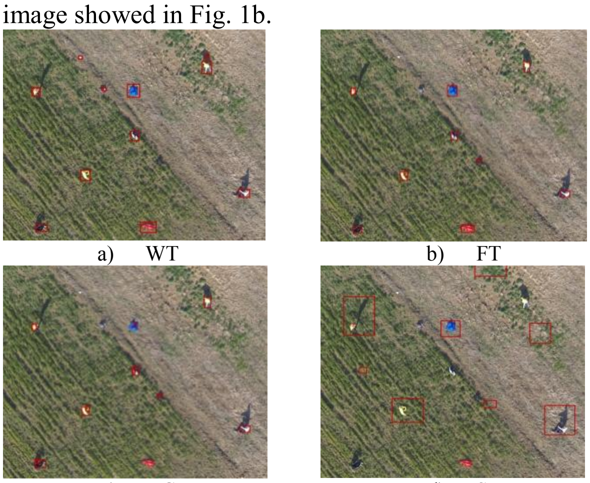
(a) (b) (c)

Fig. 3. Detection phases: (a) saliency map; (b) generating object map; (c) marking salient objects

In this simple approach, binarization of saliency map with manually selected threshold value is one of the most important operations that affect the quality of detection. For every tested model, we used threshold value that segment most of the target objects with minimal number of remaining objects that could be characterized as false alarms. After binarization further analysis had been done using connected component labeling. This step is crucial if we want to remove those objects that are considered as too large or too small. Parameters of approximate dimensions and size of objects depends largely on characteristics of optical camera on UAV and on altitude of UAV platform. After generation of object map we can easily make bounding box of each object (Fig. 3c).

In qualitative analysis we used visual inspection to determine the quality of selected algorithms for the detection of salient objects in relation to ground truth. It should be noted that partial object detection (for instance if algorithm detects only shirt on human) is evaluated as correct hit. Similarly if algorithm achieves multiple detection of one same object we evaluated as one correct hit.

In Fig. 4 we illustrated objects with bounding boxes on parts of image showed in Fig. 1b.



a) WT b) FT c) LC d) SR

Fig. 4.Middle part of image showed in Fig 1b.

It can be noted on Fig. 4 that WT correctly identified all 9 objects (humans, some jacket and trousers). FT correctly identified 8 objects; LC identified 7 objects and SR only 4 objects. Generally, on tested image set, SR model was not very accurate in detection of multiple salient object or small size objects. LC model demonstrated very good accuracy but surprisingly there was a problem with red shade objects that model regularly had been unable to detect. Model FT and WT showed very good accuracy so these two models are evaluated on more complex or heterogeneous set of images like shown in Fig. 1 c and d.

(a) (b)

Fig. 5. a) WT on right part of image showed in Fig 1c, b) FT on right part of image showed in Fig 1c

In Fig. 5 a) and b) we were interested in finding parked vehicles. Although model FT and WT on this particular image managed to detect all objects, WT model had fewer false alarms. However, in very challenging heterogeneous images where the requested objects are very small WT model demonstrated great potential. In Fig. 6 we can see part of terrain where model FT failed to detect objects while WT model correctly identified all target objects.

Fig. 6. Left part of image showed in Fig 1d.

Some of the objects are almost perfectly integrated into the environment (the man in light-blue clothing hiding in the bush at bottom in Fig. 6). Bounding box of the object is only 25 * 20 pixels. On some images WT model managed to detect objects whose bounding box was only 15x10 pixels.

B. Quantitative analysis

In order to evaluate saliency models in SAR operations in a quantitative way, a Precision and Recall measure is adopted in this paper. Precision and Recall are defined as:

$$Precision = \frac{TP}{TP + FP} \quad (1)$$

$$Recall = \frac{TP}{TP + FN} \quad (2)$$

where FP is false positive (objects falsely regarded as salient in respect with ground truth), TP is true positive (correctly identified salient objects) and FN is false negative (salient objects falsely regarded as background).

On 10 images that have relatively simple background we analyzed all 4 models. Images contained 63 target objects.

TABLE 1.Comparison of recall and precision for a set of input images with uniform background

Model	TP	FP	FN	Recall (%)	Precision (%)
WT	57	43	6	90%	57%
SR	19	116	44	30%	7%
FT	56	82	7	89%	41%
LC	52	140	11	83%	27%

From Table 1 we can see that the best result showed WT model (recall 90% and 57% precision) followed by the FT (recall 89%) whose precision was considerably lower (41%). These two models are further evaluated on set of 10 images that have cluttered or complex background. These images contained 39 target objects including 13 cars and 26 humans.

TABLE 2.Comparison of recall and precision for a set of input images with complex background

Model	TP	FP	FN	Recall (%)	Precision (%)
WT	35	44	4	90%	44%
FT	27	628	12	69%	4%

From Table 2 we can see that WT model demonstrated excellent detection results on images with a cluttered environment. Recall remained at the same level while the precision expectedly was slightly lower (more false alarms). On other hand, the detection quality of FT model is significantly degraded.

V. PERFORMANCE CONSIDERATIONS AND OPTIMIZATIONS

Fast and timely discovery of the missing person/s is critical in the search and rescue operations because sometimes, time makes the difference between life and death. So, we conducted runtime analysis of algorithms on a desktop workstation computer containing the Intel Core i7-5930K processor (frequency 3.5 GHz), 32GB of memory. For CUDA computation we used NVIDIA GF110GL Tesla C2050.

The first three algorithms in Table 3 have proved to be very fast on pictures with color resolution of 4000x3000 pixels. These three algorithms are written in C++ and executed in a sequential manner without taking into account any possible optimizations (vectorization of data) or using parallelization. On the other side WT algorithm in its original form was implemented in Matlab so for an equal footing in the comparison we decided to translate it into C ++ code.

Nevertheless, execution speed of WT algorithm proved to be very poor (147.94 seconds). However, because of exceptional quality in the detection of objects we decided to consider the possibilities for parallelization of this model on GPU processor with CUDA programming model [22].

TABLE 3. Performance of tested saliency models in seconds

Algorithm	FT (C++)	LC (C++)	SR (C++)	WT (C++)
Time(s)	4.5 s	1.01s	0.5s	147.94 s

WT model utilize low-level features obtained from the wavelet transform domain. The main phases of this model are illustrated in Fig. 7.

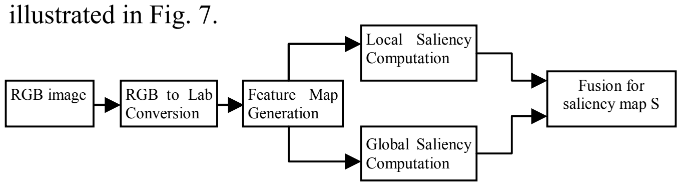


Fig. 7.Architecture of WT model

After evaluation of this model, we realized that bulk (98.6%) of computation is carried out in operations of creating wavelet maps (1.20%), creating local map features (15.60%), and creating global map features (81.80%). Creating local and global map of features depends on wavelet maps. Therefore, in implementation we focused primarily on these three critical phases.

Wavelet maps are created with a multi-level DWT transformation. DWT transformation is traditionally implemented using convolution or FIR filter banks. In relatively recent time a mathematical formulation based on spatial design of wavelets and factorization scheme has been proposed. This new approach is called wavelet transformation based on lifting scheme [23]. It allows in-place computing, reduction of memory dependency and requires half computation operations in regards to convolution based DWT. We implemented DWT and inverse DWT based on lifting scheme in C++ programming language and, as in the original paper, we used “db5” Daubechies wavelet for calculation. Number of levels of DWT in WT saliency model depends on image size and each level is calculated recursively in regards to the previous level. For each wavelet calculation, the image dimensions are scaled by half. Thus, for example, in image with resolution 4000x3000 pixels we get 12 levels per channel and each level represents one wavelet feature map. These wavelet maps are bases for local and global saliency map. So, for every particular level or wavelet map we had to do an inverse DWT transformation to the initial level. In this way we got 12 full resolution maps for every channel (3 channels – 36 maps). Combining all these maps on pixel level we get one local saliency map. For calculating global saliency all maps are first concatenated and then on entire feature set we compute the likelihood of the features at a given location (x, y). This can be defined by the probability density function (PDF) with a normal distribution.

For matrix calculus in global saliency computation we used C++ Armadillo library that supports BLAS, LaPACK standardized tools. Similarly, for CUDA we use cuBLAS library.

In Table 4 we can see summary of our implementation WT algorithm in C++ and CUDA for image with resolution of 4000x3000 pixels.

TABLE 4.Time performance of model for images of resolution 12 Mpx

Phases	**C++**	**CUDA**	**Speedup**
Wavelet Feature Maps	10.94 s	0.29 s	37.72
Local Saliency Map	109.38 s	3.21 s	34.07
Global Saliency Map	25.55 s	1.88 s	13.56
Overall	**145.87 s**	**5.38 s**	**27.09**

In Table 4 we can see that for image of resolution 4000x3000 pixels execution time of CUDA implementation is 5.38 seconds. CUDA implementation demonstrates much better results and speed ups. Speed up of 27.9 times in comparison with C++ implementation is rather significant.

This runtime is similar to the FT model which is very promising because it enables execution within a reasonable amount of time or almost real-time performance.

VI. CONCLUSION AND FUTURE WORK

In this paper, we analyzed algorithms for the detection of salient objects in the natural air images for the purpose of search and rescue operations. It is rather difficult and complex problem in which computer-assisted suggestions of locations of the target objects would be of great help. Inspecting a large number of high-resolution images in seeking for small objects and details can be a very tedious task and as a result there may occur many errors. “Looking and not seeing” is a well-known problem in human search. That problem will be, at least partially, avoided in an image processing system.

After a qualitative analysis of several models on images collected from drones, the most successful one proved to be the algorithm WT described in [15] . However, from the point of performance speed proved to be inadequate for natural high resolution images. Fast and timely discovery of the missing person/s is a critical item in the search and rescue operations because it sometimes makes the difference between life and death. Therefore, we examined the prospects and possibilities for parallelization on hybrid architectures using the graphics processor as a coprocessor to the main processor. It turned out that the most important parts of the algorithm were very suitable for parallelization on GPUs with CUDA programming model and we managed to achieve speedup of execution by 27 times in comparison to sequential implementation mode.

In future work, we will investigate further optimization steps of this algorithm, since we did not use all potential prospects of optimization. For instance we didn’t use fast shared memory or registers in GPUs. Also we didn’t fully exploit the potential of standard multi-core processors. Also we will explore advantages of some other models of saliency detection of objects and try to make a model that is adapted and specialized only for this purpose. For instance the authors in [24] managed to achieve state-of-the-art performance unifying and integrating the simpler saliency models.

REFERENCES

[1] N Bruce and J Tsotsos, "Saliency based on information maximization," in *Advances in Neural Information Processing Systems* , 2005.

[2] Itti L., Koch C., and Niebur E., "A model of saliency-based visual attention for rapid scene analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 11, pp. 1254–1259, November 1998.

[3] Yantis S., "How visual salience wins the battle for awareness," *Nature neuroscience*, vol. 8, no. 8, pp. 975–977, 2005.

[4] R. Achanta, S. Hemami, F. Estrada, and S. Susstrunk, "Frequency-tuned salient region detection," in *CVPR*, 2009.

[5] Geis S. (2007, September) Rescue Crews Find No Sign Of Missing Adventurer. [Online]. http://www.washingtonpost.com/wp-dyn/content/article/2007/09/05/AR2007090502116.html

[6] Hradecky S. (2009, June) Crash: Air France A332 over Atlantic on Jun 1st 2009, aircraft impacted ocean. [Online]. [Online]. http://avherald.com/h?article=41a81ef1/0022&opt=0

[7] A. Borji, M–M. Cheng, H. Jiang, and J. Li, "Salient Object Detection: A Survey," *Transactions on Pattern Analysis and Machine Intelligence*, 2014.

[8] R. Achanta, F. Estrada, P. Wils, and S. Susstrunk, "Salient region detection and segmentation," *Comp. Vis. Sys.*, 2008.

[9] R. Margolin, A. Tal, and L. Zelnik-Manor, "What makes a patch distinct?," in *Computer Vision and Pattern Recognition (CVPR)*, 2013.

[10] C. Yang, L. Zhang, H. Lu, X. Ruan, and M.-H. Yang, "Saliency detection via graph-based manifold ranking," in *in CVPR*, 2013.

[11] X. Hou, J. Harel, and C. Koch, "Image signature: Highlighting sparse salient regions," *IEEE TPAMI*, vol. 34, no. 1, 2012.

[12] M.-M. Cheng, N. J. Mitra, X. Huang, P. H. S. Torr, and S.-M Hu, "Global contrast based salient region detection," *IEEE TPAMI*, 2015.

[13] Y. Zhai and M.Shah, "Visual attention detection in video sequences using spatiotemporal cues," *ACM Multimedia*, pp. 815-824, Oct. 2006.

[14] X. Hou and L. Zhang, "Saliency detection: A spectral residual approach," in *Computer Vision and Pattern Recognition, 2007*, 2007, pp. 1-8.

[15] N. İmamoğlu, W. Lin, and Y. Fang, "A Saliency Detection Model Using Low-Level Features Based on Wavelet Transform," *Multimedia, IEEE Transactions*, vol. 15, no. 1, pp. 96 - 105, January 2013.

[16] J. Sokalski, T.P. Breckon, and I. Cowling, "Automatic Salient Object Detection in UAV Imagery," in *In Proc. 25th International Conference on Unmanned Air Vehicle Systems*, 2010, pp. 11.1-11.12.

[17] F. Liu and M. Gleicher, "Region enhanced scale-invariant saliency detection," in *Proceedings of the ICME 2006*, 2006, pp. 1477–1480.

[18] H. Turić, H. Dujmić, and V. Papić, "Two-Stage Segmentation of Aerial Images for Search and Rescue," *Information Technology and Control*, vol. Vol.39, no. No.2, pp. 139-145, 2010.

[19] R. J. Koester, *Lost Person Behavior*. Charlottesville, Virginia USA: dbS Productions, 2008.

[20] T. Liu, J. Sun, N. Zheng, X. Tang, and H.-Y. Shum, "Learning to detect a salient object," in *in CVPR*, 2007, pp. pp. 1–8.

[21] J. Li, M. Levine, X. An, X. Xu, and H. He, "Visual saliency based on scale-space analysis in the frequency domain," *IEEE TPAMI*, vol. 35, no. 4, pp. 996–1010, 2013.

[22] Nvidia. (2015, September) CUDA C Programming guide. [Online]. http://docs.nvidia.com/cuda/pdf/CUDA_C_Programming_Guide.pdf

[23] I. Daubechies and W. Sweldens, "Factoring Wavelet Transforms into Lifting Scheme," *The J. of Fourier Analysis and Applications*, vol. vol. 4, pp. pp. 247–269, 1998.

[24] X. Han, S. Satoh, D. Nakamura, and K.i Urabe, "Unifying Computational Models for Visual Attention Yields Better Scores than State-of-the-art Models," in *INCF Japan Node International Workshop: Advances in Neuroinformatics 2014*, Wako, JAPAN, 2014.