

An improved watershed segmentation algorithm with thermal markers for multispectral image analysis

C.R. Viau^a, P. Payeur^a, A.-M. Cretu^b

^aSchool of Electrical Engineering and Computer Science, University of Ottawa, 800 King Edward Ave., Ottawa ON Canada; ^bDépartement d'informatique et d'ingénierie, Université du Québec en Outaouais, 101, Saint-Jean-Bosco, Gatineau (Québec) Canada

ABSTRACT

As part of a broader research effort in multispectral image analysis, an improved segmentation algorithm based on the classical *Watershed* concept was developed. A requirement for this research was to develop a segmentation algorithm that could effectively extract objects of interest in both visual and thermal image pairs. The classical *Watershed* algorithm can be enhanced with “markers” identifying clusters of pixels belonging to the same object or to the background. There are several ways to create the markers and the proposed *Watershed with Thermal Markers* allows the user to extract objects of interest from both visual and/or thermal dataset using an initial seed extracted from the thermal image.

Keywords: visual thermal image analysis, multispectral, segmentation, Watershed

1. INTRODUCTION

The process of image segmentation consists of separating foreground objects in an image or scene from their background surroundings. This is often a critical first step in many computer and machine vision applications. Segmented images can subsequently be used to perform feature extraction, object detection and recognition, classification, motion estimation and tracking, as illustrated in Figure 1.

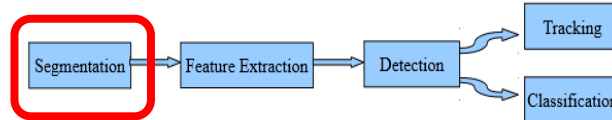


Figure 1. Segmentation as a critical initial step is several machine vision processes

The subject of image segmentation has been thoroughly studied; however the selection and performance of the algorithms are very specific to the intended applications. As an example, in a traffic sign recognition application, segmentation processes may use specific colors and shapes as the main discriminating factors to detect signs in the driver's field of view. The same algorithms would not necessarily be suitable for a Magnetic Resonance Imaging (MRI) processing or satellite-based remote sensing applications.

This paper proposes a novel methodology to create markers for the *Watershed* algorithm using a thermal image to support segmentation of objects in visual imagery. The *Watershed with Thermal Markers* algorithm was developed to support a research¹ investigating how features from visual and thermal imagery could be used jointly to improve the recognition rates of commonly found objects in an office setting. Naturally, the choice of objects was limited to those that radiate thermal energy. A multispectral dataset (visual and thermal) was generated and specific features were extracted to train several *Support Vector Machine* (SVM) classifiers. The SVM's class prediction abilities were evaluated separately on the visual, thermal and the multispectral datasets. For the purpose of this research, several classic segmentation algorithms such as the Basic Threshold, K-means, Contours and Watershed with Distance Transformation Markers were evaluated using representative samples from the training and testing multispectral dataset. Experimentation reported in this paper demonstrated that these algorithms did not provide the segmentation capabilities required to meet the objectives of the research that targets applications where machine vision systems on board mobile robots are trained to detect objects in unknown and challenging environments. As a result, the new segmentation algorithm called *Watershed with Thermal Markers* was developed.

The remainder of this paper is organized as follows: Section 2 presents a brief review of related works in the field of image segmentation. Section 3 discusses the implementation of several classic segmentation algorithms and compares their performance against the multispectral dataset. Section 4 presents the *Watershed with Thermal Markers* algorithm and demonstrates its performance against the multispectral dataset. Finally, Section 5 summarizes the experiments and suggests future work.

2. LITERATURE REVIEW

In a 2008 publication², Zhang *et al.* stated that over 1000 references had been published on the subject of segmentation algorithms and, at the time, over 150 of those were specifically designed for visual images. Some of the more common segmentation algorithms are based on histogram thresholding, feature clustering, edge detection, regions (region growing, region splitting/merging), fuzzy techniques and neural networks (supervised and unsupervised).

One type of algorithm commonly used is the *K-means* algorithm^{3,4} which attempts to segment n data points into k clusters. This segmentation algorithm is an unsupervised learning technique but requires a general understanding of the dataset in order to determine the expected number of clusters. The centers of the k clusters are initialized randomly and eventually converge to final locations. The segmented results from a dataset can vary based on the number of clusters and their initial centers. The final location of the cluster centers is determined when the cluster error function is minimized. When the standard *K-means* algorithm is applied to imagery, it typically requires that the color depth be converted to 8-bit greyscale imagery resulting in the potential loss of clustering information. Some of the principal drawbacks of the standard *K-means* are how to determine the correct number of clusters and the random initialization of the cluster centers. Many authors have focused their research on addressing these two issues and as a result several variations of *adaptive K-means* algorithms were proposed.

Chen *et al.*⁵ proposed an adaptive *K-means* algorithm that detects the number of clusters and their initial centers by analyzing the image's histogram. Using a false-peak mean shift, their proposed algorithm detects the relevant peaks (number of clusters) in the histogram and their location (initial centers). A set of conditions were applied to determine the relevancy of the peaks based on their size and location with respect to other peaks. Similarly to Bhatia's⁶ approach, this technique does not require any prior knowledge of the imagery. However, the algorithm as presented requires an image to be converted into greyscale prior to processing, which results in information loss.

Can *et al.*⁷ proposed to use *Scale Invariant Feature Transform* (SIFT) features with the *Bag-of-Features* (BOF) technique for detection and tracking of sea-surface targets in infrared (IR) and visual band video streams. They used the *K-means* algorithm to generate clusters in the visual band. This manual process involved the input of an operator to select a k value based on the number of ships in the sensor's field of view. The detection and tracking was performed in the individual bands and did not combine the features or information from the different bands to improve the tracking results. The training and testing of the BOF was used to track the target from frame-to-frame as opposed to recognize the various classes of objects.

Another common segmentation algorithm is the *Watershed*⁸ algorithm which was inspired by the field of topography whereby a geographical region is decomposed into peaks and valleys. The classic analogy is when water is dropped over an area and flows downwards to the lowest point which is called a catchment basin. As the water continues to flow, several localized basins (minima) eventually merge and create larger basins leaving only the highest points (maxima) or watershed lines unsubmerged. In image processing, the image topography is defined by the greyscale intensities of the pixel and this concept is used as a segmentation technique. Using classic mathematical morphology operations, the concepts of local minima and maxima, catchment basins and watershed lines can be extracted from digital images. The Watershed segmentation concept has since been exploited and robust algorithms^{9,10} have been developed. The *Watershed* algorithm demonstrates a lot of potential for segmenting complete objects as it considers edges and gradient changes in the imagery unlike other thresholding algorithms that are only concerned with individual pixels.

Gupta and Mukherjee¹¹ proposed a segmentation algorithm based on *Enhanced Fuzzy C-Means* clustering for automatic detection system using thermal imagery. *Fuzzy C-Means* is closely related to *K-means* whereby a data point belongs to a cluster with a certain degree of certitude (fuzzy) instead of belonging to just one cluster (*K-means*). In their algorithm, the optimum number of clusters was estimated using the validity measures Global Silhouette (GS) Index and Separation Index (SI). The GS index was calculated for a large number of clusters (up to 20) and the number of clusters with highest index was chosen as the optimum number. The SI also provided a cluster quality measure. Although not specifically

indicated, this type of implementation likely has a significant processing cost as compared to the standard and *Adaptive K-means*. This type of algorithm could improve segmentation results but it is unclear how effective it could be when segmenting both visual and thermal imagery.

Hasanzadeh and Kasaei¹² proposed a multispectral segmentation method based on size-weighted fuzzy clustering and membership connectedness. This advanced fuzzy clustering technique took into consideration the local and global position of the image pixels as part of the segmentation process. This approach was developed primarily for thermal imagery and used the spectral and spatial content of the image to improve the clustering performance over the standard and adaptive *K-means* and *Fuzzy C-means* algorithms. This proposed algorithm has shown high potential for thermal imagery but it is unclear how it would perform on greyscale visual imagery.

3. ASSESSMENT OF CLASSIC SEGMENTATION ALGORITHMS

In terms of segmentation, numerous implementations of *K-means*, *adaptive K-means* and fuzzy variations have been proposed. The main issue with this type of algorithm is that they cluster pixels based on their pixel intensities and not on their spatial location within the image which makes it difficult to extract complete objects from their background. Conversely, the *Watershed* algorithm shows more potential to extract complete objects because it is based on gradients within the image. The majority of the segmentation algorithms reviewed were specifically for visual imagery and their performance on thermal imagery is unknown. The selected algorithm for this research had to effectively extract the same objects from both visual and thermal image pairs. A new variation of the *Watershed* was developed specifically for this research and is described in Section 4.

Segmentation of the foreground objects from their background is a critical first step for feature extraction and classification of an object. Several classic segmentation algorithms such as the *Basic Threshold*, *K-means*, *Contours* and *Watershed with Distance Transformation Markers* were implemented and tested using representative samples from the training and testing multispectral dataset.

3.1 Basic Threshold

The classic *Basic Threshold* segmentation algorithm was implemented using the *segmentation*¹³ function from the OpenCV library. In this algorithm, a user-defined threshold separates the pixels in an image (visual or thermal) into two groups based on the pixel intensity level. The pixels in the image with intensity levels above the user-defined threshold are assigned a color of white while those below are assigned a color of black. This classic algorithm works on color images but must they must be converted them to 8-bit greyscale prior to segmenting. The conversion to greyscale results in a loss of information as compared to other types of algorithms. Furthermore, the algorithm does not take into consideration the state of adjacent pixels as part of the segmentation process.

3.2 K-means

The *K-means* algorithm was implemented using the *kmean*¹³ function from the OpenCV library. In this implementation, the image (visual or thermal) was first converted to greyscale and blurred to facilitate the clustering of pixels with similar intensities. The blurring function is implemented using the OpenCV *blur* function¹³ which implements a normalized box filter. The algorithm separates the n pixels in the image into k clusters. For the purposes of this experimentation, the k value was set to 2 in order to obtain two classes of objects (foreground objects of interest and background).

3.3 Contours

The *Contours* algorithm was implemented using the *findcontours* and *drawcontours*¹³ functions from the OpenCV library. In this implementation, the image (visual or thermal) was first converted to greyscale and blurred using the OpenCV functions. The blurring function was implemented using the OpenCV *blur* function¹³. A *Canny* edge detection algorithm was then used to identify primary edges in the image prior to the *findcontours*¹³ algorithm that links these edges to highlight the outlines of various connected components in the scene.

3.4 Watershed with Distance Transform

The *Watershed* implementation for this research was based on an example¹⁴ in the open literature whereby the basic watershed algorithm was enhanced with *markers* identifying clusters of pixels belonging to the same object. There are

several ways to create the markers but in this example the *distanceTransform*¹³ function from the OpenCV library was used. The *Distance Transform* works on a binary image and converts each of the white pixels to greyscale value representing the smallest distance to the background (black pixels). An example of the *Distance Transform* operation is illustrated Figure 2.

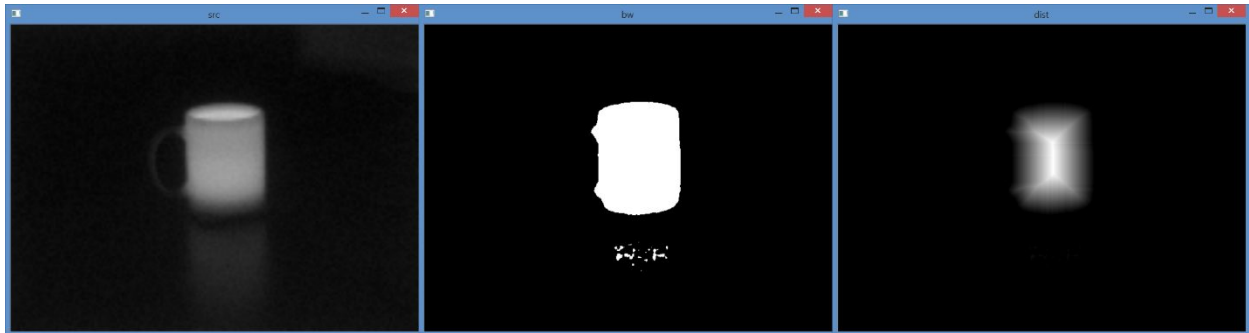


Figure 2. Example of the Distance Transform calculation. Original image (left), binary (center), distance transform (right)

The *Watershed with Distance Transform* implementation consisted of first converting the image to binary using a threshold of 40 (as suggested by the original author), and then applying the *distanceTransform* function to the resulting binary image. A final threshold operation (using a threshold value 127 on a scale of 0 to 255 in order to separate foreground radiating objects from the cooler background) was applied to the output of the *distanceTransform* function to create the markers for the watershed algorithm.

3.5 Performance assessment

A version of the *Basic Threshold*, the *K-means*, the *Contours* and the *Watershed with Distance Transform* algorithms were implemented as part of this research to find a segmentation algorithm suitable for both visual and thermal images. Samples of the segmentation algorithm results are compared and illustrated in Figure 3 to Figure 6.

Figure 3 illustrates a sample image pair (visual image on the top row and thermal image on the bottom row) of a mobile phone. This sample was specifically selected because the visual image represented a dark object on a light background, which should not have been a real challenge for any segmentation algorithm. However, in the thermal image the radiance of the mobile phone was barely greater than its background. It can be observed that in this specific example, the *K-means* algorithm performed better over the other three algorithms in both the visual and thermal spectrum. The *Basic Threshold*, *Contours* and *Watershed with Distance Transform* correctly identified the outline of the object in the visual spectrum but were susceptible to the reflection of the light on the cell phone. In the thermal spectrum, the *Basic Threshold* provided a mediocre representation of the object while the *Contours* and *Watershed with Distance Transform* could not segment any part of the object.

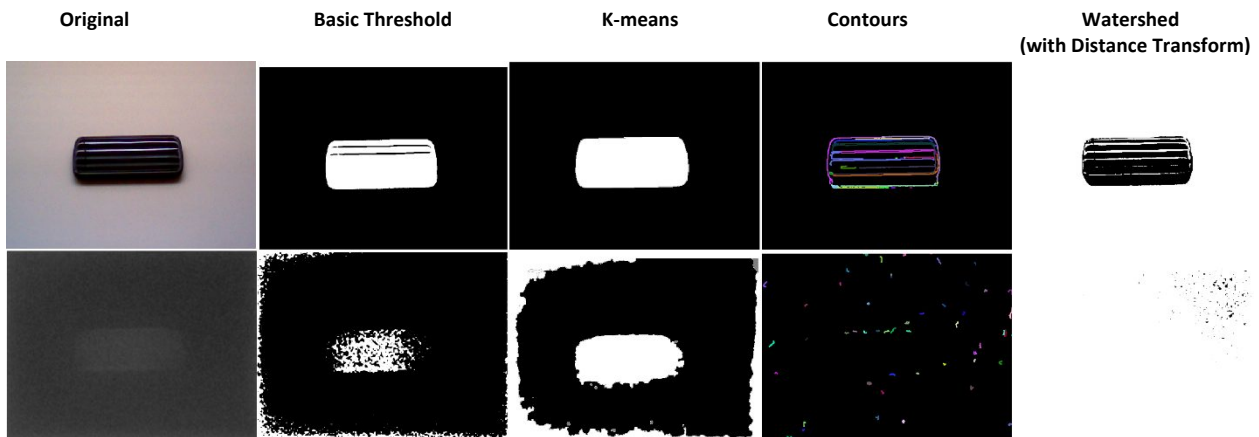


Figure 3. Sample segmentation results (dark object on light background)

Figure 4 illustrates the performance of each segmentation algorithm against a training dataset sample of a portable heater. This sample was specifically chosen to evaluate the segmentation capabilities of the algorithms on a dark object against a light multi-textured background. For the purpose of this research, the segmentation algorithm had to extract the complete outline of the object from its background in the visual and thermal spectrums. The challenges in this sample were the blinds and the other small dark object in the middle left-hand side of the scene. In the visual spectrum, the *K-means* provided the better segmentation of the object as it removed the majority of the blinds in the background and provided nearly a complete filled outline of the object. The other three algorithms all provided a good outline of the object but could not remove the background blinds from the segmentation. In the thermal spectrum, the *Basic Threshold* and the *Watershed with Distance Transform* provided a good representation of the object but the contour of the back of the heater was very grainy and not well defined which could make it difficult to extract dimensions and measurements from this segmented image. The *K-means* provided well defined outlines but only segmented the highly radiating elements of the object. Similarly, the *Contours* algorithm provided well defined outlines but struggled to capture the complete object in the thermal image.

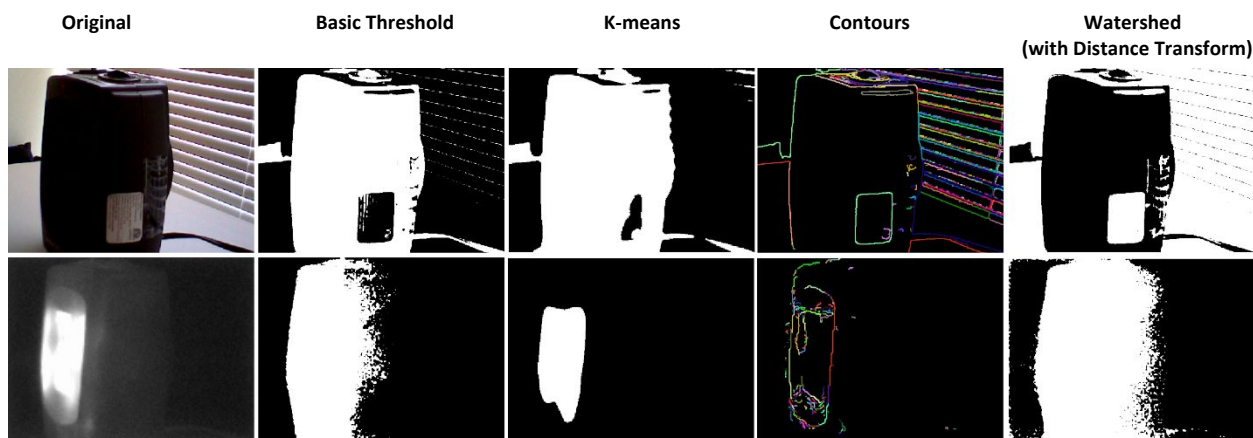


Figure 4. Sample segmentation results (dark object on light multi-textured background)

Figure 5 illustrates a sample dataset of a desk lamp segmented by the various algorithms implemented.

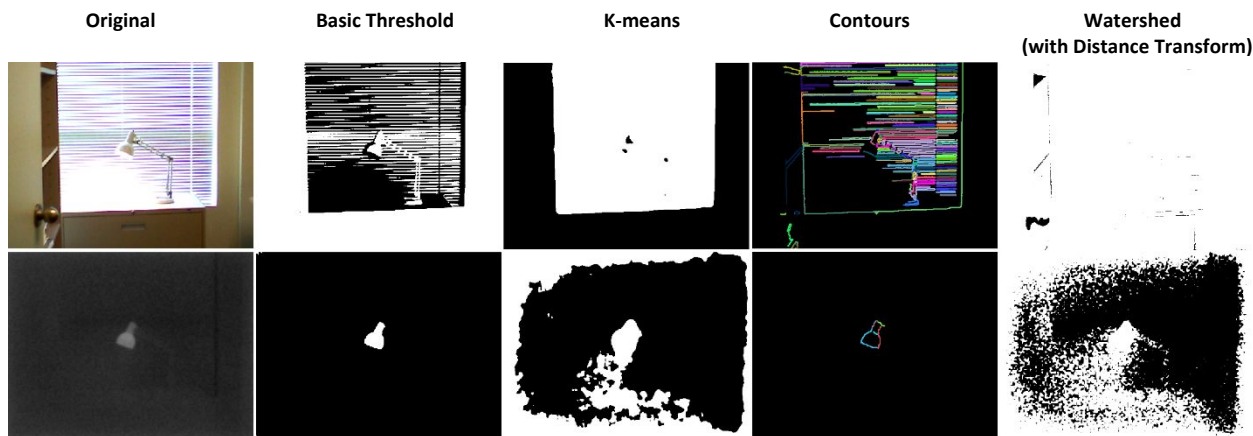


Figure 5. Sample segmentation results (light object on light background)

This sample was specifically selected because it illustrated a light colored object in front of a light multi-textured background. This is in contrast to the previous two examples presented in Figure 3 and Figure 4. It was expected that this type of image would be a greater challenge for the segmentation algorithms. In the visual spectrum, none of the algorithms were able to correctly segment the desk lamp from the background. The best approximation was probably the *Basic Threshold* but this result could not be used to easily extract features because of the large number of clusters in the

image. The *Contours* algorithm provided a similar result again with a large number of *Contours* which would make it difficult to automatically identify the desk lamp. In the thermal image, the *Basic Threshold* and the *Contours* provided the best segmentation results in comparison to the other two algorithms. The shape of the lamp was clearly outlined and all the background clutter was removed. However, none of the algorithms provided adequate results in both the visual and thermal spectrum.

The last segmentation examples are illustrated in Figure 6 and were probably the most challenging for the algorithms. This dataset sample was specifically selected because it illustrated various thermally radiating objects of interest (a coffee cup, a cellphone, a lamp, a heater, and a laptop charger) disseminated in a very cluttered and multi-textured environment. The visual image segmentation results showed that none of the algorithms tested were capable of extracting just objects of interest. A human observer could probably find the objects in the segmented images, but this would likely be difficult for an automated process. Conversely, the thermal image provided a very clear location of each of the objects of interest and all the algorithms were capable of identifying at least 3 of the 5 objects. The *Basic Threshold* and the *Watershed with Distance Transform* likely provided the better results for an automated feature extraction application. The *Contour* algorithm provided a general location of the objects but extracted many additional unnecessary outlines. Once again, none of the tested algorithms provided adequate segmentation capabilities in both the visual and thermal datasets.

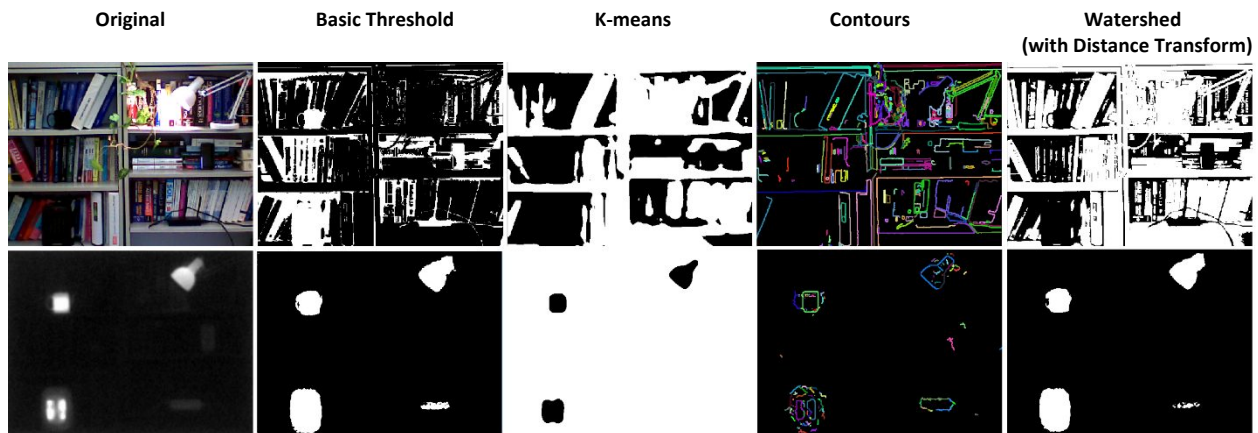


Figure 6. Sample segmentation results (light and dark objects on multi-textured background)

In order to automatically segment objects of interest in both the visual and thermal spectra, an alternate algorithm was developed. The segmentation results presented in Figure 3 to Figure 6 demonstrate that in the visual spectrum, an algorithm based on pixel intensity values only works if the object has similar colors and is presented against a contrasting background. Threshold-based algorithms used on visual imagery do not consider spatial content of the image nor the state of adjacent pixels and as a result were deemed unsuitable. Conversely, in thermal imagery the radiation emitted from a source directly affects its surrounding and consequently, a relationship exists between adjacent pixels with a similar greyscale intensity levels. In thermal imagery, a threshold-based algorithm can effectively segment related pixels from an object simply based on the image's greyscale thermo-related intensity.

The *Watershed* algorithm seemed to provide the most potential segmentation capability in the visual spectrum as this region growing algorithm takes into account the relationship between adjacent pixels. The concept of using *markers* to help the *Watershed* algorithm triggered the idea that perhaps the thermal image, which can be easily segmented, could be used as initial markers to enhance the segmentation in the visual image. This new segmentation algorithm was named *Watershed with Thermal Markers*.

4. NEW WATERSHED WITH THERMAL MARKERS ALGORITHM

The *Watershed with Thermal Markers* algorithm proposed in this research uses the thermal image to produce markers that can be used by the *Watershed* algorithm to segment either the visual or thermal images. The flowchart of the algorithm is presented in Figure 7.

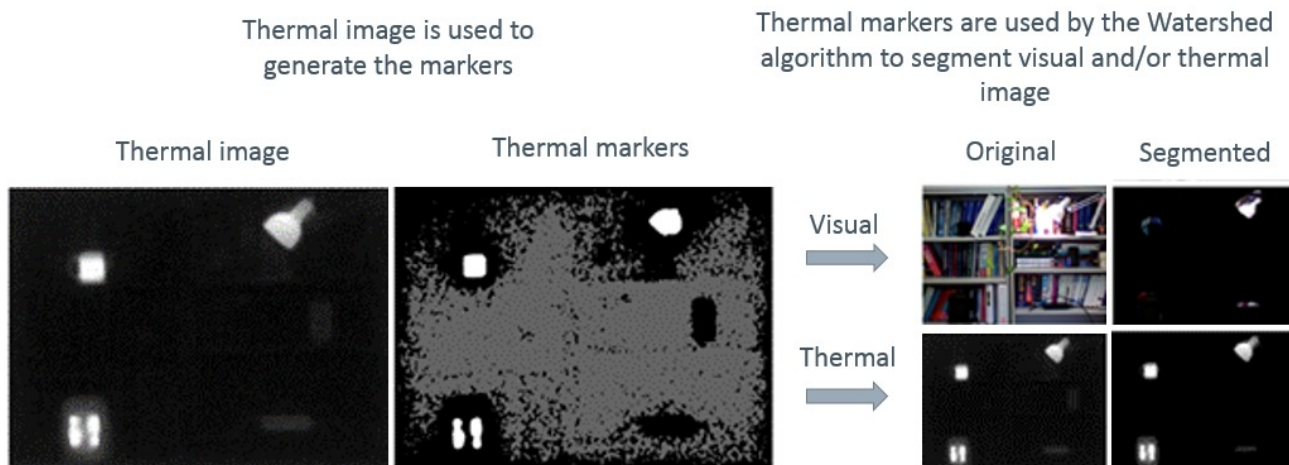


Figure 7. Watershed with Thermal Markers flowchart

The thermal markers are generated from the thermal image by separating the pixels into three greyscale intensity groups based on two user-defined thresholds. The pixels with a greyscale intensity above the upper thresholds are considered to be part of the object of interest in the thermal image and make up the first *marker* (color “white”). The pixels below the lower threshold are considered to be part of the background and make up the second *marker* (color “grey”). The rest of the pixels between the lower and upper thresholds are considered unassigned (color “black”) and could belong to the object(s) of interest or to the background. An example of the markers and unassigned pixels generated from the sample image from Figure 6 are illustrated in Figure 7.

Once the thermal markers are generated, the *Watershed* algorithm can be applied to either the visual or thermal image to complete the segmentation process. Figure 8 illustrates a segmentation example where the background and foreground thresholds are not optimized. On the left side of the figure is the original image from the visual band while the middle image shows the markers generated from the thermal band image. In the middle figure, the white pixels identify the foreground object while the grey pixels identify the background objects. The black pixels have not yet been assigned as either foreground or background objects. When these specific markers are used in conjunction with the *Watershed* algorithm, the resulting segmentation from the visual image is illustrated in the far right image of Figure 8. The segmentation by the *Watershed with Thermal Markers* can be further improved by adjusting one or both thresholds, which results in reducing the number of unassigned pixels (black) as illustrated in Figure 9. The algorithm is currently not automated and requires user interaction to optimize the segmentation results by adjusting the upper and/or lower thresholds.

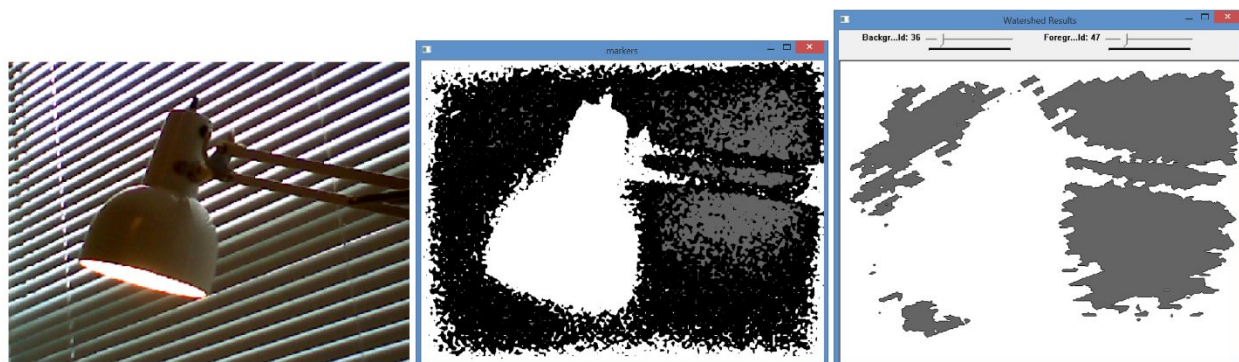


Figure 8. Original visual image (left), thermal markers (middle), and *Watershed with Thermal Markers* segmentation results (right)

The segmentation examples presented earlier in Section 3.5 were reassessed against the *Watershed with Thermal Markers* and illustrated in Figure 10 (visual band) and Figure 11 (thermal band).



Figure 9. Improved segmentation using adjusted thresholds. Original visual image (left), thermal markers (middle) and *Watershed with Thermal Markers* segmentation results (right)

It can be observed from these sample results that the new algorithm can segment the objects of interest in both the visual and thermal datasets. Note that in these samples, the segmented objects are “colored” white while the background is colored a shade of grey to easily identify the object from its background. In the actual implementation, the background is colored black and the objects of interest maintain their greyscale values in order to compute intensity features.

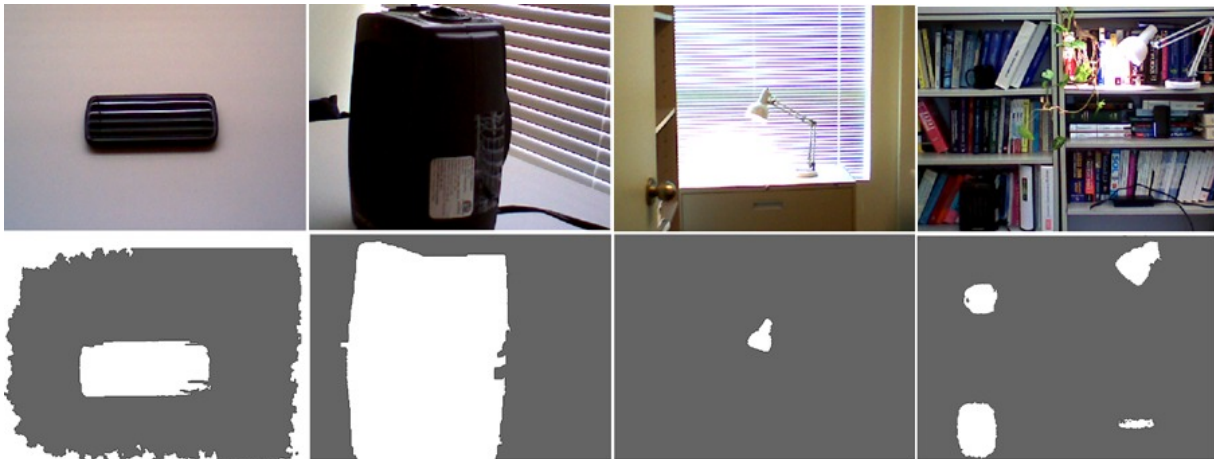


Figure 10 *Watershed with Thermal Markers* segmentation results on various objects from the visual dataset

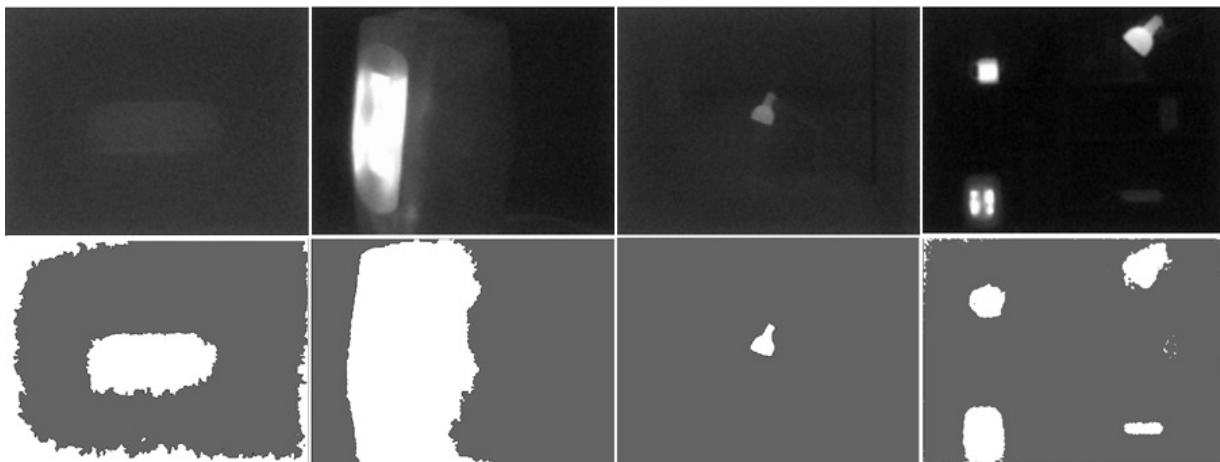


Figure 11. *Watershed with Thermal Markers* segmentation results on various objects from the thermal dataset

The *Watershed with Thermal Markers* algorithm was further evaluated on a variety of other visual and thermal image pairs to assess its performance on other imagery not used for the main research work. Figure 12 demonstrates the algorithm's ability to extract a family pet from its cozy environment. It can be observed that even with a cage in the foreground and among all the busy background, the algorithm can segment the animal of interest by taking advantage of the thermal band to orient the segmentation process in both the visual and thermal spectra.

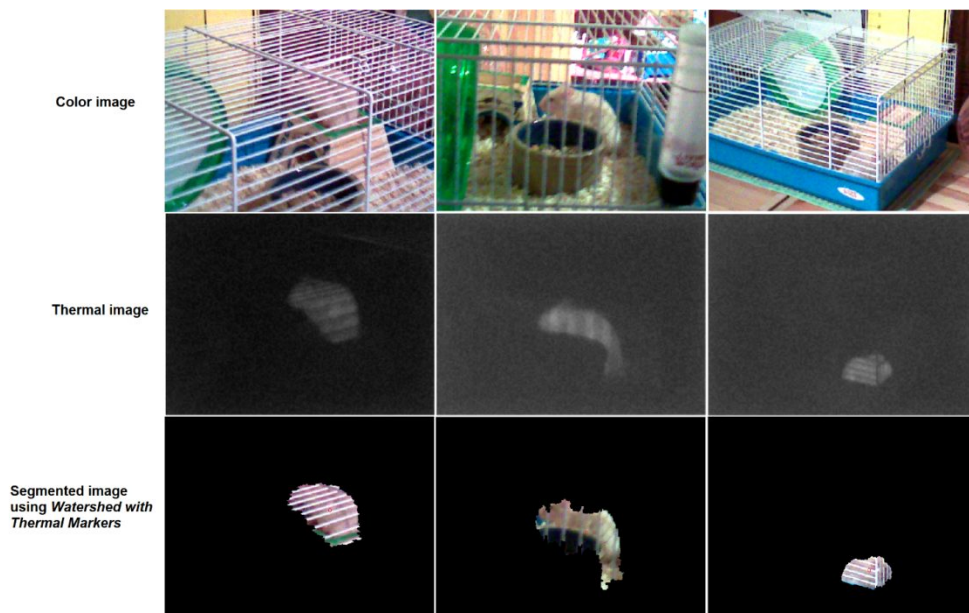


Figure 12. *Watershed with Thermal Markers* segmentation results on an animal (hamster) in a cage

The algorithm was also evaluated on outdoor scenes as illustrated in Figure 13. In this type of imagery, the algorithm proves very efficient to remove the background sky and extract radiating sources from both the visual and thermal imagery.



Figure 13. *Watershed with Thermal Markers* segmentation example to remove sky and backgrounds

5. CONCLUSIONS AND FUTURE WORK

This research work demonstrated that the classical *Watershed* algorithm can be enhanced with "markers" when applied jointly on multispectral images (visual and thermal bands). Those markers can help identifying clusters of pixels belonging to the same object or background. There are several ways to create the markers and in the proposed approach presented in this paper, the markers are generated from the thermal imagery. In the current implementation, the algorithm requires the input of an operator to find the best upper and lower thresholds in order to optimize the performance of the segmentation. An extension of this work could be to investigate the feasibility and subsequently the robustness of an automatic threshold detection and selection in order to fully automate the algorithm.

The proposed *Watershed with Thermal Markers* allows the user to extract objects of interest from both visual and/or thermal dataset using an initial seed extracted from the thermal image. The algorithm was specifically designed to support a research effort in the field of multispectral image analysis and its performance was demonstrated against a multispectral dataset comprised of various thermally radiating objects of interest (a coffee cup, a cellphone, a lamp, a heater, and a laptop charger). The concept of thermal markers could be used to support other segmentation algorithms requiring markers or initial seeds. The proposed *Watershed with Thermal Markers* algorithm could find numerous other applications in medical image analysis, satellite imagery analysis, as well in surveillance and security systems.

REFERENCES

- [1] Viau, C., Payeur, P., Cretu, A.-M., "Multispectral image analysis for object recognition and classification," Proc SPIE 9844 Automatic Target Recognition XXVI, Baltimore, USA (2016).
- [2] Zhang, H., Fritts, J. E., Goldman, S. A., "Image segmentation evaluation: A survey of unsupervised methods," Comput. Vis. Image Underst. 110(2), 260–280 (2008).
- [3] Hartigan, J., [Clustering algorithms], Books On Demand, [S.I.] (1975).
- [4] Hartigan, J. A., Wong, M. A., "Algorithm AS 136: A K-Means Clustering Algorithm," J. R. Stat. Soc. Ser. C Appl. Stat. 28(1), 100–108 (1979).
- [5] Chen, H., Wu, X., Hu, J., "Adaptive K-means clustering algorithm," 15 November 2007, 67882A – 1 to 67882A – 9.
- [6] Bhatia, S. K., "Adaptive K-Means Clustering," Proc. Seventeenth Int. Fla. Artif. Intell. Res. Soc. Conf. Miami, 695–699 (2004).
- [7] Can, T., Karali, A. O., Aytac, T., "Detection and tracking of sea-surface targets in infrared and visual band videos using the bag-of-features technique with scale-invariant feature transform," Appl. Opt. 50(33), 6302–6312 (2011).
- [8] Vincent, L., Soille, P., "Watersheds in digital spaces: an efficient algorithm based on immersion simulations," IEEE Trans. Pattern Anal. Mach. Intell. 13(6), 583–598 (1991).
- [9] Meyer, F., "Color image segmentation," Int. Conf. Image Process. Its Appl. 1992, 303–306 (1992).
- [10] Meyer, F., "Topographic distance and watershed lines," Signal Process. 38(1), 113–125 (1994).
- [11] Gupta, S., Mukherjee, A., "Infrared image segmentation using Enhanced Fuzzy C-means clustering for automatic detection systems," 2011 IEEE Int. Conf. Fuzzy Syst. FUZZ, 944–949 (2011).
- [12] Hasanzadeh, M., Kasaei, S., "A Multispectral Image Segmentation Method Using Size-Weighted Fuzzy Clustering and Membership Connectedness," IEEE Geosci. Remote Sens. Lett. 7(3), 520–524 (2010).
- [13] "Welcome to opencv documentation! — OpenCV 2.4.11.0 documentation.", <<http://docs.opencv.org/2.4.11/>> (21 March 2016).
- [14] "Count and segment overlapping objects with Watershed and Distance Transform | OpenCV Code.", <<http://opencv-code.com/tutorials/count-and-segment-overlapping-objects-with-watershed-and-distance-transform/>> (17 August 2015).