An Approach to Interesting Objects Detection in Low Quality Image Sequences for Fisheries Management

Zuojin Li, Liukui Chen* and Jun Peng
Chongqing University of Science and Technology
College of Electrical and Information Engineering
20 East Road,University City,Chongqing
China
*cqustclk@163.com

Lei Song*
Unitec Institute of Technology
Department of Computing
139 Carrington RD, Mt Albert, Auckland
New Zealand
*lsong@unitec.ac.nz

Abstract: In order to extract interest objects in low quality image sequences from fisheries management, this paper proposes a new significant feature extraction method based on cascade framework. This algorithm involves pre-processing image sequences, clipping interesting areas, extracting SURF features, removing boundary features, and acquiring significant features with interesting objects. We apply our algorithm to fisheries management for counting and matching ships and cars, the proposed method can efficiently detect multiple objects from real-scene video frames with averaged accuracy 91.63%.

Key–Words: SURF Features; Interest Objects; Features Extraction; Multiple Objects Detection; Low Quality Image Processing.

1 Introduction

Fisheries management is essential to the sustainable development of environment, which aims to balance the sustainability of fisheries stocks and the impacts of fishing on the environment with the economic opportunities they offer [1][2][3][4]. With the increasing income, the number of household ships is hiking up in New Zealand year by year. Statistics show that more than 50% households in the country own ships up to the year 2010. Large quantity of ships on the sea is harmful for fisheries resources. Therefore, how to monitor the ships on the sea is a big challenge for fisheries management.

Camera surveillance system has been widely used for monitoring environment of many real world applications, so is fisheries management. However, the surveillance for ships entering or exiting the port is not efficient, because surveillance relies on watch man is not only time consuming but also not accurate due to the status/mood of watch man varies over time. Thus, how to detect interest objects in low quality image sequences from monitoring environment with the help of image processing, computer vision, and machine learning has become a important and difficult issue. In this paper, we presented a new method to automate the process of image analysis and objects detection in real fisheries management scene.

In literature, computer vision technologies has been extensively studied for automating the process of objects detection. It mostly relies on the accuracy of feature extraction, including object searching, segmentation, matching [5][6][7][8][9]. Feature-based object detection is very important in computer vision area. SIFT [10] is one of the best performing detector method through the use of a Difference-of-Gaussian (DoG) detector in combination with a scale-space of the image and the assignment of a main orientation per interest point of objects.The sift achieves features of scaling, rotation and Corner from image data. ASIFT [11] is a fully affine invariant extension of SIFT, which uses an affine map to simulate changes of the camera viewpoint. SURF [12] is a detector combination inspired by SIFT. SURF is more efficient to compute while showing comparable performance and in some car detection also outperforms SIFT. ORB [13] is a detector combination which is constructed for real-time systems and is therefore faster than SIFT and SURF while similar matching performance is achieved with car detection and recognition. BRISK [14] is like ORB a binary detector combination which aims at high-quality description with low computational complexity like SURF, achieving comparable matching performance. However these presented feature extraction methods are effective with clear Scene or high quality image, and some good robustness of experimental results are expressed on single object image, for example only car. Some unexpected features are emerged on processed image that only for our interest objects when these above methods are applied in case of engineering (fisheries management with intelligent video surveillance) scene.

Recently, Stalder et al [15] propose a new cascade confidence filter for image feature extraction and object tracking. This method can narrow down the target search scope by detecting objects features of interest, background statistical features and time consistency. But the dependence of this filter on background frames greatly restricts its application. Moreover, this method can only extract the possible areas of interest objects rather than their significant features.

To mitigate the gap, this paper proposes a brand-new algorithm for extracting significant features of interest objects based on cascade structure from low quality static image sequences. Its not restricted by image sequences and adaptable to images with low signal-to-noise ratio.

This remainder of this paper is organized as follows: Section 2 shows the details of the proposed algorithm. Section 3 presents experimental results and section 4 concludes the paper.

2 Method

2.1 Overview

The method describes the whole process of detection interest objects with extracting significant features from image frames, which can be applied to solve other similar problems. As the environment for fisheries surveillance is usually complex, images acquired are low in signal-to-noise ratio. As we are more interested in ships going to the sea, we define the interest areas in the camera surveillance of fisheries management as the critical zones for exiting the seaport and store their coordinate information in our data documents. Because the invariance principle can well reflect the robustness features of ships of different sizes, in different locations and under different conditions, and in order to accelerate the computing process, this paper adopts SURF operator to extract the invariance features of the object areas. Because of the low signal-to-noise ratio of the images, SURF operator may take fuzzy boundary points as feature information [16], and that is why a lot of feature points are irrelevant. This paper will screen boundary features in the original feature sets and maintain the invariance feature of the inside area. Though features screening can enhance the features of interest objects, their data can not be extracted significantly. In order to find out the significance of the density of feature points in interest objects, this paper adopts OPTICS algorithm to calculate the core-distance between feature points and take the feature set with the shortest core-distance as the significant feature zone. The whole process is shown in Figure 1.

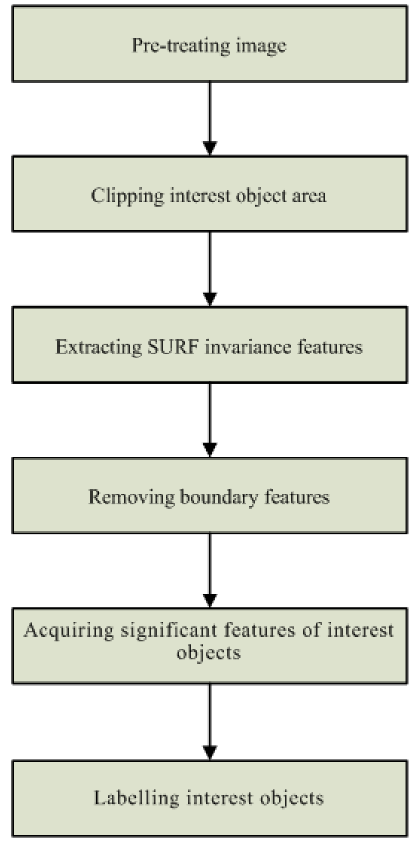


Figure 1: The flowchart of proposed method based on cascade framework.

2.2 Pre-treating images

Histogram equalization is characterized with retaining details and removing noises [17], which is widely applied in image processing. Its theoretical expression is shown below: When supplying a desired histogram hgram, histeq chooses the grayscale transformation T to minimize

$$|c_1(T(k)) - c_0(k)|, \tag{1}$$

where c_0 is the cumulative histogram of A, c_1 is the cumulative sum of hgram for all intensities k. This minimization is subject to the constraints that T must be monotonic and $c_1(T(\alpha))$ cannot overshoot $c_0(\alpha)$ by more than half the distance between the histogram counts at α. Histeq uses the transformation $b = T(\alpha)$ to map the gray levels in X (or the colormap) to their new values.

2.3 Clipping interest object areas

The visible area of images under the fisheries management surveillance is large, where exist many non-interest zones. Based on the monitoring tasks aiming at ships, we determined roads of ships/cars and launching areas of ships as the two interest zones, marked as A and B. The standardized coordinates of Zone A and Zone B are shown in the following matrix:

$$ROI_A = \begin{bmatrix} 85 & 197 & 242 & 341 & 199 & 125 & 85 \\ 201 & 163 & 188 & 222 & 293 & 244 & 201 \end{bmatrix}, \quad (2)$$

$$ROI_B = \begin{bmatrix} 303 & 434 & 601 & \cdots & 297 & 168 & 303 \\ 219 & 242 & 235 & \cdots & 353 & 274 & 219 \end{bmatrix}, \quad (3)$$

In this matrix, the 1st and 2nd rows respectively represent the coordinate X and Y of the image.

2.4 Extracting SURF invariance features

The data of camera surveillance images is huge, this research adopts a quick method of invariance feature extraction. Surf feature extraction is an improved algorithm based on SIFT feature extraction [18][19]. The major process of SURF [20] algorithm is shown as below:

Step 1: The interest points is found with using Hessian matrix for determination of feature location and scale. Given a point $x = (x, y)$ in image I, the Hessian matrix $H(x, \sigma)$ in x at scale σ is defined as follows

$$H(x, \sigma) = \begin{bmatrix} L_{xx}(x, \sigma) & L_{xy}(x, \sigma) \\ L_{xy}(x, \sigma) & L_{yy}(x, \sigma)) \end{bmatrix} \quad (4)$$

where $L_{xx}(x, \sigma)$ is the convolution of the Gaussian second order derivative $\frac{\partial^2}{\partial x^2} g(\sigma)$ with the image I in point **x**, and similarly for $L_{xy}(x, \sigma)$ and $L_{yy}(x, \sigma)$.

Step 2: An integral image is generated by building approximate Gaussian kernel to be a box filter for scale determination.

Step 3: The SURF features is described by using Haar wavelet response $\sum dx$, $\sum |dx|$, $\sum dy$, $\sum |dy|$.

As we can see, Surf is quite different from SIFT in terms of regional maximum value, multi-scale decompositions and choice of statistics in the main direction, and SURF enjoys obvious advantage in computing speed. Therefore, SURF algorithm is quite valuable in engineering application.

2.5 Removing boundary features

For given feature points (x_0, y_0), calculate:

$$Zeros = \sum(image(x_0 + \Delta x, y_0 + \Delta y) \equiv 0, \\ (\Delta x, \Delta y) < \delta), \quad (5)$$

in which $(\Delta x, \Delta y) < \delta$ represents a 10×10 rectangular area, and when $Zeros > 60$, (x_0, y_0) point is the boundary point, which is thus removed.

2.6 Automatic clustering of significant features of interest objects

The step before removes some boundary noise features, but many feature points in non-interest objects still exist within the interest areas, such as noise features, non-complete object features, background features, and interference features with low signal-to-noise ratio. This paper adopts OPTICS algorithm [21], which can automatically cluster feature points of interest objects, and thus improves the significant features extraction for interest objects.

The process of the OPTICS algorithm is shown as below:

Step 1: Create two sequences: ordered sequence and result sequence. Ordered sequence stores core objects and their directly reachable objects in ascending order according to their reachability-distance,

$$\begin{cases} UNDEFINED, if\ |N_\epsilon(o)| < MinPts \\ otherwise \\ max(core - distance(o), distance(o, p)) \end{cases}, (6)$$

where p and o are points in ROI, N_ϵ represents the ϵ-neighborhood of o, and $MinPts$ is a natural number. The result sequence is stored for the output order of sample points.

Step 2: When all the points in sample set D have been processed, the algorithm ends; otherwise, choose an un-processed core point (i.e. the point outside of the result sequence) and find its directly reachable sample point, and if the latter is not in the result sequence, put it in and rank it according to the reachability-distance.

Step 3: If the ordered sequence is empty, skip back to Step 2; otherwise, choose a first sample point from the sequence (i.e. the point with the shortest reachabiliry-distance) to expand and store it in the result sequence, if it is not in it.

3.1 Judge whether the expanded point is a core point; if not, go back to Step 3; otherwise, find all the reachable points of this expanded point;

3.2 Judge whether the reachable point is in the result sequence; if so, skip it; if not, go to the next step;

3.3 If the reachable point exists in the ordered sequence, and if the new reachability distance is less than the old one, replace it with the new distance and re-rank the ordered sequence;

3.4 If there is no such reachable point in the ordered sequence, insert this point and re-rank the sequence.

Step 4: Automatic clustering algorithm of significant features for interest objects:

4.1 Compose a new matrix with the output core distances and feature points in the OPTIC algorithm and rank them according to the $core-distances_{\epsilon,MinPts}(p) =$

$$\begin{cases} UNDEFINED, \\ if Card(N_{\epsilon}(p)) < MinPts \\ MinPts - distance(p), \\ otherwise \end{cases} ,(7)$$

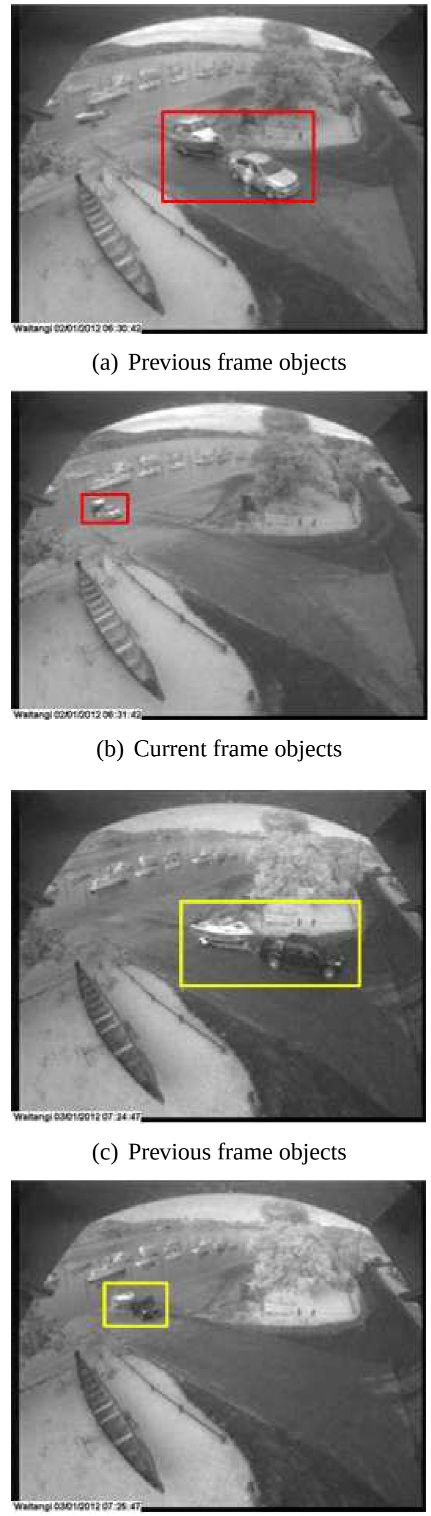
where p is a point in ROI, ϵ is a distance value, $N_{\epsilon}(p)$ represents the ϵ-neighborhood of p, $MinPts$ is a natural number and $MinPts-distance(p)$ indicates the distance from p to its $MinPts$' neighbor.

4.2 Fetch the set of feature points with shortest core distances.

3 Experimental Results

3.1 Scene from fisheries management

A scene under the camera surveillance of fisheries management is as shown in Figure 2. Restricted by external conditions, the color information of the images is almost lost and the interest points like ships/cars are shown with low signal-to-noise ratio. When the ships are getting closer to the seaside, it is even impossible to identify the ships with human vision. In limited image frame sequences, objects in current frames can not match with those in previous frames, so human identification for ships is quite difficult. As shown in Figure 2, the interest objects in the squares in Figure 2(a) and Figure 2(b) and those in Figure 2(c) and Figure 2(d) are the same targets, but it is nearly impossible to identify the relevance between them with human vision. The targets size and position in the different frames change a lot, so the visual features are weak. According to the statistics of this industry, it is quite common that the same ship in the same scene can not be correctly identified at different time. Therefore, extraction of significant features of interest objects based on computer vision technology is a key to ship control for fisheries management.



(a) Previous frame objects

(b) Current frame objects

(c) Previous frame objects

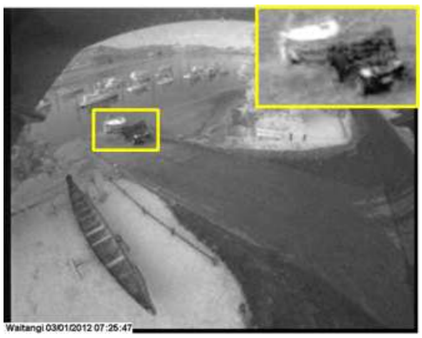
(d) Current frame objects

Figure 2: One target is quite different in visual features in previous and next frames. Red squares and yellow squares suggest the same target respectively.

3.2 Magnify image contrast

To increase the correctness of subsequent extractions, we amplify the input images. Figure 3 shows the comparison between the results before and after the equalization with histogram proposed in Section 2.2. We can see the signal-to-ratio of the target in the distance has been markedly magnified.

(a) Comparison of red-square target before and after the amplification.



(b) Comparison of yellow-square target before and after the amplification.

Figure 3: Comparison of targets before and after the amplification.

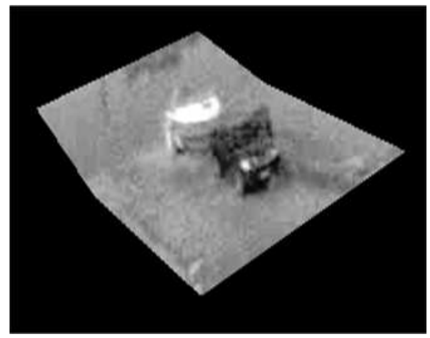
3.3 Selected ROI

To narrow down the scope for feature extraction, this paper defines the interest areas for subsequent image analysis according to the entry and exit areas of seaports under the camera surveillance of fisheries management. With the regional coordinates given in Section 2.3, we clip the corresponding interest areas as shown in Figure 4. As we can see, the target objects are in the defined areas.

3.4 Feature extraction

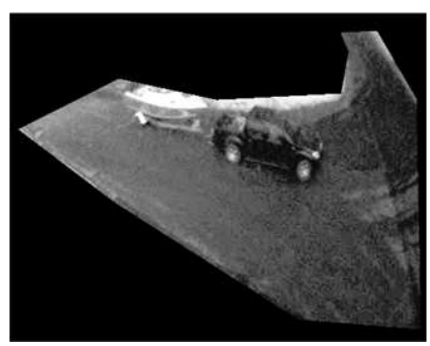
Figure 5(a) and 5(b) represent the extraction results of SURF feature algorithm from Figure 4(b) and 4(d). Note that, although feature points of the interest objects have been extracted as we expected, the surroundings however still show some SURF feature points.

(a)



(b)

(c)

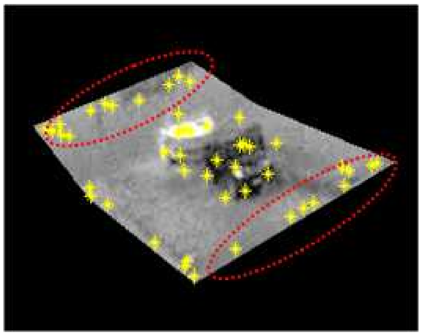


(d)

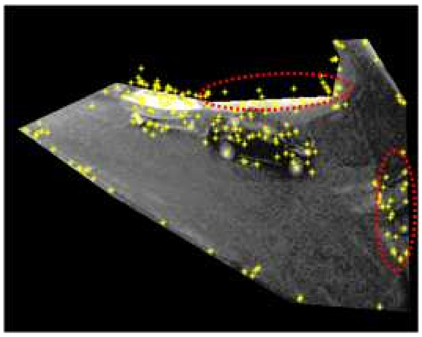
Figure 4: Image (b) and (d) are the interest areas of (a) and (c) respectively.

3.5 Removing boundary features

Evidently, in Figure 5, boundary SURF feature points occur in non-interest target areas, so the next step is



(a)



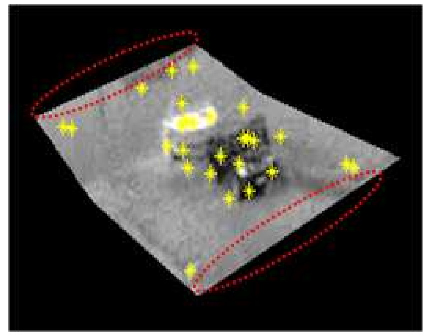
(b)

Figure 5: Image (a) and (b) show SURF feature extraction results of interest area images from Figure 4(b) and 4(d). The highlight areas are noise surf feature on the boundary.

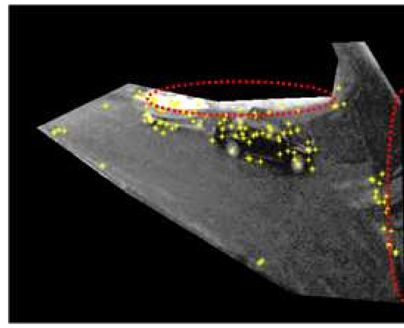
to remove these points with the method introduced in Section 2.5. Figure 6 shows the result of boundary smoothing of Figure 5, in which these points disappear.

3.6 Discovering significant features of interest objects

Though in Figure 6 the irrelevant boundary feature points are removed, we can still see background SURF feature points, which are not in the interest object areas, either. In order to cluster interest object feature in SURF feature set, we apply the property of automatic density clustering of OPTIC algorithm to further smooth the SURF feature points in target areas and directly extract SURF feature points on interest targets. As shown in Figure 7, red dots are extracted interest points, which are all located within the interest area.



(a)



(b)

Figure 6: Results of the boundary SURF feature points have been removed.

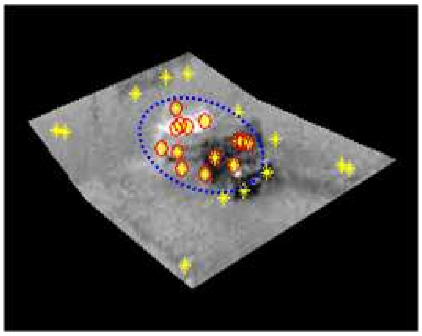
3.7 Results of experimenting on the real-scene image data

Table 1 indicated the experiments results on data in 1^{st} January, 2012 to evaluate the performance of the proposed method. The experimental result shows that the best detection time is from 22:00 to 23:59, followed by period from 13:00 to 15:59, and from 10:00 to 12:59, the accuracies are 96%, 95% and 94% respectively. Similar to other periods, the detection accuracies are over 90% but during the period from 00:00 to 03:59 which is 84% unfortunately. This is because the positions of background objects are affected dramatically by high tide and low tide in a dark environment.

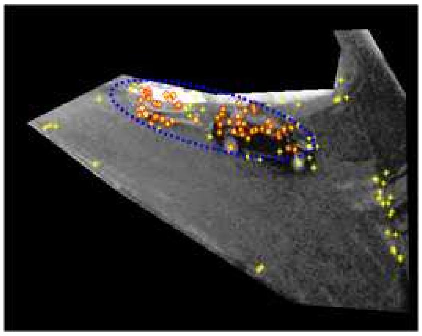
Table 1: The Experimental Results on All Data

Time	Number of images	Number of mis-detections	Correct rate
00:00-03:59	240	38	84%
04:00-06:59	180	17	90%
07:00-09:59	180	14	92%
10:00-12:59	180	10	94%
13:00-15:59	180	8	95%
16:00-18:59	180	16	91%
19:00-21:59	180	16	91%
22:00-23:59	120	4	96%

As shown in Table 1, the detection accuracy of the proposed method on the test data is about 91.63%, whereas the average detection accuracy on



(a)



(b)

Figure 7: Image (a) and (b) show extraction results of significant features of interest objects from Figure 6(a) and 6(b). The blue oval shows the detection of interest objects in ROI.

the same data with the previous approaches based on the FSAT Features from Accelerated Segment Test algorithm, MSER Features from Maximally Stable Extremal Regions algorithm, Harris Features from Harris-Stephens algorithm, MinEigen Features from Minimum Eigenvalue algorithm are 85.01%, 73.56%, 86.88%, 79.22% respectively.

4 Conclusions

Based on cascade framework algorithm and the complex background of fisheries environment surveillance, this paper proposes a new algorithm of clustering SURF features. Experiments prove that this method can efficiently screen significant features of interest objects from a mass of feature sets. It discards features on non-interest objects and is independent of pixel correlation between image frames. It can be directly used for automatic segmentation and separation of background objects, automatic recognition and tracking of interest objects, and counting and control of interest objects. Therefore, this paper provides some theoretical grounds for ship analysis in fisheries management.

Acknowledgements: The first author thanks the computing department of Unitec NZ for data support during author's study. This work is also supported in part by Scientific and Technological Research Program of Chongqing (Grant No. KJ131423, No. KJ131422 and KJ132206), Doctor Special Project from Chongqing University of Science and Technology (Grant No. CK2011B05 and CK2011B09) and Natural Science Foundation of Chongqing (Grant No. cstcjjA40041 and cstc2014jcyjA40006).

References:

[1] C. W. Clark, *The worldwide crisis in fisheries: economic models and human behavior*. Cambridge University Press, 2006.

[2] C. Costello, S. D. Gaines, and J. Lynham, "Can catch shares prevent fisheries collapse?" *Science*, vol. 321, no. 5896, pp. 1678–1681, 2008.

[3] G. Heal and W. Schlenker, "Economics: sustainable fisheries," *Nature*, vol. 455, no. 7216, pp. 1044–1045, 2008.

[4] B. Worm, R. Hilborn, J. K. Baum, T. A. Branch, J. S. Collie, C. Costello, M. J. Fogarty, E. A. Fulton, J. A. Hutchings, S. Jennings *et al.*, "Rebuilding global fisheries," *science*, vol. 325, no. 5940, pp. 578–585, 2009.

[5] F. Porikli, F. Bashir, and H. Sun, "Compressed domain video object segmentation," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 20, no. 1, pp. 2–14, 2010.

[6] C.-Y. Chung and H. H. Chen, "Video object extraction via mrf-based contour tracking," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 20, no. 1, pp. 149–155, 2010.

[7] J. C. Nascimento and J. S. Marques, "Performance evaluation of object detection algorithms for video surveillance," *Multimedia, IEEE Transactions on*, vol. 8, no. 4, pp. 761–774, 2006.

[8] P. Lobato Correia and F. Pereira, "Classification of video segmentation application scenarios," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 14, no. 5, pp. 735–741, 2004.

[9] P. Li and C. Wang, "Object of interest extraction in low-frame-rate image sequences and application," *Optical Engineering*, vol. 51, no. 6, pp. 101–112, 2012.

[10] K. Mikolajczyk and C. Schmid, "Scale and affine invariant interest point detectors," *International Journal of Computer Vision*, vol. 60, no. 1, pp. 63–86, 2004.

[11] J.-M. Morel and G. Yu, "Asift: A new framework for fully affine invariant image comparison," *SIAM J. Img. Sci.*, vol. 2, no. 2, pp. 438–469, Apr. 2009.

[12] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, "Speeded-up robust features (surf)," *Computer vision and image understanding*, vol. 110, no. 3, pp. 346–359, 2008.

[13] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, "Orb: an efficient alternative to sift or surf," in *Computer Vision (ICCV), 2011 IEEE International Conference on*. IEEE, 2011, pp. 2564–2571.

[14] S. Leutenegger, M. Chli, and R. Y. Siegwart, "Brisk: Binary robust invariant scalable keypoints," in *Computer Vision (ICCV), 2011 IEEE International Conference on*. IEEE, 2011, pp. 2548–2555.

[15] S. Stalder, H. Grabner, and L. Van Gool, "Cascaded confidence filtering for improved tracking-by-detection," in *Computer Vision–ECCV 2010*. Springer, 2010, pp. 369–382.

[16] M. Takács, S. Szénási, and Á. Szeghegyi, "Fuzzy logic control problems simulation based on parametrizied operators," in *Romanian-Hungarian Joint Symposium on Applied Computational Intelligence (SACI), Temesvar*, 2006, pp. 25–34.

[17] S. Szénási, Z. Vámossy, and M. Kozlovszky, "Evaluation and comparison of cell nuclei detection algorithms," in *Intelligent Engineering Systems (INES), 2012 IEEE 16th International Conference on*. IEEE, 2012, pp. 469–475.

[18] G. Bradski and A. Kaehler, *Learning OpenCV: Computer vision with the OpenCV library*. O'Reilly Media, Inc., 2008.

[19] H. Bay, T. Tuytelaars, and L. Van Gool, "Surf: Speeded up robust features," in *Computer Vision–ECCV 2006*. Springer, 2006, pp. 404–417.

[20] L. Juan and O. Gwun, "A comparison of sift, pca-sift and surf," *International Journal of Image Processing (IJIP)*, vol. 3, no. 4, pp. 143–152, 2009.

[21] M. Ankerst, M. M. Breunig, H.-P. Kriegel, and J. Sander, "Optics: Ordering points to identify the clustering structure," in *ACM SIGMOD Record*, vol. 28, no. 2. ACM, 1999, pp. 49–60.