An Automatic Video-based Drowning Detection System for Swimming Pools Using Active Contours

3 authors, including:

Nasrin Salehi
Isfahan University of Technology
6 PUBLICATIONS   **29** CITATIONS

An Automatic Video-Based Drowning Detection System for Swimming Pools using Active Contours

Nasrin Salehi and Maryam Keyvanara *
University of Isfahan, Faculty of Computer Engineering, Department of Artificial Intelligence, Isfahan, 81744, Iran
Email: saalehi.n69@gmail.com, maryam.k1ara@gmail.com

Seyed Amirhassan Monadjemmi
University of Isfahan, Faculty of Computer Engineering, Department of Artificial Intelligence, Isfahan, 81744, Iran
Email: amonadjemi@yahoo.co.uk

Abstract— Safety in swimming pools is a crucial issue. In this paper, a real time drowning detection method based on HSV color space analysis is presented which uses prior knowledge of the video sequences to set the best values for the color channels. Our method uses a HSV thresholding mechanism along with Contour detection to detect the region of interest in each frame of video sequences. The presented software can detect drowning person in indoor swimming pools and sends an alarm to the lifeguard rescues if the previously detected person is missing for a specific amount of time. The presented algorithm for this system is tested on several video sequences recorded in swimming pools in real conditions and the results are of high accuracy with a high capability of tracking individuals in real time. According to the evaluation results, the number of false alarms generated by the system is minimal and the maximum alarm delay reported by the system is 2.6 sec which can relatively be reliable compared to the acceptable time for rescue and resuscitation.

Index Terms— Drowning Detection, Contour, Color Space Analysis, Real-Time Image Processing

I. INTRODUCTION

Video surveillance can be used as a tool for monitoring and security. Observing public and private sites has increasingly become a very sensitive issue. The visual monitoring capabilities can be employed in many different locations to help people live more safely. Video-based surveillance systems are designed and installed in places such as railway stations, airports, and even dangerous environments. Image processing, pattern recognition and machine-vision based methods are efficient ways for real-time intelligent monitoring of the objects or events of interest [1-4].

The existing surveillance systems deliver valued information in monitoring of large areas. Applying intelligence in video surveillance systems allows real-time monitoring of places, people and their activities. The tracking approach can change with varying targets and can change from a single camera to multiple camera configurations [1, 2 and 4]. Tracking methods in video surveillance use different parameters such as objects' motion, position, path of movement and velocity [4, 5], biometrics such as skin color or clothes color [4, 6] and many more. The tracking must be robust and overcome occlusion and noise which are common problems in monitoring [4-6].

One important environment that the need for monitoring systems is crucially sensed is the swimming pool. Each year many people including children are drowned or very close to drowning in the deeps of the swimming pools, and the life guards are not trained well enough to handle these problems [7]. This raises the need for having a system that will automatically detect the drowning person and alarm the lifeguards of such danger. Real-time detection of a drowning person in swimming pools is a challenging task that requires an accurate system. The challenge is due to the presence of water ripples, shadows and splashes and therefore detection needs to have high accuracy.

A. Related Work

In swimming pool monitoring intelligent systems, different approaches have been proposed. Most methods perform background processing on input video frames. Some apply background subtraction and image denoising to detect the drowning person [8, 9]. In [9], a Gaussian Mixture Model is used for describing the pixels and the parameters of the model are updated with the EM algorithm. Also, neural networks can be trained to
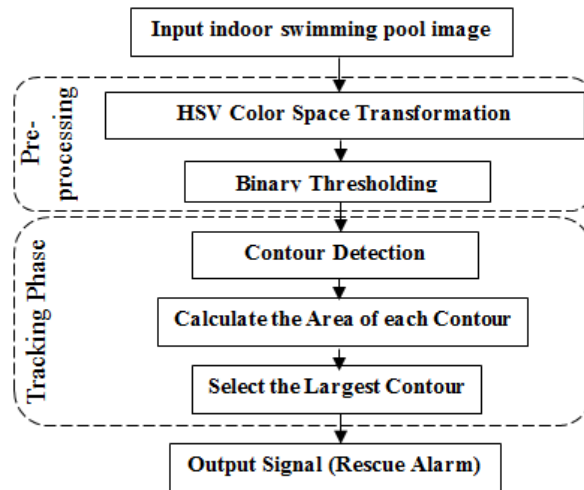
Fig. 1. The workflow of the drowning Detection method proposed in this paper.

classify near-drowning and normal swimming patterns [10]. However this requires to have a large dataset of both groups of behavior. The dataset is obtained in [10] by attaching a pressure sensor to a swimmer imitating drowning behavior and normal swimming.

Pattern recognition algorithms are also very useful in swimmer detection. In [11] a background model that has prior knowledge about swimming pools is employed. This hierarchical model operates on behavioral traits common in almost all troubled swimmers. Another vision-based system, depend on on detection of swimmers' body parts. This approach uses local motion and intensity information from image frames [12]. In [13] the YCbCr color model is selected for detection of the water polo players in water where luminance is separated and the Cb and Cr components are analyzed. Moreover, underwater ultrasonic sensors can detect drowning people up to 70 meters below water in the swimming pool along with a underwater video detection unit that locates and finds the victims [14].

This research presents a vision-based approach for detecting a drowning person and alarming the life guards of such situations. In this study, the person swimming in the pool is detected and tracked using the HSV color space properties and contour-based methods. As soon as the moving target remains under water for more than a determined period of time, an alarm is sent to the lifeguard rescues. The HSV color space is selected over other color spaces because it is more effective in segmenting the swimmer in various light conditions from the background.

The paper is organized as follows. In section II, the proposed method is presented in details. Experimental results, including discussions and reported performance results, are given in section III. Finally, conclusions are summarized in section IV.

II. PROPOSED METHOD

In this paper we have proposed a method for automatic real-time detection of a person drowning in the swimming pools. The overview of the proposed algorithm in this paper is presented in Fig. 1.

Our system is based on real time video analysis of the cameras installed around the swimming pool in a way which the entire swimming pool can be covered. Each camera is mounted on pool walls oriented downwards with a sharp angle, so that it can minimize the effect of lightening system which causes occlusions and foreshadowing. In this work, a ODROID-XU as a distributed system is installed in the swimming pool to collect all the video signals collected from cameras and process them using computer vision methods. The used hardware including the distributing system known as ODROID-XU, and our Logitech HD Pro C920 webcam used to record all the video sequences in this paper is illustrated in Fig. 2. The system is used to firstly detect the background of the pool and then decide to send an alarm to rescue team if a previously detected person is missing in video frames for an specific and defined period of time. In the next sections of this paper, we try to explain the concepts we used to detect and track individuals in swimming pools.

A. HSV Color Space Analysis

There are a number of color spaces that are suitable to be used in the area of video tracking and surveillance. They include RGB [15], YCbCr [13], CIE Lab and HSV [16, 17]. Each one can be used in different applications. Since the illumination data is inserted into the three color channels of the RGB color space, normalization of the RGB color space would allow a more robust tracking in this color space. This data can then be transformed into a different color space to separate the brightness effect from color information [18].
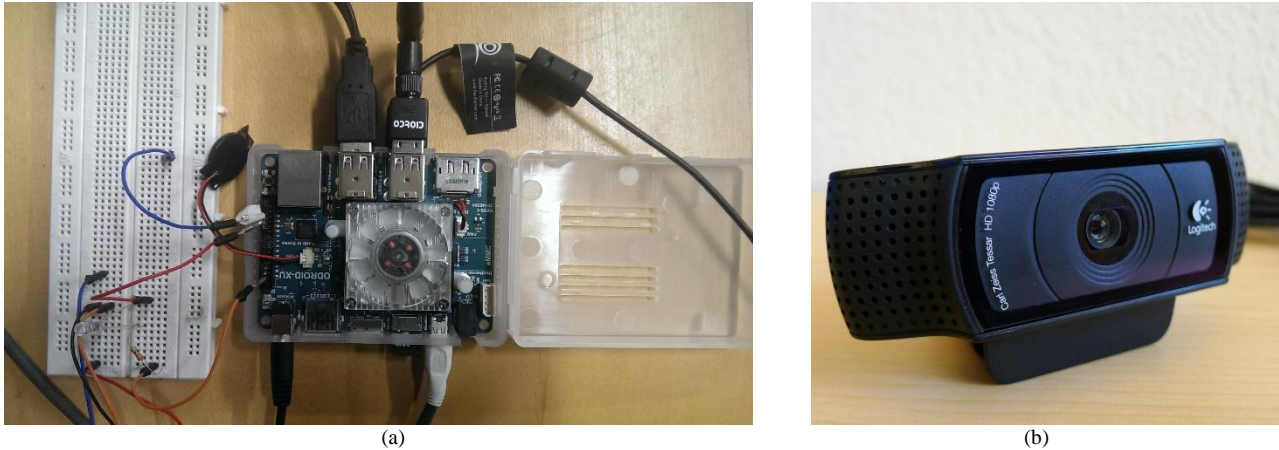
Fig. 2. Used software for the Video-Based Drowning Detection system proposed in this paper. (a) Our ODROID-XU board as the processing platform. (b) Logitech HD Pro C920 webcam used to record all the video sequences in this work.
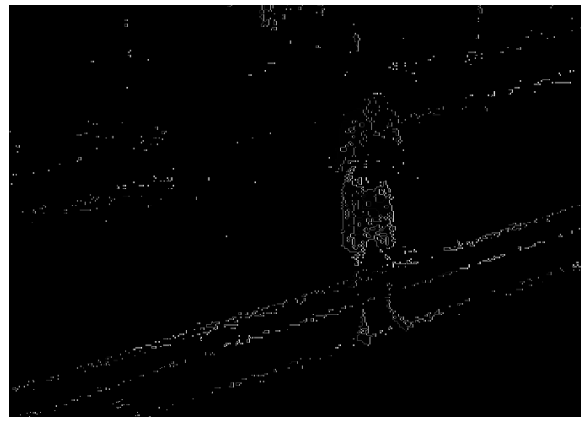

Fig. 3. A sample frame from recorded video sequences in swimming pool and the detected contours in this frame.

In the HSV color space, there are different layers of information and the luminance data is separated from the color information. The separation of brightness information from the color information makes the HSV color space very suitable for tacking purposes [18-20]. In HSV, the V channel contains the luminance information of the input image, and the H and S channels have the chromaticity information in them. These properties make this color space very effective in segmentation of the target object which is the swimmer. In addition, employment of chrominance in the HSV color space can provide the system with robust tracking. Also, the separation of the brightness information from the chrominance decreases the effect of uneven illumination in an image. Considering light intensity, HSV color model is both scale-invariant and shift-invariant [20].

Due to the vulnerability of color-based tracking algorithms and fluctuation of light conditions, for the proposed system we apply the HSV color model to find the target object which is the swimmer in the input video and also distinguish the background of swimming pool from swimmers. Before starting the detection, a single frame of the input image is given to the system. This frame should be chosen so that it is a suitable sample candidate of all the frames; that is it should contain a person swimming in the swimming pool. This will make the system have a higher accuracy during the detection process. Receiving this single frame, the object of interest which here is the human body in the blue background of the swimming pool, will be manually extracted and marked. With this prior knowledge, the appropriate values for H, S and V channels can be set and tuned. So, once the image is captured through the cameras installed around the swimming pool, its pixel values is converted to the HSV color space. Then, the HSV image obtained from every single frame is converted to binary image by a simple thresholding over HSV values. This threshold is used calculated using the prior knowledge obtained from the initial step in detection phase. As a result, the binary image will be a black and white image in which background will turn black and the foreground (which is the swimmer) would be white.

B. Contour Detection

Contours can be used to find object outlines in images and effectively track targets in videos sequences. In tracking algorithms that are based on contours, the

Table 1. Obtained results from 3 video sequences containing their frame counts along with their relative true and false alarms sent by the proposed algorithm.

Sequence	frames	True Alarm	False Alarm
No. 1	384	3	None
No. 2	603	1	1
No. 3	522	2	None

objects are tracked using their outlines as boundary contours. The contours should be updated dynamically in successive frames. In active contours concept, a closed contour is limited to the object's boundary. Hence the contour covers object region and object segmentation is reached. The contours are managed by their energy functions. This function consists of internal, external and shape energy [21, 22].

Active contour representations have been applied in different fields to track non-rigid objects [22, 23, 24]. An active contour representation is defined as in (1).

$$\Phi(x,y) - \begin{cases} 0 & (x,y)\epsilon C \\ d(x,y,C) & (x,y)\epsilon R_{out} \\ -d(x,y,C) & (x,y)\epsilon R_{in} \end{cases} \quad (1)$$

The parameters R_{in} and R_{out} are the regions inside and outside the contour C. The function d(x,y,C) returns the smallest Euclidean distance from point (x,y) to the contour C.

Segmentation is a technique that segments an image frame into sections to discover the object of interest. In segmentation algorithms, it is very important to have an efficient partitioning method. Once a video frame is segmented, the object of interest is detected for tracking. In many indoor swimming pools, the background only consists of a number of features including the water and the lane drivers. When people are swimming in the pool, the swimmers are the only objects that are distinguishable from the background due to their motion and color. Therefore the first step is to achieve an unsupervised segmentation of the empty pool.

After that the input frame is converted to a binary image, the contours in the binary image are found. Out of all the discovered contours, the one with the largest area is selected and tracked in consecutive frames. The resulting contour of the previous frame is taken as initialization in each frame. In Fig. 3 a sample of a given frame from recorded video sequences in swimming pool and the detected contours in this frame is illustrated.

An object tracker is important because it can find the motion trajectory of the target object as video frames proceed through time. This is done by identifying the position of the object in every frame of the video. In this paper the tracking procedure is done by applying HSV thresholding algorithm in every single frame and then choosing the contour with largest area available in the result binary image. So, the area that is occupied by the

target object is found by the algorithm at every instant and tracked in the subsequent frames.

III. EXPERIMENTAL RESULTS

The proposed system provides an alarm to the lifeguard rescues as soon as the tracking person is detected as being drowning. A visual indicator is used to determine whether the target being tracked in on the surface (green) or below the water (red). A red alarm along with a beep sound is generated when the swimmer is not found by the system for more than a specific number of consecutive frames regarding to the fact that the speed of different boards vary. In this research, we used an ODROID-XU board which contains Exynos5 Octa Cortex™-A15 1.6Ghz quad core and Cortex™-A7 quad core CPUs, and also a 2Gbyte LPDDR3 RAM. For video capturing purpose we used a Logitech HD Pro C920 webcam which is capable of recording in full 1080p at 30 frames per second. This hardware along with the developed algorithm to track swimming objects in pools can process about 6 frames every second. As a result, we can let the alarm go on if the swimming object is not fount after 30 consequent frames. It worth telling that we used the OpenCV library for the implementation of this software.

To evaluate the performance of our system, different footages recorded in real swimming pools were used. We used 3 sequence of videos to evaluate the proposed method. Each sequence contains different number of frames and is taken from various views of the swimming to make the evaluation results more reliable. Table I shows the obtained results from 3 video sequences containing their frame counts along with their relative true and false alarms sent by the proposed algorithm in this paper. True alarms (True Positive) represent the situations in which one person is being drowned and the system should raise an alarm to notify the lifeguard in the swimming pool. Also the false alarms (False Positive) represent the conditions in which a drowning alarm has been reported by mistake. All these situations are considered as normal situations that their importance is ranked low compare to false negative ones. In sequence No. 2, we have a drowning condition which takes 20.1 seconds. The presented system was able to detect the drowning person and its position easily, though it reported a short period (1.4 sec.) of a true false situation as a true positive which can be easily overlooked. The sequences No. 1 and No. 3 contain 3 and 2 drowning condition, which the proposed method succeeded to detect them all. Also in these two sequences we had no

Table 2. Alarm delays regarding to the length of each drowning conditions occurred in each video sequence.

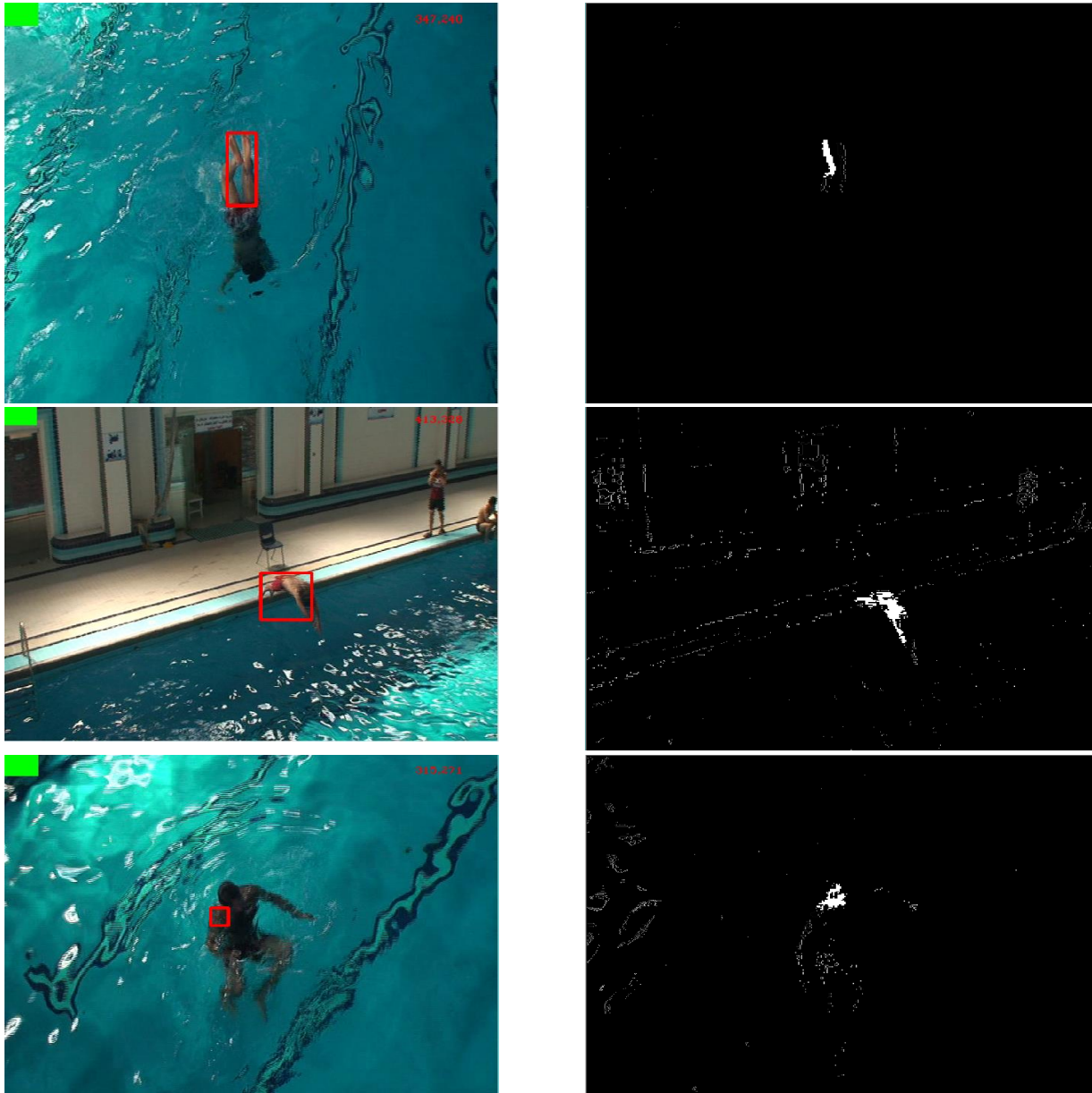| Sequence	Drowning Time (length)	Alarm Delay
No. 1	12.8 sec	1.2 sec
No. 2	20.1 sec	2.6 sec
No. 3	17. 4 sec	0.8 sec

Fig. 4. Results of applying HSV thresholding on frames' pixels in several frames of 3 video sequences.

false Alarms, and this fact can represent high performance and accuracy of our presented work. As could be seen in Table I, the number of false alarms generated by the system is minimal.

Table II provides more performance evaluation of our system by depicting the Alarm delays regarding to the length of each drowning conditions occurred in each video sequence. The average detection delay for 3 video sequences is 1.53 seconds which shows high performance and accuracy of the proposed method in this application.

Fig. 4 shows the results of applying HSV thresholding on frames' pixels in several frames of 3 video sequences. The result images are excluded from pool's background and are prepared for contour detection.
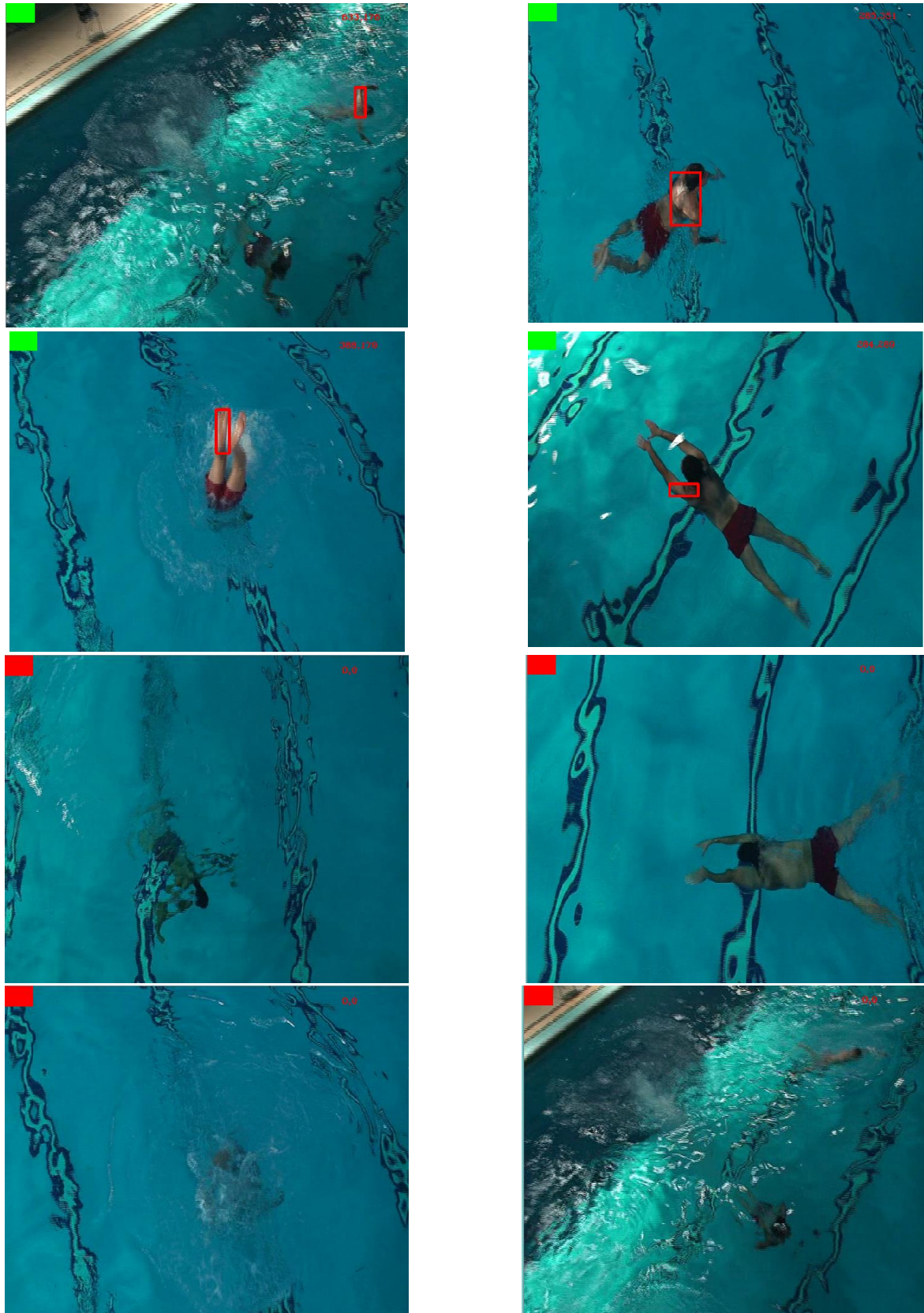
Fig. 5. Detection results for video sequences in different conditions including frames in which the object is visible on the surface of the water (the first two rows of images) and also frames in which the object is drowning (the last two rows of images).

Fig. 5 shows drowning detection results for 3 video sequences in different conditions including frames in which the object is drowning and also frames in which the object is visible on the surface of the water. In both situations, we achieved the desired results which enable us to use the proposed system for high performance drowning detection in swimming pools.

IV. CONCLUSION

In this paper, we provided a method to robust human tracking and semantic event detection within the context of video surveillance system capable of automatically detecting drowning incidents in a swimming pool. In the current work, an effective background detection that incorporates prior knowledge using HSV color space and contour detection enables swimmers to be reliably detected and tracked despite the significant presence of water ripples. The system has been tested on several instances of simulated water conditions such as water reflection, lightening condition and false alarms. Our algorithm was able to detect all the drowning conditions along with the exact position of the drowning person in the swimming pool and had an average detection delay of 1.53 seconds, which is relatively low compared to the needed rescue time for a lifeguard operation. Our results show that the proposed method can be used as a reliable multimedia video-based surveillance system.

REFERENCES

[1] Foresti, Gian Luca, Petri Mähönen, and Carlo S. Regazzoni, eds. Multimedia video-based surveillance systems: Requirements, Issues and Solutions. Vol. 573. Springer Science & Business Media, 2012.

[2] Jones, Graeme A., Nikos Paragios, and Carlo S. Regazzoni, eds. Video-based surveillance systems: computer vision and distributed processing. Springer Science & Business Media, 2012.

[3] Conde, Cristina, et al. "HoGG: Gabor and HoG-based human detection for surveillance in non-controlled environments." Neurocomputing 100 (2013): 19-30.

[4] Wang, Xiaogang. "Intelligent multi-camera video surveillance: A review." Pattern recognition letters 34.1 (2013): 3-19.

[5] Gudyś, Adam, et al. "Tracking people in video sequences by clustering feature motion paths." Computer Vision and Graphics. Springer International Publishing, 2014. 236-245.

[6] Vezzani, Roberto, Davide Baltieri, and Rita Cucchiara. "People reidentification in surveillance and forensics: A survey." ACM Computing Surveys (CSUR) 46.2 (2013): 29.

[7] Bierens, Joost, and Andrea Scapigliati. "Drowning in swimming pools." Microchemical journal 113 (2014): 53-58.

[8] Zhang, Chi, Xiaoguang Li, and Fei Lei. "A Novel Camera-Based Drowning Detection Algorithm." Advances in Image and Graphics Technologies. Springer Berlin Heidelberg, 2015. 224-233.

[9] Fei, Lei, Wang Xueli, and Chen Dongsheng. "Drowning Detection Based on Background Subtraction." Embedded Software and Systems, 2009. ICESS'09. International Conference on. IEEE, 2009.

[10] Kharrat, Mohamed, et al. "Near drowning pattern detection using neural network and pressure information measured at swimmer's head level." Proceedings of the Seventh ACM International Conference on Underwater Networks and Systems. ACM, 2012.

[11] Kam, Alvin H., Wenmiao Lu, and Wei-Yun Yau. "A video-based drowning detection system." Computer Vision—ECCV 2002. Springer Berlin Heidelberg, 2002. 297-311.

[12] Chan, Kwok Leung. "Detection of swimmer using dense optical flow motion map and intensity information." Machine vision and applications 24.1 (2013): 75-101.

[13] Pleština, Vladimir, and Vladan Papić. "Features analysis for tracking players in water polo." 16th International Conference on Automatic Control, Modelling & Simulation. 2014.

[14] Wang, Hua, and Sing Kiong Nguang. "Intelligent and Comprehensive Monitoring System for Swimming Pool." International Journal of Sensors Wireless Communications and Control 3.2 (2013): 85-94.

[15] Kim, Jong Sun, Dong Hae Yeom, and Young Hoon Joo. "Fast and robust algorithm of tracking multiple moving objects for intelligent video surveillance systems." Consumer Electronics, IEEE Transactions on 57.3 (2011): 1165-1170.

[16] Kim, Wonjun, and Changick Kim. "Background subtraction for dynamic texture scenes using fuzzy color histograms." Signal Processing Letters, IEEE 19.3 (2012): 127-130.

[17] Wang, Yu-Chen, et al. "The color identification of automobiles for video surveillance." Security Technology (ICCST), 2011 IEEE International Carnahan Conference on. IEEE, 2011.

[18] Gonzalez, Rafael C. "RE woods, Digital Image Processing." Addison–Wesely Publishing Company (1992).

[19] Oliveira, V. A., and A. Conci. "Skin Detection using HSV color space." H. Pedrini, & J. Marques de Carvalho, Workshops of Sibgrapi. 2009.

[20] Sural, Shamik, Gang Qian, and Sakti Pramanik. "Segmentation and histogram generation using the HSV color space for image retrieval." Image Processing. 2002. Proceedings. 2002 International Conference on. Vol. 2. IEEE, 2002.

[21] Pitas, Ioannis. Digital image processing algorithms and applications. John Wiley & Sons, 2000.

[22] Li, Xi, et al. "A survey of appearance models in visual object tracking." ACM transactions on Intelligent Systems and Technology (TIST) 4.4 (2013): 58

[23] Sun, Xin, Hongxun Yao, and Shengping Zhang. "A novel supervised level set method for non-rigid object tracking." Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on. IEEE, 2011.

[24] K.R.Ananth, S. Pannirselvam,"A Geodesic Active Contour Level Set Method for Image Segmentation." International Journal of Image, Graphics and Signal Processing (IJIGSP) 4.5 (2012): 31-37.

Authors' Profiles

Nasrin Salehi Nasrin Salehi, was born 1990, in Isfahan, Iran. She got her B.Sc. in computer engineering, Computer Software, from Isfahan University of Technology, Isfahan, Iran, in 2012. She also got her M.Sc. in computer engineering, Artificial Intelligence, from University of Isfahan, Isfahan, Iran, in 2015. Her research interests include Pattern Recognition, Image Processing, Machine Vision, and specifically Medical Image Processing.

Maryam Keyvanara Maryam Keyvanara, was born 1988, in Isfahan, Iran. She got her B.Sc. in computer engineering, Computer Software, from University of Isfahan, Isfahan, Iran, in 2011. She Also got her M.Sc. in computer engineering, Artificial Intelligence, from University of Isfahan, Isfahan, Iran, in 2015. Her research interests include Image Processing, Machine Learning, and Machine Vision.

S. Amirhassan Monadjemi Amirhassan Monadjemi, was born 1968, in Isfahan, Iran. He got his PhD in computer engineering, pattern recognition and image processing, from University of Bristol, Bristol, England, in 2004. He is now working as an associate professor at the Department of Computer Engineering, University of Isfahan, Isfahan, Iran. His research interests include pattern recognition, image processing, artificial neural networks, and physical detection and elimination of viruses.