Non-Invasive Video Analysis Technique for Detecting Sleep Apnea

HARISH B
Department of Computer Science and Engineering
Amrita Vishwa Vidyapeetham
Chennai, India
harishbharish6@gmail.com

***Abstract*—Sleep Apnea is a serious sleep disorder in which breathing starts and stops repeatedly during sleep. If this condition is left untreated, it can lead to serious health issues. This paper proposes a robust method for detecting sleep apnea and keeping track of the region of interest automatically. The methodology involves converting the captured video into grayscale images, and then apply Gaussian blur and Canny edge detection to detect noise-free edges of the body. Further, the current frame is compared to the previous frame to detect body movements accurately. The proposed method is experimented with different testing conditions to demonstrate its efficiency and accuracy in identifying sleep apnea. This paper also introduces a novel algorithm for automatically adjusting the region of interest within the frame. Furthermore, the advantages and limitations of the proposed algorithm are also been discussed.**

Keywords—Obstructive Sleep Apnea, Gaussian blur, Canny edge detection, contour detection, object tracking.

I. Introduction

Sleep Apnea is a sleeping disorder in which the person affected by it, involuntarily stops breathing for a period of 10 to 30 seconds repeatedly during sleep. If the condition is left untreated, it might increase the risk of getting cardiovascular diseases and high blood pressure. In the case of infants particularly, if sleep apnea is left untreated it might lead to long-term complications due to irregular oxygen supply to the brain. Therefore, it is very crucial to detect sleep apnea at earlier stages.

Sleep Apnea is categorized into two types, namely Obstructive Sleep Apnea (OSA) and Central Sleep Apnea (CSA). The fundamental cause for OSA is an obstructed nasal passage. The muscles in the back of the throat collapse, and blocks the airway while sleeping. It is one of the most common types. In CSA, the brain fails to signal the muscles to breathe and is also a rare condition. It is related to a neurological disorder.

A. *Available Solutions*

The currently accepted method for sleep apnea detection is Polysomnography (PSG). It is a standard diagnostic tool used for measuring blood oxygen level, electrocardiography, and electromyography during sleep. Polysomnography takes place in a specialized sleep centre or a hospital administered by a technician. However, it is an expensive method and a lot of sensors are required to be attached to the patient's body, which would make the patients feel uncomfortable.

Acoustic techniques make use of audio recordings to detect sleep apnea in patients. The microphone senses the breathing sound and is processed to detect anomalies in the patient's breathing. However, this technique does not effectively differentiate between the patient's breathing sound and the noise in the background, which makes it difficult to identify and process the sound of breathing in patients, which is of low strength when compared to the background noise.

A lot of research has been done on using video processing techniques to identify sleep apnea. It is a non-invasive technique and can also be set up easily at home. The previous research works [1], [2], [3], [4], show that the use of the Gaussian blur technique and Canny edge detection algorithm is an efficient way of finding subtle body movements made by the patient. However, the previous researches do not provide an effective way to keep track of the region of interest and should be periodically updated. In the proposed method, this issue has been discussed in detail and an effective algorithm is suggested to overcome this issue.

II. Methodology

In the proposed method, a specific region of interest is selected in the captured video and is automatically updated when the patient moves. First, the video is converted into frames, so that the frames can be compared to identify the movements. When the patient inhales and exhales, the diaphragm contracts and relaxes, which helps in determining whether the patient is breathing or not. Now, the color frames are converted into grayscale frames. Then, the Gaussian blur is applied to the frames, to smooth the images and remove background noise in them. Furthermore, the Canny edge detection is applied to find the edges or outline of the patient's body. An audible alarm is triggered if no movement is observed after 10 seconds. An email alert is also sent to notify that the patient is not breathing. When the body part observed moves out of the region of interest, the algorithm can detect it and will automatically adjust the region of interest with the help of the object tracking technique.

A. *Gaussian Blur*

Gaussian blur is a smoothing technique that blurs out the images. It is used to remove the background noises in the image so that the patient's body alone can be focused to find any movements. The Gaussian blur is based on the Gaussian distribution function which is defined as,

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{1}$$

Here, x denotes the distance from the origin in the horizontal axis, y denotes the distance from the origin in the vertical axis, and σ denotes the standard deviation of the Gaussian distribution. The values obtained from the Gaussian distribution function are used to form a convolution matrix, also known as a kernel. The image is then convolved with the n-by-n kernel to obtain the blurred image. The amount by which the image is getting blurred depends on the size of the kernel.

The principle of this function is that it will add more weight to the central pixel and less weight to the neighbouring pixels and sums them up. Then the result will be placed in the central pixel. This process is repeated for the whole pixel matrix and all the pixels will be replaced with their respective values. As a result, the whole image will be blurred.

B. *Canny Edge Detection*

Canny edge detection [5] is a multi-stage algorithm which is used to detect edges or outline of objects in an image. The first stage is to reduce noise in the image by using the Gaussian blur technique. The second stage is to find the intensity gradient of the image. In this stage, the image is filtered with Sobel operator in both the horizontal and vertical direction to get the first derivative in the horizontal direction (G_x) and vertical direction (G_y). The edge gradient and direction can be determined by,

$$G = \sqrt{G_x^2 + G_y^2} \tag{2}$$

$$\theta = tan^{-1}\left(\frac{G_y}{G_x}\right) \tag{3}$$

The third stage involves applying non-maximum suppression to get a binary image with thin edges. Each pixel in the image is checked if it is a local maximum in its neighbourhood in the direction of the gradient. If it is not a local maximum, then the pixel is suppressed. The final stage decides if an identified edge is a strong edge or not.

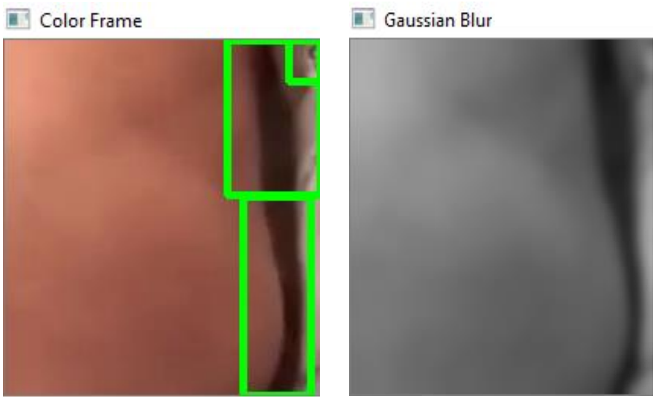

Fig. 1. Gaussian blur applied to the color frame with contour area.

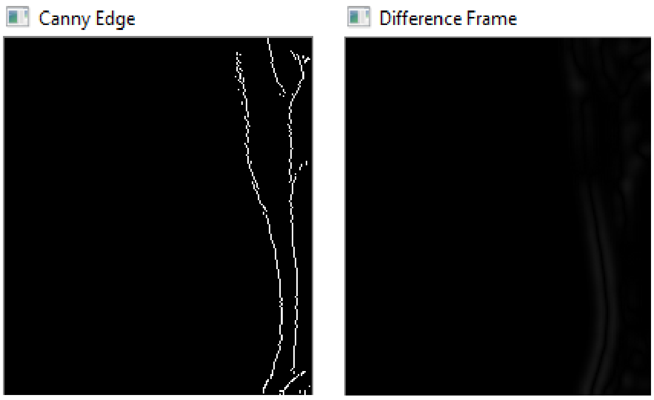

Fig. 2. Comparison between the Canny edge detection frame and the absolute difference frame.

In this stage, those edges with an intensity gradient more than the maximum threshold value are considered to be strong edges and those below minimum threshold values are considered to be weak edges. The edges which lie between the minimum and maximum threshold value are classified based on their connectivity. This method is known as edge tracking by hysteresis.

III. IMPLEMENTATION

The non-invasive video analysis technique for observing body movements turns out to be an efficient method when compared to the other state-of-the-art methods. The algorithm is implemented in the python programming language, along with the OpenCV library, which provides a wide range of functions for implementing the video processing techniques including Gaussian blur, Canny edge detection, and object tracking. The Pandas library provides the functions for storing the start and end time of sleep apnea in an excel sheet, which can be further used to identify the patient's breathing pattern and to determine how often the sleep apnea occurs.

The video is fed into the algorithm, which then converts the video into frames. The program asks the user for the region of interest in the frame, where the motion needed to be detected. Once the region of interest is selected, the frames are converted into grayscale images. The Gaussian blur can identify the noises easily in a grayscale image compared with a color image. The Gaussian blur removes all the background noises, then the Canny-Edge detection is applied to detect the edges on the image.

Once the Gaussian blur and Canny-Edge detection are applied, the present frame is compared with the previous frame for identifying the differences in the frames. These differences identified in the frames are the movements made by the patient. Then, the movements which are above a particular threshold value are identified and are highlighted in white color and the remaining parts of the frame are made black in color. Thus, the contours can be found in the frames effectively. These contours are enclosed in a green rectangular box, which denotes the part of the patient's body in the frame where the movements are detected. If no movements are detected after 10 seconds, then an audible alarm is triggered and an email alert is sent to notify that the patient is not breathing. When the program ends, the start and end times of the sleep apnea session are appended in the excel sheet along with the date. The algorithm also makes use of the object tracking method to automatically update the region of interest and keep track of it. OpenCV has 8 different trackers, each having its limitations. The trackers can be selected based on the need. The Minimum Output Sum of Squared Error (MOSSE) tracker [6] can be used to operate at higher frames per second (fps), about 450 fps, and more. The Discriminative Correlation Filter with Channel and Spatial Reliability (DCF-CSR) [7], also known as the CSRT tracker can be used to operate at a comparatively lower fps, about 25 fps, but it is highly accurate than the MOSSE tracker.

A. Object Tracking

Object tracking is an area of research under computer vision, which is used to track objects as they move across a series of video frames. In the proposed algorithm, the object refers to the observed body part of the patient. In this technique, a unique ID is assigned for the object in the initial stage, and then it tracks the objects across the frames in the captured video. It estimates the trajectory of an object across the frames.

The MOSSE tracker uses an adaptive correlation for object tracking which produces stable correlation filters when initialized using a single frame of the video. It adapts to the changes in the appearance of the object while tracking. The MOSSE tracker does not depend on changes in lighting, non-rigid transformations, pose, and scale. It also detects occlusion based upon the peak-to-sidelobe ratio, which enables the tracker to pause and resume where it left off when the object reappears. The correlation filter-based tracking system undergoes a number of steps to track the target object. The first step involves, application of the Fast Fourier Transform (FFT) to both the target object's template and the image. Then, convolution operation is performed between the target object's template and image. The result from previous step is inverted to the special domain using Inverse Fast Fourier Transform (IFFT). The position of the object corresponds to the maximum value obtained from the IFFT.

The CSRT tracker uses the spatial reliability map for adjusting the filter support to the part of the selected region from the frame for tracking. This ensures enlarging and localization of the selected region and improved tracking of the non-rectangular regions or objects. The CSRT tracker is more accurate in tracking objects when compared to the MOSSE tracker. It operates at a lower fps but gives higher accuracy when it comes to object tracking. The OpenCV library provides both the MOSSE and the CSRT tracker.

Algorithm:

1. Input the video source (either recorded or real-time monitored from the camera).

2. Convert the video into frames for processing.

3. Select the region of interest in the frame.

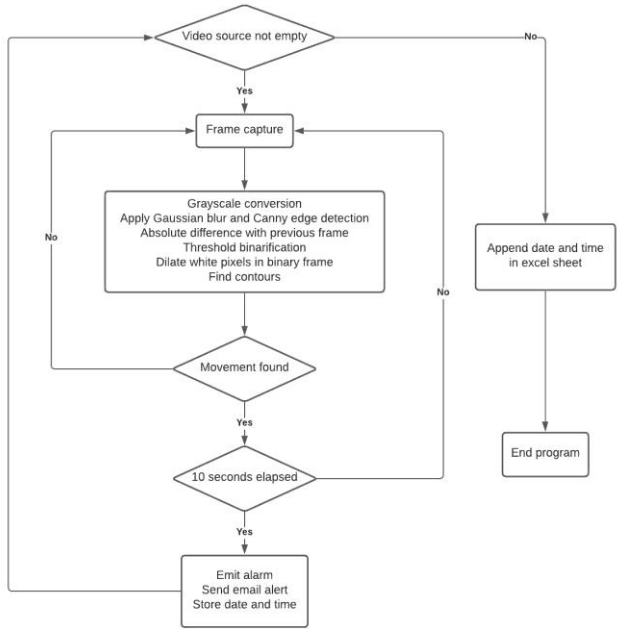

Fig. 3. Flowchart representation of the algorithm.

4. Convert the frame into a grayscale image and apply Gaussian-Blur for smoothing.

5. Apply Canny-Edge to detect the edges in the frame and once again apply Gaussian-Blur (if needed).

6. Compare the present frame with the previous frame for identifying movements.

7. Find the movements which are above a particular threshold and highlight them.

8. Find the contours and enclose them in rectangular boxes to imply there is movement in the region.

9. Alert the user, if the movement is not found.

10. If the part of the patient's body observed in the region of interest moves, automatically update the region of interest with the object tracking technique.

11. Take note of the time when the movements are not observed and append it in the excel sheet for future references.

IV. Results and Discussion

The tests were performed on both recorded and real-time monitored videos to find the efficiency of the algorithm. The frames were processed at 30 fps and 60 fps for testing purposes. At 30 fps, the CSRT tracker was able to perform quite well and was more accurate. But at 60 fps, the CSRT tracker was not able to track the objects accurately and was slow in updating the region of interest. The MOSSE tracker was able to perform quite well in both the 30 and 60 fps and is much faster in tracking the objects when compared to the CSRT tracker. The MOSSE tracker was able to outperform the CSRT tracker at higher fps. The CSRT tracker can be considered when the fps is low and also it is highly accurate at a lower fps.

TABLE I.
Percentage accuracy of the algorithm

Test	Percentage of accuracy
Sleep Apnea detection	90%
False alarm ratio	20%

Table 1 shows that the sleep apnea detection accuracy of the proposed algorithm is about 90%, thus illustrating a very high accuracy rate. The false alarm ratio is quite high because the patient's movement is not happening only because of breathing. The other body movements would interfere and the algorithm fails to differentiate it from the movements caused by breathing. This error can be reduced by observing multiple regions of interest.

One of the limitations faced by this video analyzing technique is that, when the lighting conditions are inadequate, this technique fails to identify the movements. To overcome this issue, the video can be captured in better lighting conditions, or a night vision camera can be used to monitor the patient during dull lighting conditions.

V. CONCLUSION

The results of the proposed method show that it is highly efficient in detecting sleep apnea. The novel method of automatically updating the region of interest with the object tracking technique proves to be an efficient way to overcome the issue of keeping track of objects within the frame. In the previous researches, this issue was not addressed with a working solution.

The non-invasive video analysis technique for detecting sleep apnea is a useful tool. This is especially true in the case of newborn babies. The newborn babies will be sleeping for most of the time, so it is not advised to use contact devices, since the placement of contact devices and sensors might interfere with their normal growth.

Moreover, the proposed method is non-intrusive and doesn't require any wires or sensors that need to be attached to the body. This makes it more convenient and comfortable for the patient. It is also portable and can be easily set up in either the patient's house or the hospital. Also, this detection method is inexpensive when compared to the other available solutions for detecting sleep apnea.

VI. REFERENCES

[1] G. Jayatilaka, H. Weligampola, S. Sritharan, P. Pathmanathan, R. Ragel and I. Nawinne, "Non-contact Infant Sleep Apnea Detection," 2019 14th Conference on Industrial and Information Systems (ICIIS), Kandy, Sri Lanka, 2019, pp. 260-265, doi: 10.1109/ICIIS47346.2019.9063269.

[2] S. Sharma, S. Bhattacharyya, J. Mukherjee, P. K. Purkait, A. Biswas and A. K. Deb, "Automated detection of newborn sleep apnea using video monitoring system," 2015 Eighth International Conference on Advances in Pattern Recognition (ICAPR), 2015, pp. 1-6, doi: 10.1109/ICAPR.2015.7050675.

[3] F. Mendonça, S. S. Mostafa, A. G. Ravelo-García, F. Morgado-Dias and T. Penzel, "A Review of Obstructive Sleep Apnea Detection Approaches," in IEEE Journal of Biomedical and Health Informatics, vol. 23, no. 2, pp. 825-837, March 2019, doi: 10.1109/JBHI.2018.2823265.

[4] D. L. Dorantes Olvera, D. S. López, M. Á. Martinez Prado, J. R. Reséndiz and Y. O. Rivera, "Noninvasive monitoring system for early detection of apnea in newborns and infants," 2016 IEEE EMBS Conference on Biomedical Engineering and Sciences (IECBES), 2016, pp. 494-498, doi: 10.1109/IECBES.2016.7843500.

[5] J. Canny, "A Computational Approach to Edge Detection," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. PAMI-8, no. 6, pp. 679-698, Nov. 1986, doi: 10.1109/TPAMI.1986.4767851.

[6] D. S. Bolme, J. R. Beveridge, B. A. Draper and Y. M. Lui, "Visual object tracking using adaptive correlation filters," 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Francisco, CA, USA, 2010, pp. 2544-2550, doi: 10.1109/CVPR.2010.5539960.

[7] A. Lukežic, T. Vojír, L. C. Zajc, J. Matas and M. Kristan, "Discriminative Correlation Filter with Channel and Spatial Reliability," 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 2017, pp. 4847-4856, doi: 10.1109/CVPR.2017.