

IMPLEMENTATION OF FEATURES DYNAMIC TRACKING FILTER TO TRACING PUPILS

Salvador Medrano¹, Ma. Guadalupe Medina¹, J. Juan Hernández¹ and L. Carlos Altamirano²

¹Master of computer systems, Technological Institute of Apizaco, Tlaxcala, México

²Department of computer science, Autonomous University of Puebla, Puebla, México

ABSTRACT

The objective of this paper is to show the implementation of an artificial vision filter capable of tracking the pupils of a person in a video sequence. There are several algorithms that can achieve this objective, for this case, features dynamic tracking selected, which is a method that traces patterns between each frame that form a video scene, this type of processing offers the advantage of eliminating the problems of occlusion patterns of interest. The implementation was tested on a base of videos of people with different physical characteristics of the eyes. An additional goal is to obtain information of the eye movements that are captured and pupil coordinates for each of these movements. These data could help some studies related to eye health.

KEYWORDS

Artificial vision, tracking, dynamic tracking features.

1. INTRODUCTION

The object tracking is one of the current areas of development in the field of computer vision. This type of processing can be done with static images and sets of images that form a video. There are several algorithms that allow analyze these image sequences to find features within them.

In this paper is shown the implementation of features dynamic tracking filter focused on tracking the pupils of a person in order to find answers to questions such: where are the pupils?, which displacements performed?, what features have detected movements? and the characteristics of the videos influence the efficiency of the filter?

For tracking work correctly, previously performed some filters of improve images that are applied to each frame of the video sequence, among them are converting to grayscale, segmentation for where are initially pupils, a crop of the images to make lighter and easier tracking of pupils, erosion and dilation of the cutout to eliminate possible distractions. In addition to the set of filters applied infrared technology is used to increase the efficiency of tracking and accuracy results.

This work aims to create a tool to help eye care specialists to more accurately diagnose their patients, not replace the work of doctors, let alone give a diagnosis only be a support tool.

2. TRACKING OBJECTS

One of the most important topics of computer vision is object tracking. This procedure allows to estimate the position of one or more objects at run time within a sequence of images.

There are different algorithms focused on solving these problems, some more efficient than others depending on the techniques used and factors as the quality of images being processed.

3. FEATURES DYNAMIC TRACKING

The basis of dynamic tracking features is the reduction of functional complexity when following salient features of the object, instead of the continuous region of the object or its contours. Therefore work the recognition and object tracking through the extraction elements, grouped at the highest level and then making a features matching between images. Are used as parameters features: corners, color information and texture. In features dynamic tracking of these are dynamically determined and tracked over consecutive frames by estimating feature movement and looking for the next frame.

A point in feature-based tracking is the paradox between complexity and efficiency of monitoring. For low levels features as the position of the coordinates of the edges, which are easy to extract, it is very difficult to make the trace, due to the difficulty of establishing a one to one correspondence between them. [2]

3. METODOLOGY

The complete process for the implementation of features dynamic tracking consists of some steps prior to the main algorithm, these steps are performed for increasing the final efficiency of the algorithm. Within the methodology includes the construction of a physical prototype was adopted to obtain videos that have consistent characteristics. At the end of processing as a result is obtained a text file that contains relevant information of the movements that can be analyzed in the image sequences. Complete methodology used is shown in Figure 1:

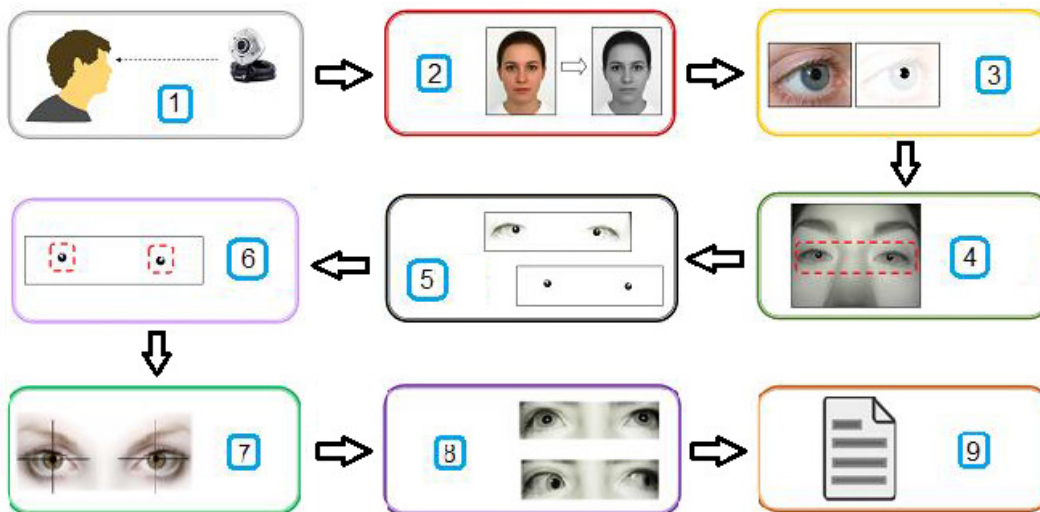


Figure 1. Metodology

The steps observed are: 1 obtaining the videos with the physical prototype, 2 grayscale conversion of the frames that make the video, 3 pupils segmentation, 4 determining the search area, 5 removing distractors, 6 tracing patterns or areas of interest, 7 determining positions of the pupils (x, y), 8 recording movements and 9 generating a text file with relevant information.

5. DEVELOPMENT

5.1. Obtaining of the videos

To form the basis of test videos a physical prototype was built with features that favoring the shots. Figure 2 shows a diagram of this prototype.

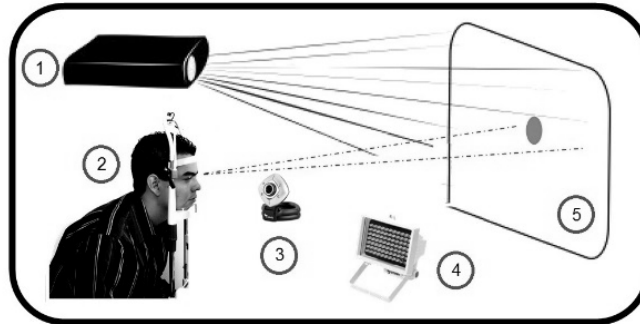


Figure 2. Physical prototype.

The elements constituting the prototype are: 1 projector, 2 chin base for persons to prevent movement of the face, 3 infrared camera, 4 infrared light and 5 animation reflected on a surface in front of the persons, this animation is controlled and allows different types of eye movements and at different speeds. The use of an infrared camera and light of the same technology provides a video with better features for processing.

5.2. Grayscale conversión.

To reduce the complexity of processing that is needed it performs the conversion of the images obtained in the first step to their equivalent grayscale because original images are in a format of three values (RGB red, green and blue for its acronym) and its analysis would require more resources.

The formula used is the average of the three original values of each pixel of the images.

$$(R + G + B) / 3 \quad (1)$$

5.3. Pupils segmentation

Segmentation is a process whose goal is to separate one or more areas of interest within the image of the background of it. Particularly in this case seeks to segment pupils for two purposes the first to know what position they are initially, this segmentation is done on the entire area of the first frame of the video sequence and the second objective is to find the centers of the pupils in the search area that is defined in step 4 of the methodology.

Thanks to the shots obtained in the first step can be segmented using thresholding method. This method is applied to the gray scale images in which the values for each pixel are between 0 and 255 so as to define a threshold value to section the areas of interest from the rest of the image.

$$Pixel(x, y) > threshold \quad (2)$$

The thresholding is shown in the Figure 3.

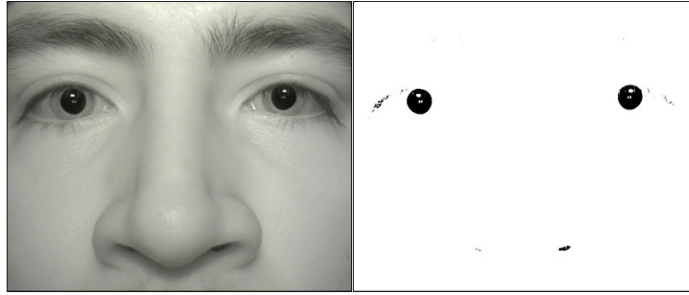


Figure 3. Pupils segmentation.

5.4. Determining the search area

Once the pupils have been segmented and it is known positions are found is possible delimit the region that later will be looking areas of interest. During the course of filter will not be an exhaustive search for the entire area of the images but only on the perimeter close to the coordinates obtained in the previous step. This is done by cropping the image. In Figure 4 shows the result of cropping.

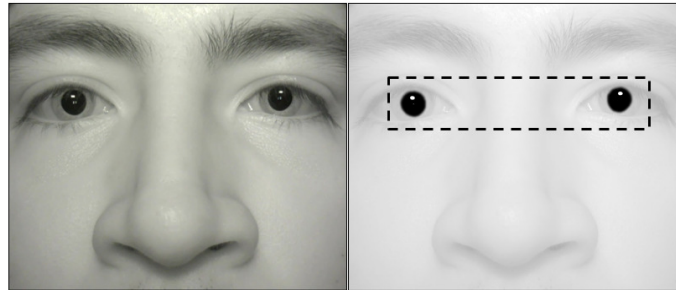


Figure 4. Delimitation of the search area.

5.5. Removing distractors.

In some special cases it may happen that the segmentation does not yield the desired result, for example when in a video appearing lashes with mascara is possible that at the moment of segement addition of the pupils display other zones in the same color range and could confuse the filter, so that these areas have to be removed. This is achieved using erosion and dilation filters. Equation (3) shows the dilation and equation (4) erosion. [3]

$$A \oplus B = \bigcup_{t \in B} A_t \quad (3)$$

$$A \ominus B = \{z \in E | B_z \subseteq A\} \quad (4)$$

Figure 5 shows the application of the two filters.



Figure 5. Pupils segmented (above), eroded (middle) and dilated (below).

5.6. Tracking areas of interest.

Once the distractions have been eliminated, have images where only seen areas of interest, the last thing to do is to determine the center of the two areas to know their positions in terms of coordinates and repeat steps 4 , 5 and 6 for subsequent frames. Figure 6 shows an sample of the features dynamic tracking filter.



Figure 6. Sample the filter.

A step that is added to the original method of the filter is for cases where the pupils disappear completely because of blinks, in these particular cases the algorithm restarts the process, in other words when the filter does not detect areas interest begins again with the segmentation process to find again the pupils as they appear.

6. EXPERIMENTATION

To determine the efficiency of the filter was used videos of people's faces with different features ranging from the length of the video, the eye movements recorded to some physical factors specific to the faces such as tone of the eyes, the kind of eyelashes and eyebrows, if they use makeup or not at the time of capture, among other features. The speed of the video is 20 frames per second. As the objective of the filter is to keep the focus on the pupils, the efficiency lies in not losing any of the position of those areas and the main reason that this problem is propitious by the eye blinking that is an involuntary movement and cannot be excluded from the videos. In Figure 7 shows some of the shots that were tested.



Figure 7. People with different characteristics tested.

The following are important features cases in which the filter was tested.

6.1. Experiment 1: pupils on dark color iris.

The filter process is shown in Figure 8:

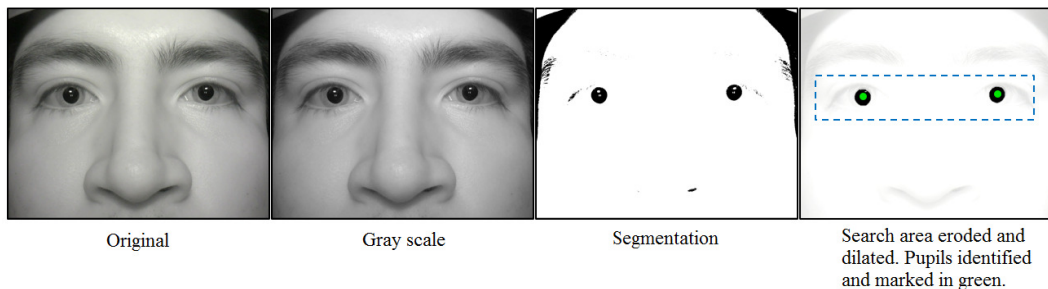


Figure 8. Filter results in people with dark color iris.

The filter tracks pupils with 100% success rate in the 13 videos that have this feature.

6.2. Experiment 2: pupils on light color iris.

The filter process is shown in Figure 9:

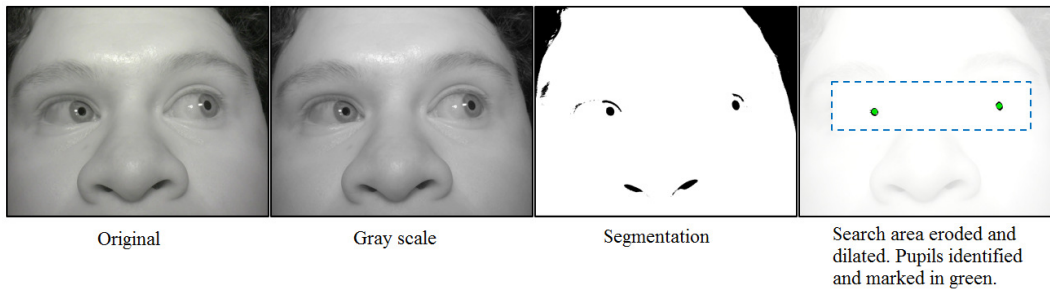


Figure 9. Filter results in people with light color iris.

The filter tracks pupils with 100% success rate in the 5 videos that have this feature.

6.3. Experiment 3: eye closed.

The filter process is shown in Figure 10:

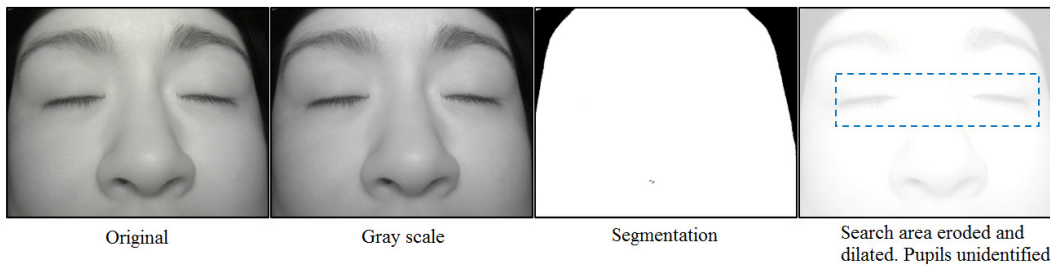


Figure 10. Filter results in scenes with eyes closed.

The eye closure is a movement that cannot be controlled, so inevitably there are scenes where see this problem. As shown in Figure 10 in the search area are not areas of interest so that in these cases, in the frames after the closing of the eyes the segmentation is done until the pupils appear again.

6.4. Experiment 4: down gaze.

The filter process is shown in Figure 11:

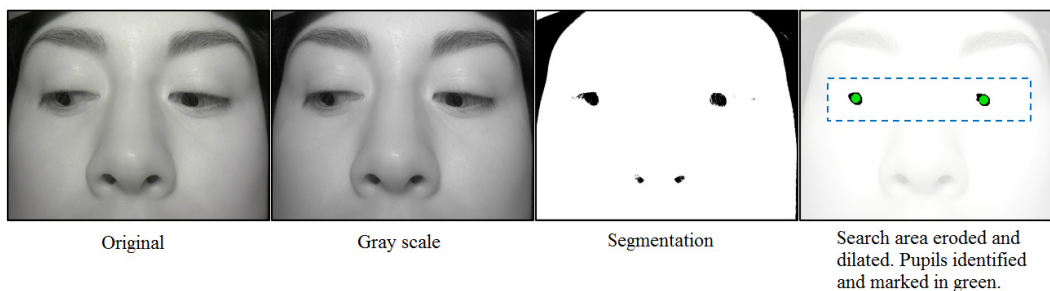


Figure 11. Filter results in scenes with down gaze.

It is important to analyze the response of the filter to this type of movement because in the down gaze the eyelashes appear above the eyes and can be confusing to the filter. In these cases it reaches 100% effective and this is thanks to infrared lighting and camera position (front of the person, at the height of the nose and eyes focused) because both features are implemented physically of the best way so that these cases are not an issue.

6.5. Experiment 5: lashes with mascara.

The filter process is shown in Figure 12:

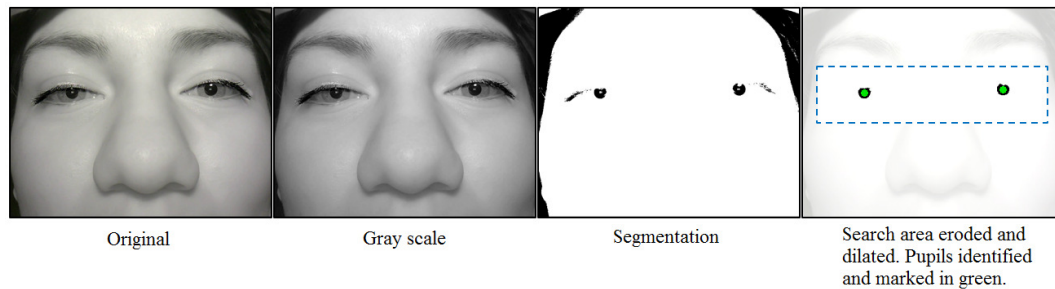


Figure 12. Filter results in persons with lashes with mascara.

The filter responded with an efficiency of 100% over the only video of a person using mascara on her eyelashes.

The selected cases are important to mention as they have different characteristics that could subtract the filter efficiency, may not be the only ones however to get the first reliable results are the most appropriate.

7. RESULTS

The filter was tested with a test basis consists of 18 videos of different people (age, sex, color of eyes and lashes with mascara). Are processed 20fps. There are two particular cases where the filter has the following behaviors: first when there is complete closure of the eyes stop scanning filter areas of interest and when appear again the filter recovers them immediately in the following frame starting again with the tracking process, the second case is when the gaze is downward, eyelashes interfere with pupils but the filter is able to recognize areas of interest without problem. The effectiveness is 98% due to the type of processing performed and the processing speed is 4.5 seconds for a video consisting of 1200 frames. The results can be seen in more detail in Tables 1 and 2.

Table 1. Velocity of processing.

Video	Duration	Frames (20fps)	Time
1	36 seg	720	2.64
2	56 seg	1120	4.10666667
3	49 seg	980	3.59333333
4	54 seg	1080	3.96
5	54 seg	1120	4.10666667
6	51 seg	1080	3.96
7	45 seg	900	3.3
8	50 seg	1000	3.66666667
9	50 seg	1000	3.66666667
10	47 seg	940	3.44666667
11	57 seg	1140	4.18
12	54 seg	1080	3.96
13	67 seg	1340	4.91333333
14	51 seg	1020	3.74
15	61 seg	1220	4.47333333
16	48 seg	960	3.52
17	64 seg	1280	4.69333333
18	58 seg	1160	4.25333333

Table 2. Comparison of errors and successes.

Comparison mistakes and successes			
Video	False positive	False negative	% Effectiveness
1	0	21	97.08333333
2	0	235	79.01785714
3	0	12	98.7755102
4	*	-	-
5	0	7	99.375
6	0	16	98.51851852
7	0	145	83.88888889
8	0	8	99.2
9	0	3	99.7
10	0	16	98.29787234
11	*	-	-
12	0	24	97.77777778
13	0	261	80.52238806
14	0	12	98.82352941
15	0	7	99.42622951
16	0	6	99.375
17	0	15	98.828125
18	0	152	86.89655172

8. CONCLUSIONS

The implemented filter throws an efficiency of 98%, in part due to the quality of the images obtained because they have features that facilitate processing also in cases of occlusion continues to respond adequately recovering well. As shown in the methodology applies two morphological filters (erosion and dilation) to increase tracing efficiency, usually dynamic tracking features are used in a manner not obligatory therefore constitutes a significant improvement to the process and results.

Addition to tracking efficiency, the filter shows acceptable processing speed because for the video with the highest number of frames (1220) the process has a duration of no more than 5 seconds.

This work can be used in the creation of a tool to support studies and other analyzes in the field of eye care.

9. FUTURE WORK

As part of the work that can be done later is to increase the videos base, with the aim of finding other cases to test the tracking response and try other filters in the methodology that can increase efficiency and speed of the application.

REFERENCES

- [1] Bernd, Girod (2013) "Morphological image processing 1", Stanford University
- [2] Gurjeet & Rajneet (2013) "Review on recent image segmentation techniques", Sri Guru Granth Sahib World University
- [3] Aksit, Kaan (2013) "Dynamic exit pupil trackers for autostereoscopic displays", Koç University
- [4] Majid, Tajeri (2012) "Real time eye tracking in unconstrained environments", Islamic Azad University
- [5] Swirski, Lech (2012) "Robust real-time pupil tracking in highly off-axis images", University of Cambridge
- [6] Mantink, Radoslaw (2012) "Do-it-yourself eye tracker: low-cost pupils-based eye tracker for computer graphics applications", West Pomeranian University of Technology in Szczecin
- [7] González & Wood (2012) "Digital Image processing", Prentice Hall
- [8] Yáñez, Javier (2010) "Tracking de personas a partir de visión artificial", Universidad Carlos III de Madrid
- [9] Platero, Carlos (2009) "Apuntes de visión artificial", Departamento de Electrónica, Automática e Informática Industrial

AUTHORS

Salvador Medrano Romero. Degree in computer science from the Technological Institute of Apizaco.



María Guadalupe Medina Barrera. Master of Science in Computer Science from the National Center for Research and Technological Development. Researcher of the Master in Computer Systems of the Institute Technology of Apizaco and faculty member PROMEP "Information systems".



José Juan Hernández Mora. Computer Engineering from the Autonomous University of Tlaxcala. Master of Science in Computer Science from the National Center for Research and Technological Development. Currently Head of Technologies Intelligent and Research Laboratory of the Institute Technology of Apizaco.



Luis Carlos Altamirano received his Ph.D. degree at National Polytechnic Institute (México) in 2002. His interest areas include: Computer Vision, Image Processing, Remote Sensing and Artificial Intelligence. He has participated in several different projects including: Oil projects and Computer Vision Projects. Currently, he is full time professor in the Computer Science Department at Autonomous University of Puebla and head of Postgraduate Department at the same Institution.

