



HAL
open science

Diagnosis of human operator behaviour in case of train driving: interest of facial recognition

Cyril Legrand, Philippe Richard, Vincent Benard, Frédéric Vanderhaegen,
Patrice Caulier

► To cite this version:

Cyril Legrand, Philippe Richard, Vincent Benard, Frédéric Vanderhaegen, Patrice Caulier. Diagnosis of human operator behaviour in case of train driving: interest of facial recognition. 30th European Annual Conference on Human Decision-Making and Manual Control, Sep 2012, Germany. 6p. hal-00993161

HAL Id: hal-00993161

<https://hal.science/hal-00993161v1>

Submitted on 19 May 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Diagnosis of human operator behaviour in case of train driving: interest of facial recognition

Cyril Legrand, Philippe Richard, Vincent Benard ; IFSTTAR – ESTAS, Villeneuve d'Ascq, France
Frédéric Vanderhaegen, Patrice Caulier ; LAMIH UMR CNRS 8201, Valenciennes, France

Abstract:

This paper shows a state-of-the-art of facial recognition with reliable algorithms more particularly with Viola and Jones contributions [3] based on the use of Haar-like features [4] suited to temporal constraints of rail drive. Estimations of eyes and mouth positions are provided in relation to face. The operations are done on pixels of those pictures in order to determine the blinking and the yawns betraying a sleepiness of the operator. The definition of neutral and surprise emotions corresponding respectively to normal or degraded situation permits to test the algorithms.

Keywords: human factors, human stability, facial recognition, railway safety

1. Introduction

Thanks to technological progress with regards to systems safety, accidents/incidents due to technical failures become less frequent. In spite of this, serious or even catastrophic accidents continue to occur in transportation field such as the train collision at Hal in Belgium [1].

It highlights, through the feedbacks (REX data), the essential role of human operator for this kind of accidents/incidents in the field of guided systems in particular transport. This role can be harmful in the case of human error (violation or unconscious error) or beneficial like in the accident of TER Lille-Rouen at Saint Laurent Blangy [2] underlining the inconsistent of some safety procedures. That is why, it is necessary to consider human factors, not only, during the design of a guided transport system but also throughout its life-cycle in order to evaluate its safety and improve its performances.

In this context, a diagnostic model of human operator behaviour is proposed and allows to attain these targets. In order to observe precursor signs of emotional state changing, the facial recognition can be used.

2. Using facial recognition in the service of railway safety

The first application of the face recognition is the people identification, used, for example, by the police in order to compare faces with a database of criminals and arrest them. This technology has greatly increased in recent decades.

It is developed in a lot of field like telephony with the famous smartphone iPhone of Apple. In this case, it will be applied in transports. Some manufacturers like BMW or Mitsubishi Motors (EMIRAI Concept car) have worked on cars with face recognition on-board in order to display information about driver's behaviour. Contrary to

automobile domain, the recognition topic is useful in railway transport for safety. The facial recognition is used for human operator impact analysis. The rate of facial expressions recognition is very high on a fixed position and in delayed time. Nevertheless, if the goal is to detect a change of emotional state, it must be realized on line i.e. during a driving simulation. This diagnostic will be done in four steps.

Step 1: face detection (Viola and Jones algorithm)

The Viola and Jones algorithm [3] is very used in face recognition. In fact, this method is more efficient than the others similar. The detection can be done in real time. The procedure is based on the classification of pictures. It is logical to think that it can be possible with computing sums of pixels but Viola and Jones, for constraints of computing speed, use features (see Figure 2-1) based on Haar-like features [4]. To compute them, some masks are used. They scan a picture and define the difference of sums of pixels of areas inside rectangles to find a change of texture or a contour.



Figure 2-1 : Example of masks used in Viola and Jones algorithm

In view of quantity of data (180 000 features for 24x24 test window resolution), they propose to compute features on the basis of a new representation of a picture: the integral image.

Thanks to the integral image, the features are computed faster than pixel-by-pixel computation.

Thus, a picture containing a face can be extracted. It is advisable to convert into grayscale to improve the filter's efficiency used in step 2.

Step 2: facial expression retrieval

The Figure 2-2 shows the different steps of facial expression retrieval. Three areas are defined: LE for Left Eye, RE for Right Eye and Mouth.

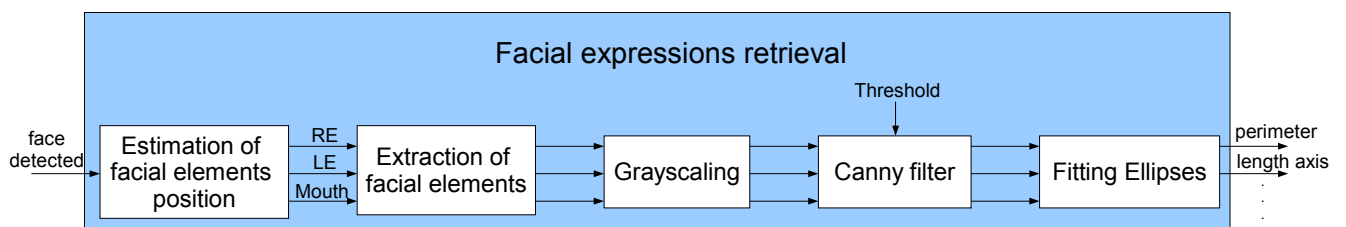


Figure 2-2 : Facial expressions retrieval

At this moment, a face is detected but the facial expressions must be extracted from this face.

To achieve this goal, an estimation of face elements (see Figure 2-3) based on Viola and Jones algorithm results is used:

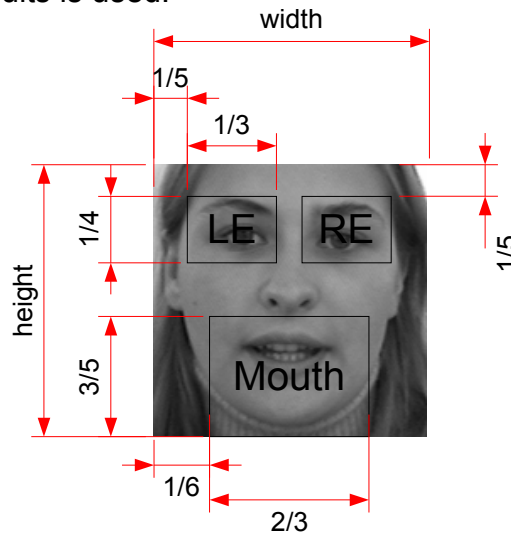


Figure 2-3 : Estimation of face elements

The data are given in relation with the height and width of the picture.

This estimation of this areas' location is based on a typical face. It is simply a mean relative position of eyes and mouth in relation to different faces.

After converting the picture into grayscale, these areas are extracted from the detected face and several treatments are used on each area.

The Canny filter (see Figure 2-4) is a contours detector often used in computer vision. This filter is chosen for its performances. Other filters exists (Sobel, Prewitt, ...) but they are less efficient.

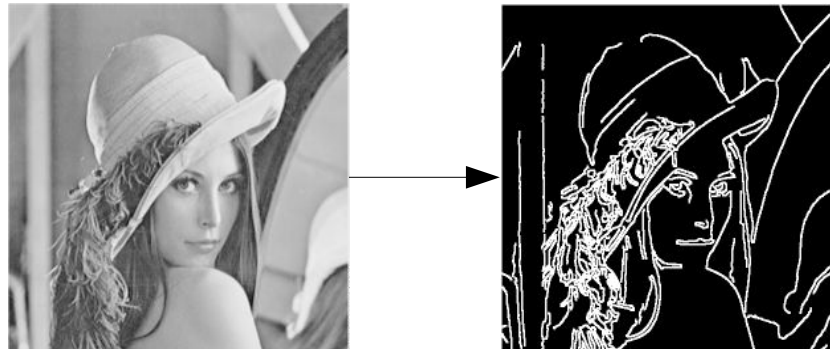


Figure 2-4 : Canny filter in operation

This filter is applied to LE, RE and Mouth areas. Then, some ellipses are drawn on these areas depending of the detected contours (see part 3 for results). The mouth and eyes shape is similar to an ellipse. With these ellipses, the features (perimeter, length of axis, angles, etc...) of eyes and mouth can be computed easily.

Step 3: interpretation of facial expression

With the distances compute in step 2, facial expressions can be founded. Regarding to eyes, it has to distinguish between simple blinking and sleepiness. The blinking is an unconscious and necessary function of eye which consists of open/close of eyelid

in order to avoid eye dryness. It can be detected and misinterpreted. On one hand, contrary to this, the beginning of a drowsy state is a long phenomenon; the upper eyelid slowly closes. On the other hand, an increasing of the frequency of blinking can indicate that the operator go into a sleepiness state.

Like eyes, an ambiguity is possible between yawns and speech. Indeed, both are defined by an opening of the mouth. A yawn can be distinguished between speech by the opening amplitude of the mouth and the opening time of this action. It is a long and large opening whereas the lips movements are more variables during a speech. However, in the train driving context, the operator is not supposed to speak in the cab except if he/she is called. In this case, he/she places the microphone in front of his/her mouth and it is not detected anymore.

Step 4: diagnostic of impact on railway safety

The sleepiness is one of the states of human operator which is the most critical during rail driving. The surprise state is important: it can be signified that the operator has seen something whereas the system can not detect. So, it is interesting to know the state of the two entities to ameliorate safety. However, detect these human states is not enough. This impact on the railway safety must be known. To better understand it, the concept of human stability [5] will be used. It is characterized by several states (stable, instable, indeterminate) and transitions between two states (leap, break, indeterminate). Stability sequences can be detected regarding the occurrence of a disturbance. Two kinds of stability were defined: recommended signature (in the case of the disturbance can be control) and risked signature (in the contrary case).

The human stability criteria are clearly asserted and classified in three facets: State, Performance, and Behaviour. The object of this study is focused on the facet State. In fact, the monitoring of variations in facial expression allows to define a leap or a break of stability in terms of emotional state like sleepiness.

To apply these different steps, experimentations have been led.

3. Experimental results

To carry out these different algorithms, a library of functions for real time computer vision, OpenCV, is used [6].

The experimentations are divided in two steps. First, the BioID database [7] is used to test the face detection. This database contains 1521 grayscale pictures of 23 different people (men/women, wear glasses or not, bearded or not...) with a 384x286px resolution. These pictures were caught with different room's brightness and distance between video camera and people.

Table 3-1 : Detection results

	Fault detections	No detection	Correct Detection
Results	3,3 %	3 %	93,7 %

These results coincide with detection rates advanced by Viola and Jones [3]. Throughout this test, some factors have been identified:

- If a part of a face is hidden or not be in shot, the face is not detected,

- If the face is over/underexposed (too and not enough bright), the face is not detected,
- Some elements of decor can be associated with a face (some inscription on a wall...) corresponding with fault detections in the table 3-1.

The second part of the experimentation is the detection and the interpretation of facial expressions. It is based on videos of some willing people from IFSTTAR, The French Institute of Science and Technology for Transport, Development and Networks. Before establishing a diagnostic on line, the tests are led off line and, now, out of train driving context. The subjects simulate the different emotions.

The Distance between EyeBrow and Eye (Debe) and the Diameter of the Mouth (DM) are measured.



Figure 3-5 : Example on an eye

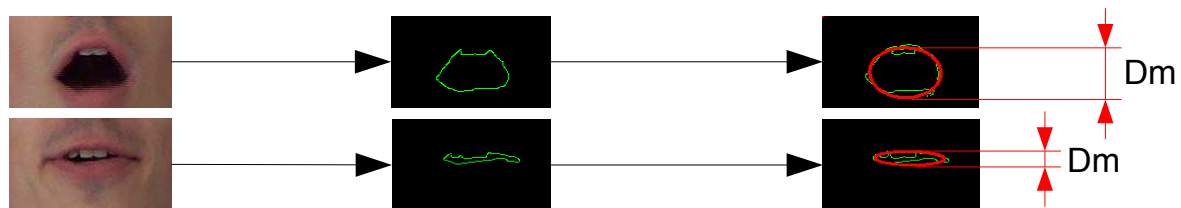


Figure 3-6 : Example on a mouth (weak and strong opening)

Debe is computed by the center of ellipses corresponding with eyebrow and eye. Dm is simply the axis of the ellipse and corresponds with the opening of the mouth.

Debe and the perimeter of eye – corresponding to the contour drawn by upper and lower eyelid – are observed. If a decreasing then a constancy of these features are noticed, a sleepiness state can be interpreted. Inversely, an increasing can signify that the operator is in a surprise state.

4. Conclusion and perspectives

The facial recognition is a possible method to detect changes of human operator states in terms of emotions. It is a good solution to observe external signs of intrinsic states. Nevertheless, the results coming from this method are not enough. In addition to the fault detections and some people characteristics (wear glasses, haired people, etc), each operator do not react in the same way. This facial expressions recognition can bring some clues to identify states of human stability but it must be a complementary method to reach this goal.

Other non-intrusive means of detection exist: measure of breath and heart rates, infrared for the body temperature. A fusion of all output data is necessary and a multi-criteria study must be led.

This base of human factors impacts diagnostic allows to identify the state of the human operator. The operator and the system do not apprehend the situation in the same way. Endsley [8] defines in her work the concept of situation awareness (see

Figure 4-7).

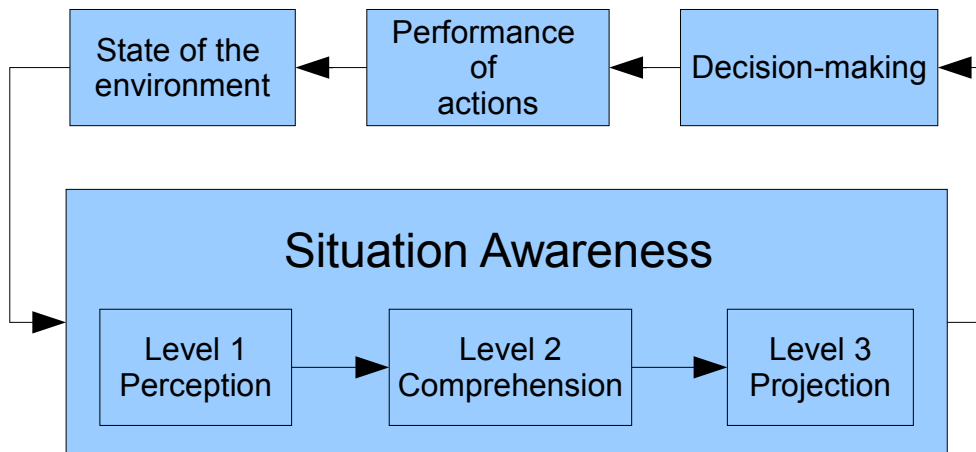


Figure 4-7 : Simplified Situation Awareness model

The Situation Awareness is divided in three steps: first the human operator perceives a given situation, then, comprehends it and, finally, acts in consequence. His/her representation of the situation is able to give some information of an emotional state. A perception can correspond with a stability signature [5] given and leads to an action.

At this moment, warnings can be displayed on the operator interface. Some barriers can be put between the moment when the operator senses the situation and he/she acts.

5. References

[1] BEA-TT, (2006), Rapport d'enquête technique sur la collision survenue le 9 Juin 2005 au passage à niveau 83 à Saint Laurent Blangy, *Accident Report*.

[2] ERA, (2010), Railway Safety Performance in European Union, *ERA technical report*.

[3] P. Viola and M. Jones, (2001), Rapid Object Detection using a Boosted Cascade of Simple Features, In *International Conference on Computer Vision*.

[4] C. Papageorgiou et al, (1998), A General Framework for Object Detection, In *International Conference on Computer Vision*.

[5] P. Richard, (2012), Contribution à la formalisation et à la détection de la stabilité humaine au regard de la sécurité : application aux systèmes de transport guidé, *mémoire de thèse, UVHC, Janvier 2012*.

[6] Open Source Computer Vision (OpenCV). <http://opencv.willowgarage.com/wiki/>

[7] BioID face database. <http://www.bioid.com/downloads/software/bioid-face-database.html>

[8] Endsley, (1995), Toward a Theory of Situation Awareness in Dynamic Systems. In *Human Factors*.