Study and Prediction of Human Behavior Based on Face-Recognition

Christophe Lwanyi Ashimalu[1], Simon Ntumba Badibanga[1], Pierre Kafunda Katalay[1], Joséphine Bukaya Tshibangu[2], Alphonse-Marie Gashita Mayimbi[3]

[1]Department of Mathematics, Statistics and Computer Science, University of Kinshasa, Kinshasa, Democratic Republic of the Congo
[2]Department of Education, University of Muene Ditu, Muene Ditu, Democratic Republic of the Congo
[3]Department of Education, National Pedagogical University, Kinshasa, Democratic Republic of the Congo
Email: abbechristophelwanyi@yahoo.fr

Abstract

The question relating to the Detection and Prediction of human behavior using Artificial Intelligence tools is the main focus of our research. More specifically, we have studied the link between emotion and human behavior in order to model an artificial intelligence that is not only capable of detecting but also, and above all, predicting human behavior. Predicting what someone is about to do next based on their body language is natural for humans, but not for computers. When we meet another person, they may greet us with a handshake or a fist bump. However, we don't know which gesture will be used, but we can read the situation and react appropriately. In a new study, researchers at Columbia Engineering unveil a computer vision technique to give machines a more intuitive sense of what will happen next by taking advantage of higher-level associations between people, animals and objects. In order to validate the scientific reliability of the results emanating from the various theories, this article will follow a constructivist approach. Constructivism is an epistemological position based on the relativity of the notion of truth or reality. Reality is defined by the representation of a subject's experience of reality. As Edgar Morin so aptly put it, "subject and object are indissociable, but our way of thinking excludes one through the other" [1]. This is the "dialogical" principle, the maintenance of duality within unity. Constructivism is based on subject-object interaction, with research "no longer defined by its object, but by its project" [2]. The model we have chosen, constructivism, is a reliable one, especially as it takes into account the three elements needed to assess the quality of an artificial intelligence model. The training data must be of good quality, the algorithm chosen must be relevant and robust, and the prediction error of the model generated must be as low as possible. In this article, we have shown that there is an intrinsic link between

emotion and human behavior. Thus, understanding emotion would increase the possibility of success in predicting the behavior that should follow. Through literature survey and data analysis in related fields, it is found that emotion acquisition will be carried out using a facial feature algorithm for point capture and combined with machine learning for detection, analysis and prediction. This is why, when studying an individual's facial features, it is possible to combine several techniques such as the *HoG* vector, the convolution matrix or the Gabor filter in order to identify the essential features of the face. This will enable us to determine the individual's psychological state and, consequently, predict the possible behavior they may display. Applying this to the field of security would seem to be one of the best ways of curbing crime, especially in public places. To do this, you need the logistics to capture several photos and process them in real time.

Subject Areas

Artificial Intelligence, Machine Learning, Psychology, Sociology

Keywords

Detect, Predict, Face Recognition, Constructivism, Human Behavior, Emotion, Interaction, Gabor Filter, Convolution Matrix

1. Introduction

Artificial intelligence is as old as humanity itself. Many myths tell of stories and rumors of artificial beings endowed with intelligence or consciousness by master craftsmen. Yet the seeds of modern Artificial Intelligence can be traced back to the work of philosophers who attempted to describe the human thought process as the mechanical manipulation of symbols.

In an age when we are increasingly surrounded and invaded by communication devices, applications, chatbots and/or virtual assistants, robots; it has become essential for IT tools (machines) to be able to communicate emotions if our relationships with them are to be optimal.

Emotions drive every aspect of our existence: our well-being, our health, our relationships with others and our decision-making all depend on our emotional state. People with very high emotional intelligence can build better relationships with others, are more reliable and more convincing.

This article will attempt to build an epistemological bridge between computer tools and human behavior, while circumscribing the latter within an emotional framework.

2. Development

2.1. Behavior

Behavior can be broadly defined as a way of being, acting and/or reacting of

human beings, of a group, of animals, it is not only an attitude, a behavior, it is also a way in which something functions, evolves in certain circumstances. For the purposes of this study, we will use just one definition. Human behavior can be defined as all the physical actions and emotions associated with an individual or even a social group. In other words, they are the responses to the most varied stimuli in daily life, whether internal or external in origin. There are several types of behavior worth mentioning: innate, acquired, associative, private and public.

However, there is an intimate link between behavior, thought and emotion. An emotion can be defined as our brain's response to the events that surround us. This response is pre-programmed and will generate physiological states, which in turn give rise to specific behaviours: anger, joy, disgust, sadness, ... This is what is shown in Figure 1.

2.2. Emotions

Our emotions are an integral part of our daily lives: they influence our decision-making, but also the way we remember our past experiences and the way we eat! But what are they? Is there really a link between emotion and human behavior? Views are divided depending on which school of thought you belong to.

Clearly, an emotion is a reaction that human beings experience in response to events or situations. The type of emotion a person may experience is determined by the circumstances surrounding its triggering. For example, people feel joy when they receive good news. A person feels fear when threatened [3].

According to several studies, emotions have a strong influence on our daily lives. We make decisions based on whether we are happy, angry, sad, bored or frustrated. We choose activities and hobbies based on the emotions they arouse. Emotions are therefore mental and physiological states associated with psychological changes, internal or external stimuli, natural or learned.

Figure 1. Link between behavior, thought and emotion.

According to most modern theories, emotions are a multi-component process, *i.e.* articulated in several components and with an evolving time course [4].

Understanding emotions can help us to navigate through life with greater ease and stability, and above all to predict certain behaviours.

2.3. Types of Emotion

In their book, Discovering Psychology, authors Don Hockenbury and Sandra E. Hockenbury suggest that an emotion is a complex psychological state involving three distinct components: a subjective experience, a physiological response and a behavioral or expressive response [5]. The definition of any psychological entity usually presents major difficulties, and the concept of emotion is far from being an exception to this rule.

A particular problem in the quest to define emotion arises from the fact that statements often relate to only one aspect of emotion. In fact, the concept of emotion is used differently depending on whether it is considered in reference to the stimulus aspect, the subjective experience, a phase of a process, an intermediate variable or a response. Our aim is not to get lost in the scholarly and/or partisan definitions of emotion. But rather to understand the concept of emotion in order to highlight the intrinsic link between it and human behavior.

As well as trying to define what emotions are, researchers have also tried to identify and classify their different types. Descriptions and ideas have changed over time, so here we would like to look at a few types of emotion, with an emphasis on the link between them and human behavior.

- In 1972, the psychologist Paul Ekman suggested that there are six basic universal emotions in human cultures: fear, disgust, anger, surprise, happiness and sadness [6].
- In the 1980s, Robert Plutchik introduced another system for classifying emotions, known as the wheel of emotions (Figure 2). This model demonstrated how different emotions can be combined or mixed together, much in the same way as an artist mixes primary colors to create other colours [7].
- In 1999, Ekman extended his list to include a number of other basic emotions, including embarrassment, excitement, contempt, shame, pride, satisfaction and amusement [6].
- In fact, for decades, several theories have been formulated to explain what an emotion is. The aim here is not to present a complete list of the major currents, but rather to mention a few representatives, highlighting the key ideas underlying each classification.

2.3.1. Emotions, the Evolutionary Perspective

Evolutionist theories, very popularly defended by authors such as Cosmides & Tooby, Izard, Nesse and Öhman, follow on from Darwin's work on emotions, which can be found in his 1872 book entitled The Expression of Emotions in Man and Animals [8].
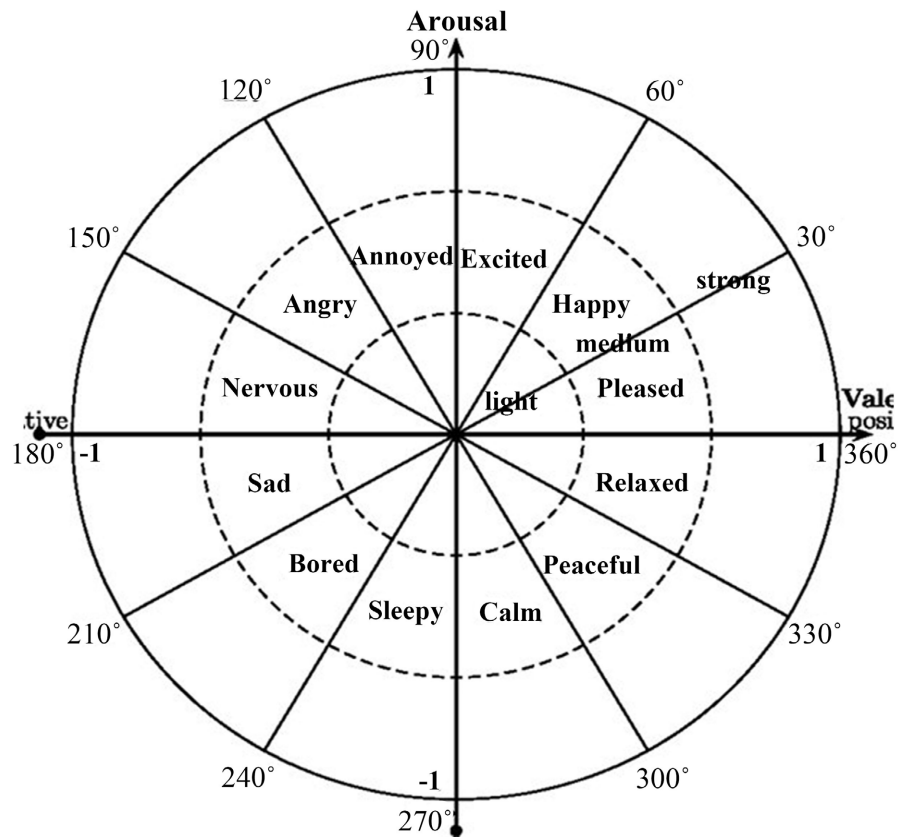
Figure 2. Robert Plutchik's wheel of emotions.

According to Darwin, emotions are biologically and phylogenetically determined. They are the result of successive adaptations of the organism to its environment. So, whatever the animal species, including man, when they are subjected to the same situations, they set up the same expressive patterns. Darwin defends the idea of the universality of emotional expression, according to which the same emotion, such as anger, would be expressed in the same way regardless of culture or species.

Expressions of emotion, like reflex actions or instincts, are part of innate and involuntary behavior, which, however, was first acquired. The transformation of acquired behavior into innate behavior therefore depends on the key notion of habit, which is inseparable from that of heredity [9].

It should be noted that for the authors Tooby and Cosmides, emotions develop in response to different sets of recurring situations. Emotional expressions, such as facial movements, are only produced for certain emotions. These are emotions whose spontaneous and automatic expression would have had an evolutionary interest in the survival of the species [10].

Ekman even went so far as to say that facial expression is the linchpin of human communication [11]. Knowing how to read faces makes social relations easier, and misinterpreting a facial mimic can lead us to adopt behavior that is ill-suited to the situation.

In monkeys, when a dominant male chases another male and the latter makes a grimace (expression of fear), the dominant male will stop chasing him. Conversely, if the dominant male makes the same grimace, he expects the subordinate male to come and kiss him. In this sense, facial expression informs the individual of our intentions but also of the behavior we expect from him.

2.3.2. Somatic Theory: James-Lange

The James-Lange theory is a hypothesis on the origin and nature of emotions. It is one of the first theories of emotion in modern psychology. It was developed by the philosopher John Dewey and named after two nineteenth-century scholars, William James and Carl Lange [12] [13]. The basic premise of the theory is that physiological arousal causes the experience of emotion [14].

Instead of experiencing an emotion and a subsequent physiological (bodily) response, the theory proposes that the physiological change is primary, and the emotion is then experienced as the brain reacts to the information received via the body's nervous system. The author proposes that each specific category of emotion is attached to a unique and different pattern of physiological arousal and emotional behavior in response to an arousing stimulus.

In his 1884 article [15], William James maintains that feelings and emotions are secondary to physiological phenomena. As we go through different experiences, our nervous system develops physical reactions associated with them. The emotional reaction would therefore depend on how these physical reactions are interpreted.

Examples of these physical reactions include an increase in heart rate, trembling, stomach upset, ... These in turn generate other emotional reactions such as anger, fear and sadness [14]. This is the case of "an emotion-inducing stimulus (snake) which provokes a pattern of physiological responses (increased heart rate, accelerated breathing, …). This pattern is interpreted as a particular emotion (fear). This theory is supported by experiments in which manipulating a physiological state induces the desired emotional state" [16].

A Finnish study published on 31 December 2013 in the journal Proceedings of the National Academy of Sciences has produced the first body map of emotions. To produce this map, researchers conducted a study with 773 Finnish, Taiwanese and Swedish volunteers [17].

The volunteers took part in five experiments to test their sensory reactions to certain emotions: starting with a stimulus. The parts of the body in which their sensations were strongest were identified and mapped. In the first experiment, the participants listened to words in their mother tongue. In the second and third experiments, they watched images and films.

In the last two experiments, participants had to recognize different emotions from faces and body temperature maps. The results were averaged to produce the first body map of emotions. Here's a sample that sums up all these ideas (Figure 3).

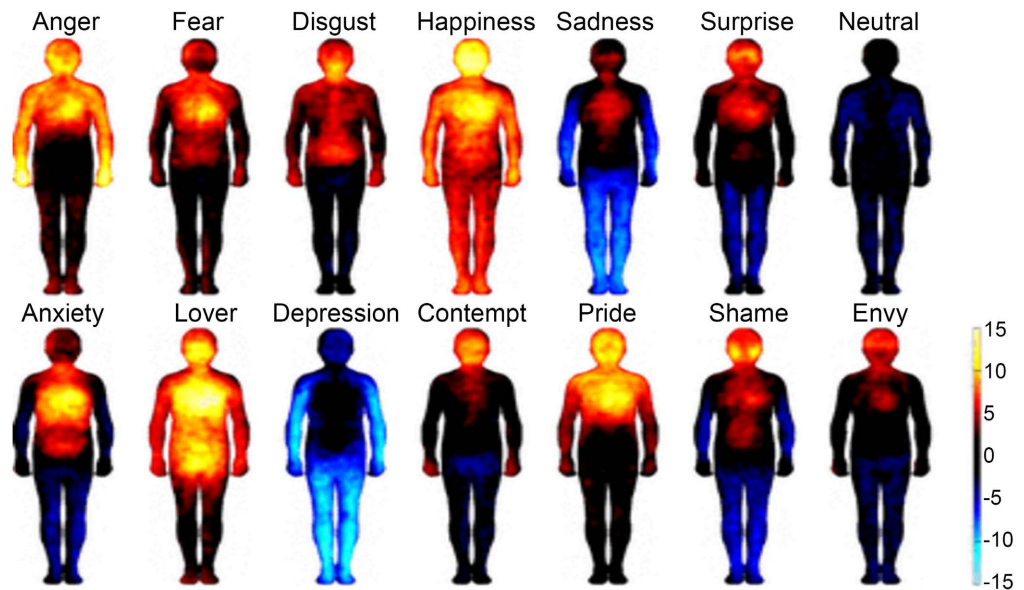**Figure 3.** Body map of emotions.

2.3.3. Psychological Theory: The Interruptive System

The American researcher Herbert Simon has distinguished himself in the fields of economics, organizational sociology, psychology, … At the same time, it is worth noting the concept of bounded rationality, which he invented and for which he was awarded the Nobel Prize in Economics in 1978. He is a specialist in cognitive psychology, who developed a theory in 1967 of the interruptive system of linear decision. He defined three groups of needs in real time for an individual:

1) Needs that arise in the face of uncertain events (sudden noise or visual stimuli) that could signal danger;

2) Physiological needs, which are internal stimuli such as hunger, thirst or exhaustion;

3) Cognitive associations, which are strong stimuli derived from memory associations, for example, the memory of a fear.

For Simon [18], "The procedure of rational calculation is interesting only when it is not trivial—that is, when the substantially rational response to a situation is not immediately obvious. If you put a quarter and a dime in front of a subject and tell him that he can take one or the other, not both, it's easy to predict which he'll choose but difficult to learn anything about his cognitive procedures".

The procedure becomes important to study when the agent does not have complete information. In this case, it cannot find the optimal solution and will stop searching for information when it has found a solution that satisfies its needs (satisficing) [19]. If the study of procedures and organizations is important for Simon, it is because, despite everything, it is important to make the best possible decisions and therefore to follow processes that lead to the solution closest to the optimum (Figure 4).
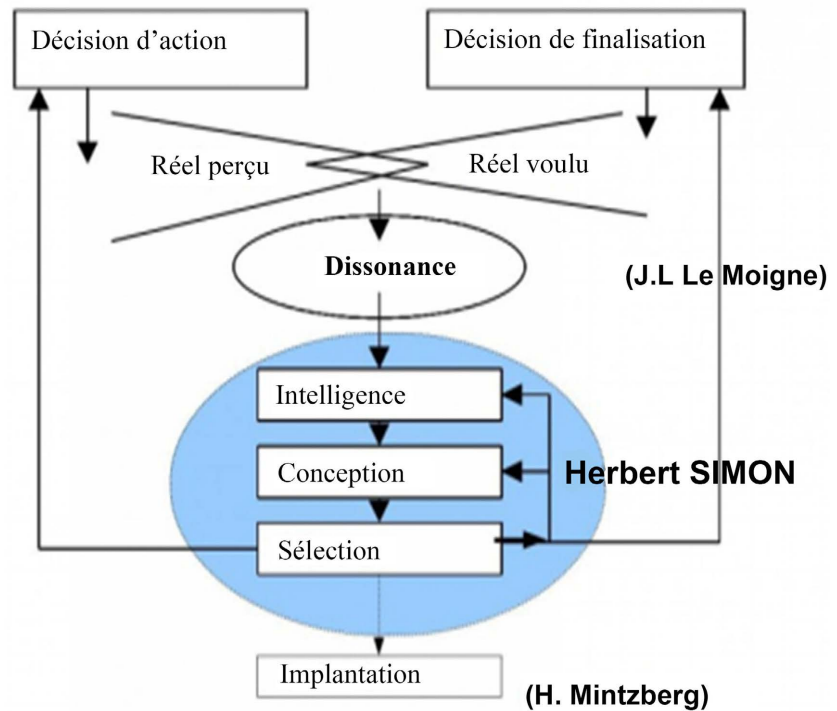
Figure 4. Herbert Simon's model: procedural rationality of decision making.

2.3.4. Neurobiological Theory

According to this theory, emotions are thoughts that are linked to certain activities in areas of the brain that manage our attention, motivate our behavior and determine the meaning of what is happening around us. Pioneering work by Paul Broca, James Papez and Paul D. MacLean suggests that emotions are associated with a group of structures in the center of the brain called the limbic system, which consists of the hypothalamus, the cingulate cortex, the hippocampus and other structures. More recent research has shown that some of these structures are not as dedicated to emotions, while some non-limbic structures are more involved in emotions.

For example, guilt and shame (two related emotions) share certain neuronal networks located in the frontal and temporal lobes of the brain, but their structures are distinctly different. Guilt arises when our behavior conflicts with our conscience. Shame is felt when we think we have damaged our reputation [20]. In an fMRI study, scientists at Ludwig Maximilian University in Munich found that shame triggers strong activity in the right side of the brain but not in the amygdala.

In contrast, in a state of guilt, there is an increase in activity in the amygdala and frontal lobes, but very little neuronal activity in either cerebral hemisphere. The researchers concluded that shame, with its socio-cultural factors, is a more complex emotion, whereas guilt is only linked to social norms learned by the individual [21].

Other emotions such as fear and anxiety have long been characterized as being formed exclusively by the most primitive brain areas and associated with the

fight-flight response. They have also been considered as expressions adapted to defensive behavior when a threat is encountered. Although defensive behaviours have been observed in a wide variety of species, Blanchard *et al.* found equivalence in the expression of defensive behaviours and their function in humans and mammals [22]. Figure 5 is an illustration explaining how the brain processes information.

As soon as a potentially dangerous stimulus is presented again, certain brain structures repeat the same reaction (hippocampus, thalamus, ...). The amygdala therefore plays an important role in generating the next behavioral response. Notably in terms of the neurotransmitters that respond to the dangerous stimulus. These biological functions of the amygdala are not limited to fear conditioning or the processing of aversive stimuli, but are also observed in other processes. This may define the amygdala as a key structure for understanding potential behavioral responses to danger in humans and mammals more generally [23].

Neurobiology is a branch of biology that studies the structure and function of the nervous system. Often referred to as neuroscience, it brings together several medical and scientific fields. One of the major challenges of neuroscience is to understand human cognition. The discoveries made are leading to a better understanding of psychiatric and neurological pathologies such as schizophrenia, autism, multiple sclerosis, epilepsy, Parkinson's and Alzheimer's disease.

Researchers are also studying the link between obesity (or bulimia) and the brain, aware of the neurobiological dimension of diet. Mastering all these interrelationships will give us a better understanding of human behavior and enable us to predict its outcome using Artificial Intelligence tools.
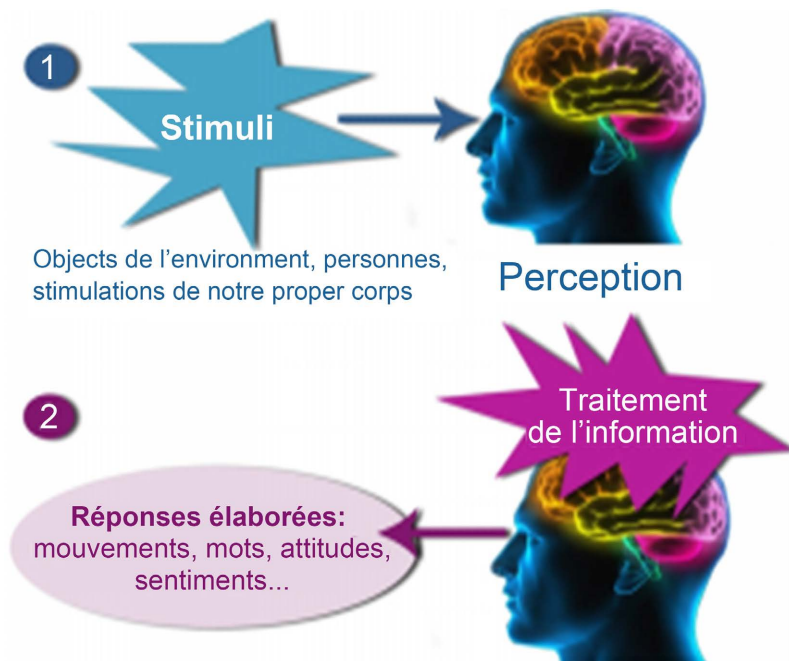

Figure 5. Information processing by the brain.

2.3.5. Emergence of This Field

Understanding mass feelings is the first step, the logical next step being the ability to refine the analysis at the level of the individual. Here too, technology is proving to be a very useful ally. In the future, biosensors or connected biological sensors will provide new information. This will make it possible to complete and refine our knowledge of an individual's emotions.

A biosensor could measure body temperature, iron levels in the blood and perspiration levels, with the additional data then enriching the visual (face) and sound (tone of voice) sensors to refine the analysis of emotions. One application could be, for example, hospital nurses being able to monitor not only the physiological state of patients, but also their emotional state, without any prior interaction; and so being able to adjust their behavior and tasks according to the overall state of the patient.

New technologies could feed the algorithms with more accurate and precise data. For individual data, Endotronix has developed a solution for remotely measuring and monitoring the state and behavior of the heart. As for external data, the French company Alpha MOS and the Californian company Aromyx, for example, are building electronic noses. This technology could ultimately reshape the perfume and food industries.

The same technology could also be used for security. And this is what this work seeks to demonstrate in the lines and chapters that follow.

2.3.6. Facial Expressions

According to the specialist site, *Expressions faciales et micro-expressions* [24], there is a question on everyone's mind: how do we know if facial expressions are reliable? In order to detect lies, we need to be aware that the face contains three messages [Ibidem], the quintessence of which is as follows:
-   Authentic information,
-   That which the liar wishes to show,
-   The information that the liar wishes to conceal.

In view of the above, it is only right to point out that the difficulty will lie in decoding genuine facial expressions from those that are false. Facial expressions generally occur spontaneously, sometimes even without the individual being aware of it.

It is also very important to note that facial expressions are not solely dedicated to or related to emotions. There are in fact thousands of facial expressions that underline what is being said, which are illustrative gestures such as winking, or raising one's eyebrows to nod in agreement, as if to say: ah, yes, I see!

As you can see, the jungle of facial expressions is so dense that it deserves special study and attention in order to avoid making errors of judgement. So before any interpretation, whether it's a gesture, a posture or, of course, a facial expression, you should never forget to refer to the baseline.

At this stage, it's essential to review a number of emotions, known as micro expressions. Facial micro expressions are quick flashes of emotion (lasting less

than half a second) that appear when the speaker wishes to conceal an emotion or when they are experiencing a high-stakes situation (when, for example, they are in a situation where they have a lot to lose, a lot to win, when they are under a lot of stress, …). Figure 6 displays seven types of micro expressions.

1) Joy—easy to recognize, difficult to fake

Joy can be expressed towards others or kept to oneself. In this case, the face is tilted slightly downwards. According to the author of Lie to Me [25][1], it may be trivial, but when a person feels joy, they smile.

The opposite, *i.e.* trying to make someone believe you are happy, is much more difficult. A deliberately distorted smile often appears in this type of situation, when you want to please someone and not offend them. But it can also occur when you disagree with someone and pretend to disagree.

Sometimes, in the opposite situation, we can observe a micro expression of joy, when it should not normally appear. This expression arises when someone is proud of themselves after a reprimand, for example.

The clues: the smile and the wrinkles around the eyes, proving that the smile is genuine (Figure 7).

2) Sadness—it sags

The author goes on to say that it is sometimes important to be able to spot the sadness behind a smile. Whether it's to get a sympathetic ear or to spot the distress in which someone close to you finds themselves. Many people may pretend to be happy because they don't want to spoil the good mood of those around them.

However, it is sometimes necessary to observe this sadness and provide the appropriate help. Even in the midst of a conversation, turning sadness into joy can give the discussion a new impetus.

Clues: the upper eyelids and outer corners of the eyebrows droop. The person will give the impression of staring into space. The corners of the lips also tend to droop (Figure 8).

Figure 6. Facial micro expressions (7).

[1]Lie to Me is an American television series produced from 2009 to 2011. Created by the same producers as 24, the series stars Tim Roth as Cal Lightman, a psychologist who is an expert in non-verbal communication and infallible in understanding when someone is not telling the truth, and who uses this knowledge in the service of justice.
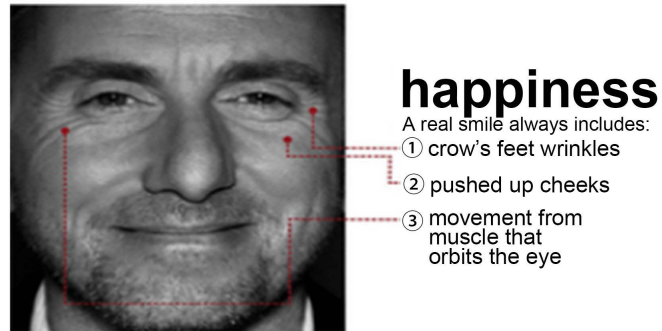
Figure 7. Happy face.

Figure 8. Sad face.

3) Anger (angry)

Anger is usually easy to spot. But sometimes it can come on all of a sudden in the middle of a banal conversation, shouting "BEWARE" [25], and with all the consequences that follow and which could have been spotted at first sight.

The clues: lips pressed together. Eyebrows pressed towards the nose. The wrinkles between the eyebrows are apparent and the most striking feature is the fixed gaze (**Figure 9**).

4) Contempt-the asymmetrical bar

It's the expression that makes people's hair stand on end. And yet, according to the author, it's the easiest to spot. And for good reason: it's the only one of the 7 emotions that is asymmetrical. This is clearly shown in the illustration. This expression is unilateral on either side (left or right). All the other expressions, whether anger or surprise, have characteristic elements that are symmetrical on the face along the vertical axis of the nose.

Clues: one side of the lip is slightly pulled back (or raised). This lip movement can be very subtle, appearing and disappearing very quickly (**Figure 10**).

5) Disgust

This expression on the other person's face is a very good clue as to whether the other person likes what you're saying or not [25].

The clues: the upper lip usually rises, revealing the teeth. This clue is usually accompanied by a wrinkle around the nose (**Figure 11**).
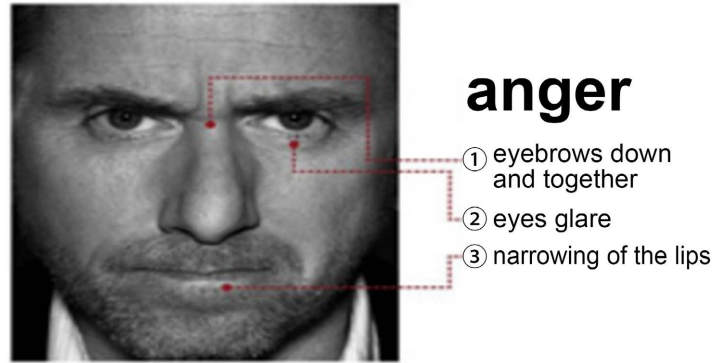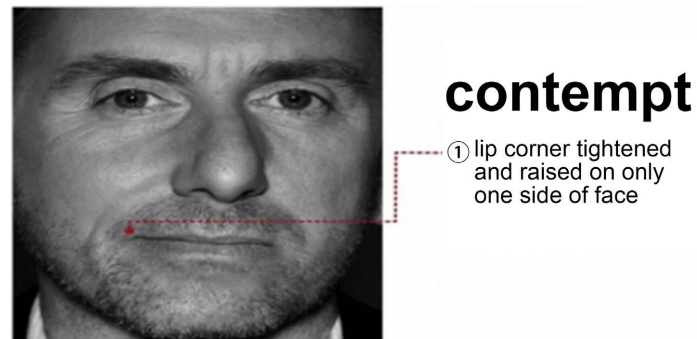
6) Fear (afraid)

anger

① eyebrows down and together

② eyes glare

③ narrowing of the lips

Figure 9. Angry face.

contempt

① lip corner tightened and raised on only one side of face

Figure 10. Face of contempt.
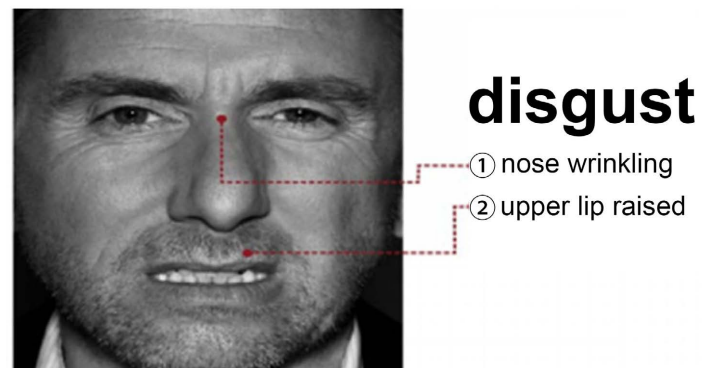

disgust

① nose wrinkling

② upper lip raised

Figure 11. Face of disgust.

It's an expression we all know. If you want to observe it, I simply invite you to surprise one of your friends. For a few milliseconds, his face will show the following expressions [25].

The clues: The lips stretch horizontally towards the ears. The lower eyelids are tense and the upper eyelids are raised. The eyebrows are also raised, as if they had arranged to meet (**Figure 12**).

7) Surprise

In the test, we often mix the two micro expressions: surprise and fear. The facial elements involved are similar and sometimes the contexts are the same. It's a very important micro expression for determining deception.

But what people don't realize is that simulated surprise is easy to identify. In fact, if someone pretends to be surprised, their facial expression will last more than a second. In contrast, genuine surprise is very quick (less than a second) before returning to a neutral state. It is very brief. If someone keeps their eyebrows up for more than a second, they are definitely lying [25].

The clues: the mouth is half-open, the eyes wide open and the eyebrows raised. Most importantly, it lasts less than a second (Figure 13).

3. Behavior Prediction

3.1. Face Recognition

Each of us has had the experience of not recognizing someone or not being recognized because of changes in the pose, because of illumination... As if to say that there are situations in which bodily expression makes the task of recognition difficult.

So it's not surprising that computer vision systems have similar problems. In fact, no computer vision system can match human performance, despite years of work by computer scientists around the world. The recognition that most closely resembles the human identification and/or authentication process is facial recognition.
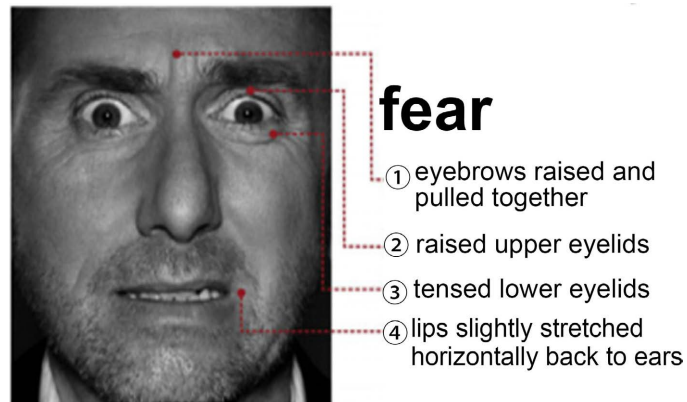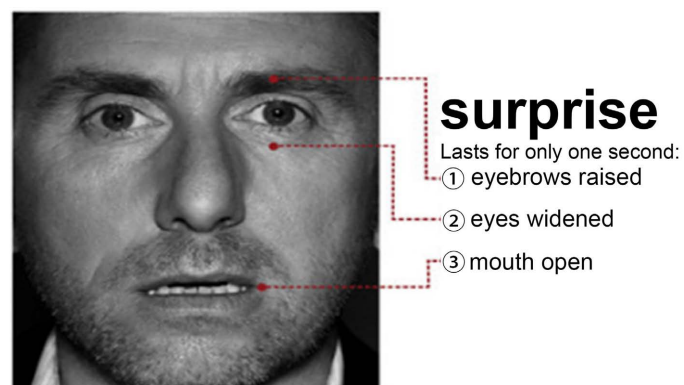

Figure 12. Face of fear.

Figure 13. Surprise face.

3.2. Position of the Problem

Facial recognition makes it possible to adapt biometric verification to all situations. It is a highly effective technology used in many security applications. Facial recognition is contactless and requires no special tools to use, which makes it ideal not only for identifying people in crowds or public spaces but also, and above all, for taking rapid salutary action in the presence of a prediction of behavior likely to endanger human life.

3.3. Biometric Analysis of Facial Data

Facial recognition seems to be the most natural biometric technology, because we recognize ourselves by looking at our face. Facial recognition systems are systems that identify individuals on the basis of facial features such as eye spacing, nose bridge, corners of the mouth, automated ears, chin, ... These features are analyzed and compared with other facial recognition systems. These features are analyzed and compared with an existing database to identify a person or verify their identity.

As you can see from this photo (Figure 14):

Improvements are regularly made to this rapidly evolving technology. These include, for example, the development of 3D sensors, recognition of faces in motion, processing of faces in profile, ...

3.3.1. Improving Biometric Accuracy

The effectiveness of facial recognition depends on several key factors:

- Image quality: the system attempts to distinguish cooperating subjects from non-cooperating subjects. Cooperating subjects are those who voluntarily allow their facial image to be captured. Non-cooperating subjects are those whose faces have been captured by surveillance cameras or witnesses using their smartphones or other means without their cooperation.

Figure 14. Biometric face analysis.

- Identification algorithms: the second key factor in improving performance is the power of the algorithms used to determine similarities between facial features. The algorithms analyze the relative position, size and/or shape of the eyes, nose, cheekbones and jaw. These features are then used to search for other images with similar characteristics.

- Reliable databases: finally, the accuracy of face recognition depends on the size and quality of the databases used: to recognize one face, it has to be compared with another. The difficulty lies in establishing points of correspondence between the new image and the source image, *i.e.* photographs of known individuals. Consequently, the more images there are in the database, the more matches can be found.

3.3.2. Adaptability and Effectiveness of Facial Recognition

Facial recognition can be used for a variety of purposes. For example, facial recognition software on a mobile device (such as a smartphone, tablet or PC) equipped with a camera allows a user to log into an account, replacing the use of a password for authentication.

In the fight against crime, this technology can be used to identify suspects. For border control, it simplifies procedures and increases security at the same time. Another use case for facial recognition access control is for sensitive sites. In the commercial sector, analysts use this technology to gather strategic demographic data.

3.3.3. Facial Recognition Algorithms

In this work, we have analyzed the face recognition algorithm known as Histogram of Oriented Gradient (*HoG*). The aim of this analysis is to present, in detail, the policy that guides the internal logic of this algorithm.

The Histogram of Oriented Gradient (*HoG*) is a function used in computer vision to detect objects. This technique calculates local histograms of the orientation of the gradient in a denser grid, *i.e.* uniformly distributed over the image zones. It has in common with the scale-invariant feature transform (*SIFT*) [26]—an algorithm used in the field of computer vision to detect and identify similar features between various digital images (landscape features, objects, people, ...).

The principle of the algorithm is to extract points from an image by describing the contours, and to obtain for each of these points the context of the shape by determining the relative distribution of the closest points using a histogram of the polars and contour orientation data, but it differs in particular by using a dense grid.

The method is particularly effective for detecting people. *HoGs* were proposed by Navneet Dalal and Bill Triggs, researchers at INRIA Grenoble, at the CVPR conference in June 2005 [27].

The basic idea behind the *HoG* descriptor is that the local appearance and shape of an object in an image can be described by the distribution of the intensity gradient or the direction of the contours. This is done by dividing the image

into adjacent areas of small windows called cells, and calculating for each cell the histogram of the gradient directions or contour orientations for the pixels in the cell. The combined histograms then form the *HoG* descriptor.

For best results, local histograms are contrast-normalized, by calculating an intensity measure over areas (called blocks) larger than the cells, and using this value to normalize all the cells in the block. This standardization provides greater resistance to changes in light and shadow.

3.3.4. Construction of a *HoG* Vector

The histogram of oriented gradients characterizes an object in an image on the basis of its contours using the distribution of local gradient orientations. As shown in Figure 15 below, *HoG* features can lead to better results even in a complex classification task.

1) Gabor filter

Before we get to that, it's important to understand the concept of the Gabor filter. Gabor filters: These are orientation-sensitive filters used to analyze frames. They generally travel in packets, one for each direction. A set of Gabor filters with a given direction provides a strong response for positions in the target image that have structures in that given direction. For example, if the target image is a periodic grid in a diagonal direction, a set of Gabor filters will only give you a strong response if its direction matches that of the grid (Figure 16).

It is widely used in character recognition and fingerprint enhancement. It can also be used in biomedical imaging to characterize the principal orientation in fibrillar structures (Figure 17).

These are not the only areas mentioned above. In this work, it is applied to images to help the algorithm predict human behavior in order to deduce features likely to be associated with dangerous behavior.

By applying the Gabor filter to an image, we obtain approximately the following[2] (Figure 18):

The sense organ responsible for vision is the eye. It enables us to interact with our environment by capturing and encoding its constituent elements. Perceiving an object involves elaborate processing that breaks down the image into several
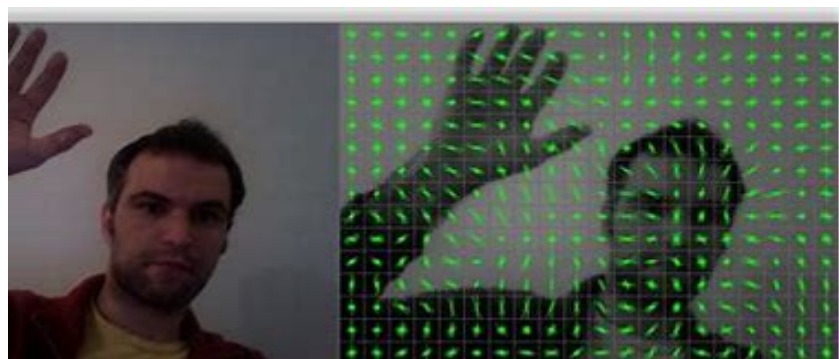

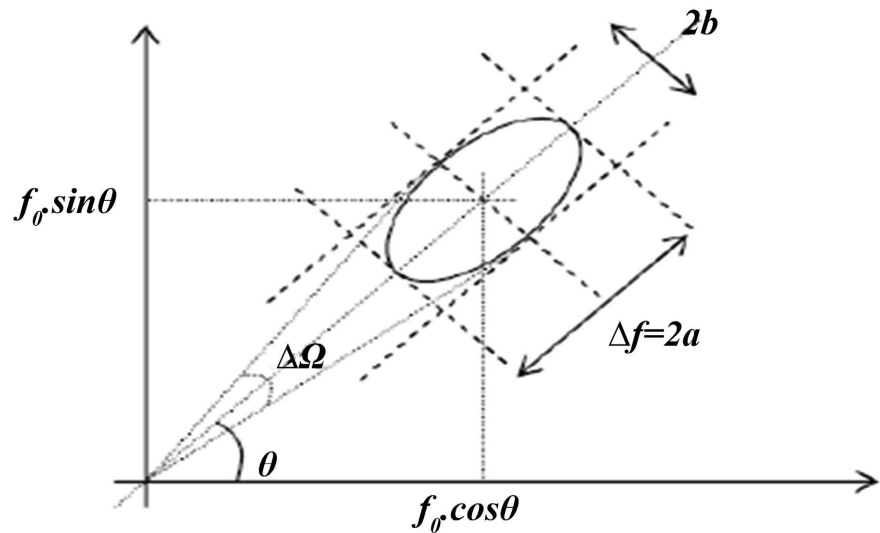**Figure 15.** Construction of the *HoG*.

[2]Documentation *MatLab a2016b*.

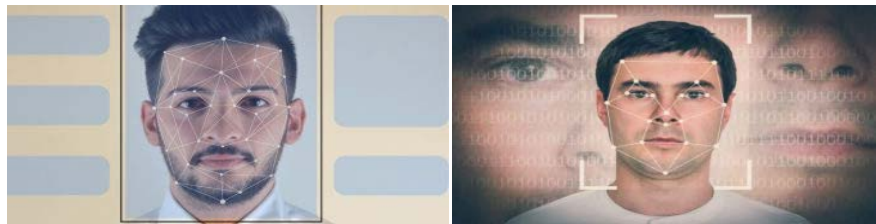**Figure 16.** Orientation of features.
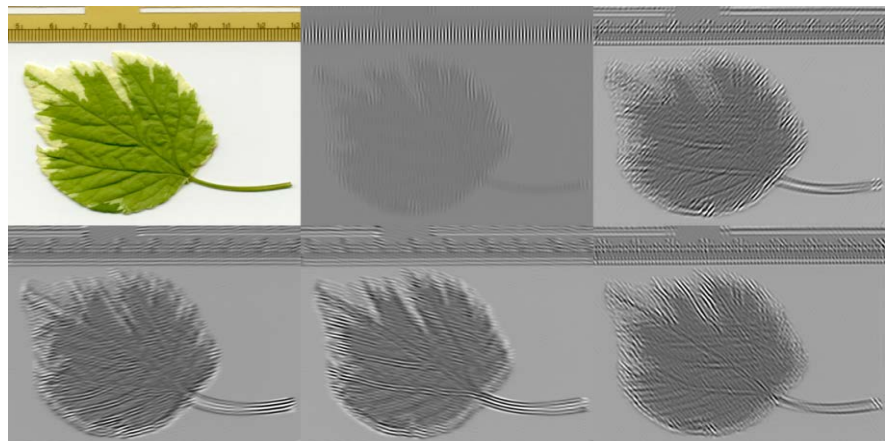

Figure 17. Gabor filter.

Figure 18. Orientation of the screen.

frequencies and orientations. Human vision is therefore undoubtedly a high level of perception. Computer vision, on the other hand, does not aim to faithfully reproduce the functioning of the human visual system, but to model its main characteristics to enable the machine to perform recognition tasks automatically. One of these characteristics is the detection and extraction of object contours.

Concern for the detection and extraction of object contours has led to the development of a number of techniques. Most of these use first-order local opera-

tors such as gradient descent (Prewitt, Sobel) or second-order operators (Laplacian) followed by a search for local maxima.

However, these techniques give insufficient results on a real image, where intensity variations are rarely sharp and require a thresholding operation to eliminate noise. New approaches have proved more effective. They proceed by optimizing the criteria, taking into account a predefined model of the contours to be detected. One example is the Hueckel method, which works on the principle of adjusting the parameters of an ideal contour model to best match the image data.

Among those who have tackled the issue, one name has stood out. This was the English physicist and inventor of holography, Dennis Gabor (1900-1979), who developed the method commonly known as Gabor filters. These filters, which operate in a similar way to human visual processing, have the advantage of being parameterizable in terms of frequency and orientation. Their study will form the final section of this doctoral dissertation.

It is also possible to find, in a scientific environment, work on the application of Gabor filters to the classification of handwriting. In the field of image processing, three types of classification techniques can be distinguished: deterministic (symbolic), probabilistic (neural networks) and ensemblistic (mathematical morphology). It is not wrong to define the Gabor filter as an association of a Gaussian curve [A Gaussian function is a function in exponential of the opposite of the square of the abscissa (a function in $\exp(-x^2)$). It has the characteristic shape of a bell curve] and a directed sinusoid (Figure 19).
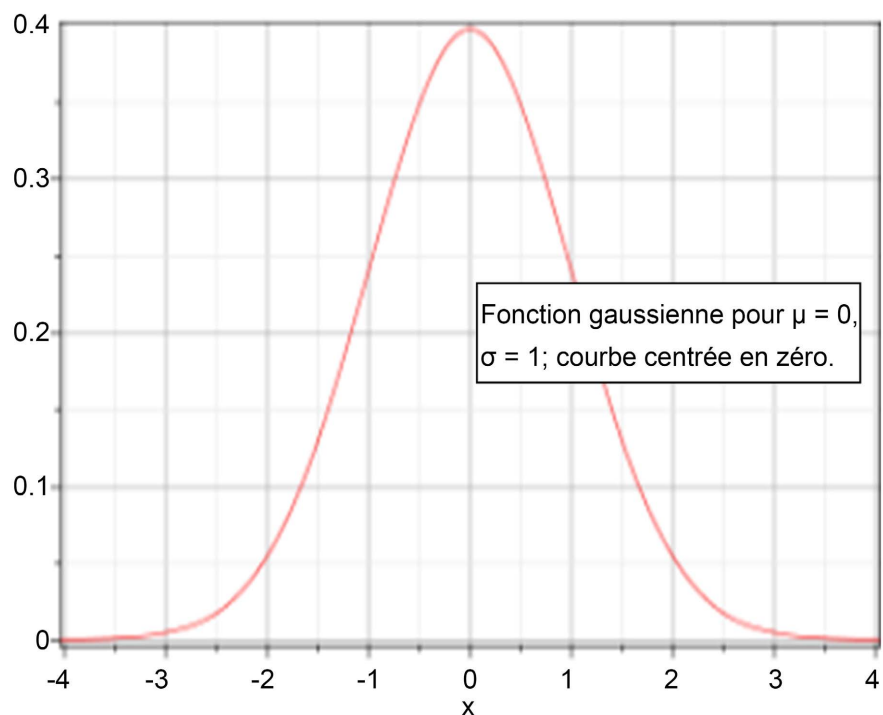

Figure 19. Gaussian function.

In image processing, we work in the 2-dimensional spatial domain. This allows us to write the Gabor function as follows:

$$G(x, y, \vartheta, f) = \epsilon^{-1/2} \left[\frac{x_\vartheta^2}{\sigma_x^2} + \frac{y_\vartheta^2}{\sigma_y^2} \right] \cos(2\pi f x_\vartheta)$$

with $\quad x_\vartheta = x\cos\vartheta + y\sin\vartheta \quad$ and $\quad y_\vartheta = y\cos\vartheta - x\sin\vartheta$

where ϑ is the orientation of the sinusoid, f its frequency and σ_x (respectively σ_y) the standard deviation of the Gaussian along the x-axis (respectively y-axis).

Applying this function to a convolution mask, we define a convolution filter which we will call the Gabor filter and which can be represented as the following image (**Figure 20**).

Consider the case of the mask where the Gabor filter of radius 21 pixels for θ = 0, $f = \sqrt{2}/10$ and $\sigma_x = \sigma_y = 7$. The application of a Gabor filter g of mask M of radius r to an image I of width m and height n, therefore boils down to the following formula:

$$g(I) = J = M * I$$

where J is a matrix of dimension m, n and for $i, j \in \mathbb{N}$, $r \le i < m - r$ and $r \le j < n - r$:

$$J_{i,j} = \sum_{k=-r}^{r} \sum_{l=-r}^{r} M_{k,l} * I_{(i-k),(j-l)}$$
$$= \sum_{k=-r}^{r} \sum_{l=-r}^{r} G(k, l, \vartheta, f) * I_{(i-k),(j-l)}$$

As can be seen, Gabor filters can be used to isolate the contours of an image with an orientation perpendicular to θ and satisfying a certain thickness, which depends on f. This justifies the fact that to detect all the contours of an image, we generally apply a set of Gabor filters to it, which we call a bench.

Input Image of a circle

A bank of 16 Gabor Filters

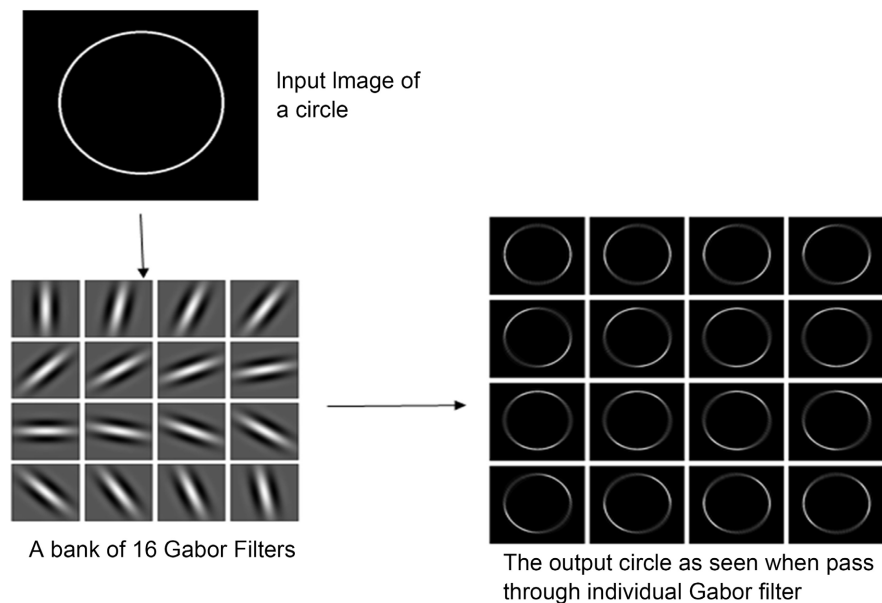The output circle as seen when pass through individual Gabor filter

Figure 20. Gabor filter.

For identification in the context of security, it is therefore possible to implement Gabor filter banks so that they can be used on different images. When the program is run, it is possible to define all the parameters needed to calculate the Gabor functions and apply the bank. In this way, using the characteristics of the basic emotions, it will be possible for the algorithm to competently predict behavior in the direction of a possible threat.

Thus, the combination of a set of Gabor filters provides greater coverage of frequency space and enables a greater number of orientations to be detected, so all the image contours can be extracted. Since the result of a filter bank is the average of the results of the filters in that bank, we can take advantage of the distributivity of the convolution product to lighten the processing.

After this lengthy analysis of the Gabor filter, we can now construct the *HoG* vector. Let's see how! The histogram of oriented gradients, also known as the *HoG*, is a feature descriptor like the Canny Edge detector, *SIFT*. The technique counts the occurrences of gradient orientation in the localized part of an image. This method is quite similar to edge orientation histograms and scale invariant function transform (*SIFT*). The *HoG* descriptor focuses on the structure or shape of an object. It is better than any edge descriptor because it uses magnitude as well as gradient angle to calculate features. For image regions, it generates histograms using magnitude and gradient orientations.

2) Calculating *HoG* features

To determine the fundamental characteristics of the *HoG* vector, we recommend that you follow the steps described below step by step:

a) The first thing is to take the input image on which you wish to calculate the *HoG* characteristics. Resize the image to an image of 128 × 64 pixels (128 pixels high and 64 pixels wide) (Figure 21).

b) The gradient is obtained by combining the magnitude and angle of the image. Considering a block of 3 × 3 pixels, we first calculate G_x and G_y for each pixel. The first G_x and G_y are calculated using the formulae below for each pixel value.

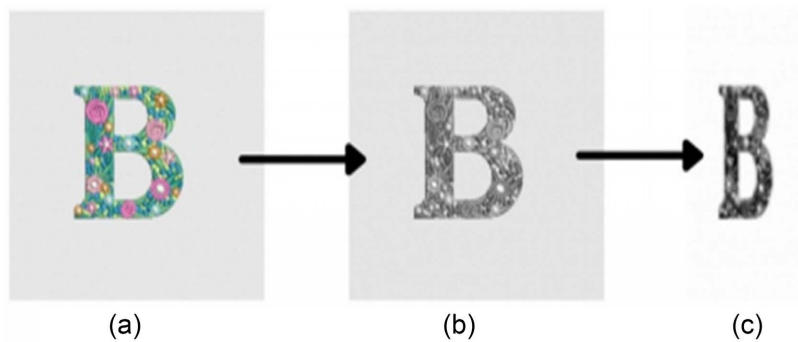$$G_x(r,c) = I(r,c+1) - I(r,c-1)$$

$$G_y(r,c) = I(r-1,c) - I(r+1,c)$$

(a) (b) (c)

Figure 21. Image transformation and resizing.

where r and c denote the rows and columns respectively.

Once G_x and have been calculated, the magnitude and angle of each pixel are calculated using the formulae below (**Figure 22**).

$$\text{Magnitude}(\mu) = \sqrt{G_x^2 + G_y^2} \quad \text{and} \quad \text{Angle}(\vartheta) = \left| \tan^{-1}\left(G_y/G_x\right) \right|$$

c) Once the gradient of each pixel has been obtained, the gradient matrices (magnitude matrix and angle matrix) are divided into 8×8 cells to form a block. For each block, a 9-point histogram is calculated. A 9-point histogram develops a histogram with 9 cells[3] and each bin has an angle range of 20 degrees.

Figure 23 shows a 9-bin histogram in which the values are assigned after the calculations. Each of these 9-point histograms can be plotted as histograms with bins producing the intensity of the gradient in that bin. As a block contains 64 different values, for the 64 magnitude and gradient values, the following calculation is performed. As we are using histograms with 9 points, so:

Number of boxes = 9 (magnitude from 0˚ to 180˚)

Step size (dimension) ($\Delta\theta$) = 180˚/Number of bins = 20˚.
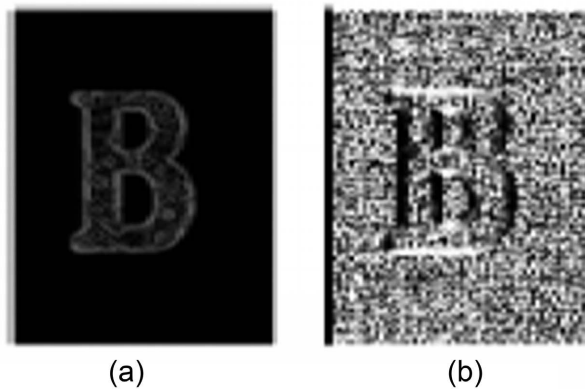

(a) (b)

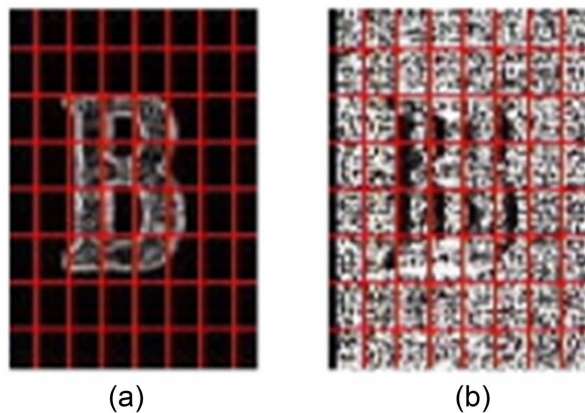**Figure 22.** Magnitude and angle display.

(a) (b)

Figure 23. Magnitude and angle blocks in the image.

[3]To generate a histogram, the range of data values for each bar must be determined. The ranges of the bars are called bins. Most of the time, the bins are of equal size. In this case, the height of the bars indicates the frequency of the data values in each bin.

Each jth bin, will have boundaries from: $[\Delta\theta j, \Delta\theta(j+1)]$ The value of the center of each bin will be: $C_j = \Delta\theta(j + 0.5)$

Value									
Cases	0	20	40	60	80	100	120	140	160

For each cell in a block, we must first calculate the jth cell and then the value that will be supplied to the jth and $(j+1)^{\text{th}}$ cell respectively. This value is given by the following formulas:

$$j = \left\lfloor \frac{\vartheta}{\Delta\vartheta} - 1/2 \right\rfloor$$

$$V_j = \mu \cdot \left[\frac{\vartheta}{\Delta\vartheta} - 1/2 \right]$$

$$V_{j+1} = \mu \cdot \left[\frac{\vartheta - C_j}{\Delta\vartheta} \right]$$

d) An array is considered as a cell for a block and the values of V_j and V_{j+1} are added in the array to the index of the jth and $(j+1)^{\text{th}}$ cell calculated for each pixel.

e) The resulting matrix after the above calculations will have the shape of $16 \times 8 \times 9$.

f) Once the histogram calculation has been completed for all the blocks, 4 blocks of the 9-point histogram matrix are grouped together to form a new block (2×2). This grouping is performed in a superimposed manner with a step size of 8 pixels. For the 4 cells in a block, we concatenate all the 9-point histograms for each constituent cell to form a feature vector of 36 (Figure 24).

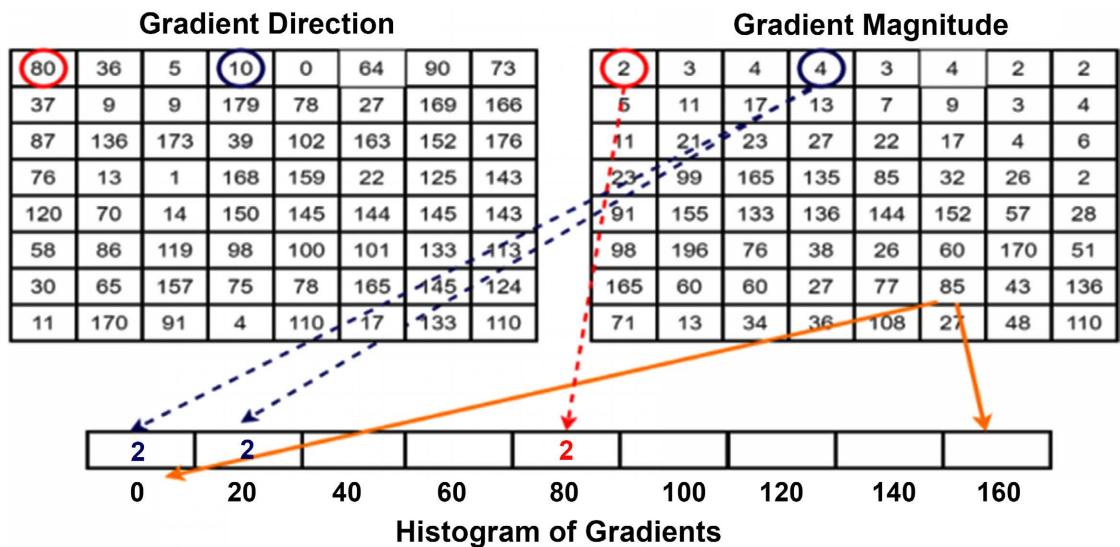$$f_{bi} = [b_1, b_2, b_3, \cdots, b_{36}]$$

Figure 24. Method for calculating 9-box histograms.

g) The f_b values for each block are normalized by the L2 standard:

$$f_{bi} \leftarrow \frac{f_{bi}}{\sqrt{\|f_{bi}\|^2 + \varepsilon}}$$

where ε is a small value added to the square of f_b to avoid the division by zero error.

h) To normalize, the value of k is first calculated by the following formulae:

$$k = \sqrt{b_1^2 + b_2^2 + b_3^2 + \cdots + b_{36}^2}$$

$$f_{bi} = \left[\frac{b_1}{k}, \frac{b_2}{k}, \frac{b_3}{k}, \cdots, \frac{b_{36}}{k} \right]$$

i) This normalization is carried out to reduce the effect of changes in contrast between images of the same object. From each block, a feature vector of 36 points is collected. In the horizontal direction there are 7 blocks and in the vertical direction there are 15 blocks. The total length of the *HoG* features will therefore be: $7 \times 15 \times 36 = 3780$. The *HoG* features of the selected image [27] [28] are obtained (Figure 25).

3) Convolution Matrix: Image Processing and Analysis

The question that many people ask is what is a convolution matrix? What is its purpose? You can get a rough idea without using the mathematical tools that very few people know.

Mathematically, convolution is a mathematical operation that takes two signals (u_1 and u_2) as inputs and returns a new signal s as an output, such that $s = u_1 * u_2$. Convolution is a treatment of one matrix by another called the convolution matrix or kernel.

The convolution matrix uses a first matrix which is the image, *i.e.* a collection of pixels in 2D rectangular coordinates (there are 3D matrices...), and a variable kernel depending on the desired effect. For obvious reasons, we will restrict ourselves to 3×3 matrices, which are the most commonly used. They are sufficient for all the effects required. If all the cells on the edges of the kernel are 0, the system considers that it is a 3×3 matrix. The filter successively studies each of the pixels in the image (Figure 26).

For each pixel, which we will call the initial pixel, it multiplies the value of this pixel and of each of the 8 pixels surrounding it by the corresponding value in the kernel. It adds up all the results and the initial pixel then takes the value of the final result.

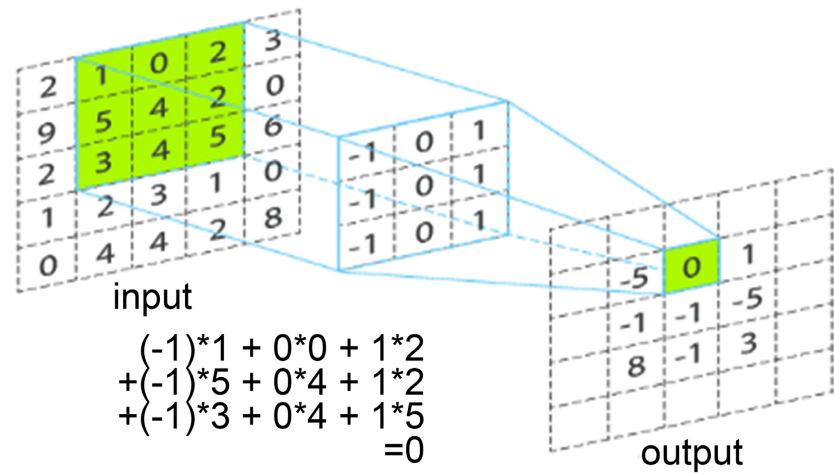**Figure 25.** Visualization of *HoG* characteristics.
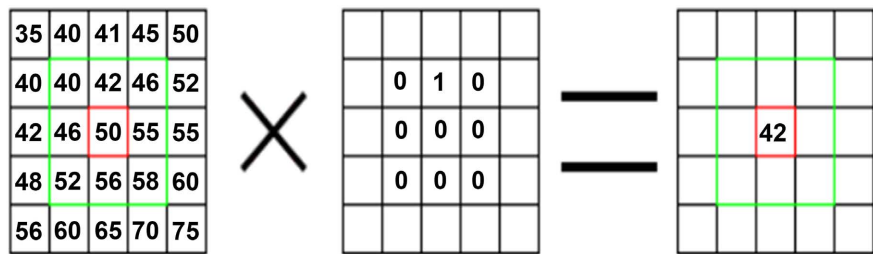
Figure 26. Convolution matrix.

Figure 27. Calculation of the values of the convolution matrix.

Here's a practical example (Figure 27).

On the left is the image matrix: each pixel is indicated by its value. The initial pixel is framed in red. The kernel's area of action is framed in green. In the center is the kernel and, on the right, the result of the convolution.

Here's what actually happened: the filter successively read, from left to right and top to bottom, the pixels in the kernel's action zone, multiplied each of them by the corresponding kernel value and added up the results. The initial pixel took the value 42: $(40 \times 0) + (42 \times 1) + (46 \times 0) + (46 \times 0) + (50 \times 0) + (55 \times 0) + (52 \times 0) + (56 \times 0) + (58 \times 0) = 42$ (the filter deposits its results on a copy of the image and not directly in the image). The graphic result is a downward shift of one pixel from the initial pixel.

The final objective of a convolution matrix is to extract characteristics specific to each image by compressing them so as to reduce their initial size. In short, the input image is passed through a succession of filters, creating new images called convolution maps.

4. Conclusions

Safety concerns continue to haunt the human mind, especially in these times of great mobility. Several models were developed in the second half of the twentieth century to reinforce security, especially at the borders of nations.

It has to be said, however, that none of these models was initially designed to

predict (changes in) individual behavior, particularly in the high-risk area of security.

This chapter has drawn on psychological studies to establish an intrinsic link between basic emotion and human behavior. According to the analyses applied to the various images, emotion fundamentally determines an individual's behavior. It shows up in the face in such a way that it cannot easily be hidden.

If emotion permeates and permeates human behavior, and if it can easily be detected through the individual's face, then it is logically possible to analyze video images to extract the essential characteristics for predicting human behavior.

To do this, a number of instruments and tools need to be brought to bear. From a software point of view, MatLab proved to be very effective in analyzing and modelling the problem. We analyzed both the Gabor filter, whose effects no longer need to be demonstrated, and the *HoG* feature vector and Convolution Matrix.

These techniques, each in their own way, have helped us to highlight the fundamental characteristics of the human face, which bear the seeds of the behavior that the individual is supposed to externalize through concrete acts and gestures. This is what enables the software to predict human behavior and, if necessary, to take safety measures to put the potential perpetrator out of harm's way.

From a hardware point of view, we were limited by the absence of tools to capture video images in order to attempt the experiment. Basic infrastructure is an essential aspect of a successful initiative like this.

In conclusion, we were content to analyze and apply our study of photo images in order to extrapolate it to the prediction of human behavior. This gave us a more or less accurate idea of the safety challenges.

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] Morin, E. (1990) Introduction à la pensée complexe. Du Seuil, Paris.

[2] Le Moigne, J.-J. (1995) Les épistémologies constructivistes. PUF, Paris.

[3] APA (2015) APA Dictionary of Psychology. 2nd Edition, American Psychological Association, Washington DC.

[4] State of Mind: Il Giornale delle Scienze Psicologiche.
https://www.stateofmind.it/psicologia/

[5] Hockenbury, D. and Hockenbury, S.E. (2007) Discovering Psychology. Worth Publishers, New York.

[6] Ekman, P. (1977) Biological and Cultural Contributions to Body and Facial Movement. In: Blacking, J., Ed., *Anthropology of the Body*, Academic Press, London, 34-84.

[7] Plutchik, R. (1984) Emotions and Imagery. *Journal of Mental Imagery*, **4**, 105-111.

[8] Darwin, C. (1872) The Expression of the Emotions in Man and Animals. John

Murray, London. https://doi.org/10.1037/10001-000

[9] Cosmides, L. and Tobby, J. (1987) From Evolution to Behavior: Evolutionary Psychology as the Missing Link. In: Dupre, J., Ed., *The Latest on the Best: Essays on Evolution and Optimality*, The MIT Press, Cambridge, 276-306.

[10] Tobby, J. and Cosmides, L. (2008) The Evolutionary Psychology of the Emotions and Their Relationship to Internal Regulatory Variables. In: Lewis, M., Haviland-Jones, J.M. and L.F. Barrett (Eds.), *Handbook of Emotions*, The Guilford Press, New York, 114-137.

[11] Farrell, R.L., Murtagh, K.E., Tien, M., Mozuch, M.D. and Kirk, T.K. (1989) Physical and Enzymatic Properties of Lignin Peroxidase Isoenzymes from *Phanerochaete chrysosporium*. *Enzyme and Microbial Technology*, **11**, 322-328.
https://doi.org/10.1016/0141-0229(89)90014-8

[12] Dewey, J. (1894) The Theory of Emotion: I: Emotional Attitudes. *Psychological Review*, **1**, 553-569. https://doi.org/10.1037/h0069054

[13] Barrett, L.F. (2017) The Theory of Constructed Emotion: An Active Inference Account of Interoception and Categorization. *Social Cognitive and Affective Neuroscience*, **12**, 1-23. https://doi.org/10.1093/scan/nsw154

[14] Cannon, W.B. (1927) The James-Lange Theory of Emotions: A Critical Examination and an Alternative Theory. *The American Journal of Psychology*, **39**, 106-124.
https://doi.org/10.2307/1415404

[15] James, W. (1884) What Is Emotion? In: Dennis, W., Ed., *Readings in the History of Psychology*, Appleton-Century-Crofts, New York, 290-303.

[16] Laird, J.D. (2007) Feelings: The Perception of Self. Oxford University Press, Oxford.

[17] Nummenmaa, L., Glrean, E., Hari, R. and Hietanen, J.K. (2014) Bodily Maps of Emotions. *Proceedings of the National Academy of Sciences of the United States of America*, **111**, 646-651. https://doi.org/10.1073/pnas.1321664111

[18] Simon, H.A. (1947) Administrative Behavior: A Study of Decision-Making Processes in Administrative Organization. Macmillan, New York.

[19] Parthenay, C. and Thomas-Fogiel, I. (2005) Science économique et philosophie des sciences: la question de l'argument transcendantal. *Revue de Métaphysique et de Morale*, **3**, 428-456. https://doi.org/10.3917/rmm.053.0428

[20] Bastin, C., Harrison, B.J., Davey, C.G. and Moll, J. (2016) Feelings of Shame, Embarrassment and Guilt and Their Neural Correlates: A Systematic Review. *Neuroscience & Biobehavioral Reviews*, **71**, 455-471.
https://doi.org/10.1016/j.neubiorev.2016.09.019

[21] Michi, P., Meindl, T., Meister, F. and Born, C. (2014) Neurobiological Underpinnings of Shame and Guilt: A Pilot fMRI Study. *Social Cognitive and Affective Neuroscience*, **2**, 150-157. https://doi.org/10.1093/scan/nss114

[22] Blanchard, D.C, Hynd, A.L., Minke, K.A. and Minemoto, T. (2002) Human Defensive Behaviors to Threat Scenarios Show Parallels to Fear- and Anxiety-Related Defense Patterns of Non-Human Mammals. *Neuroscience & Biobehavioral Reviews*, **25**, 761-670. https://doi.org/10.1016/S0149-7634(01)00056-2

[23] Steimer, T. (2002) The Biology of Fear- and Anxiety-Related Behaviors. *Dialogues in Clinical Neuroscience*, **4**, 231-249.
http://www.ncbi.nlm.nih.gov/pubmed/22033741
https://doi.org/10.31887/DCNS.2002.4.3/tsteimer

[24] https://www.eiagroup.fr/domaines-expertise/expressions-faciales-et-micro-expressions/

[25] Lie to Me (TV Series 2009-2011). https://www.imdb.com/title/tt1235099/

[26] Lowe, D.G. (1999) Object Recognition from Local Scale-Invariant Features. *Proceedings of the* 7*th IEEE International Conference on Computer Vision*, Kerkyra, 20-27 September 1999, 1150-1157. https://doi.org/10.1109/ICCV.1999.790410

[27] Dalal, N. and Triggs, B. (2005) Histograms of Oriented Gradients for Human Detection. http://lear.inrialpes.fr/people/triggs/pubs/Dalal-cvpr05.pdf

[28] https://www.youtube.com/watch?v=QmYJCxJWdEs