Horn Schunck Algorithm for Facial Expression Change Detection Classification

N. Zainudin, S. Ramli, T. Tan Zizi, N. A. Hasbullah, M. Mohd Hanafi, N. A. Wahab
Universiti Pertahanan Nasional Malaysia, Malaysia

N. Ibrahim
Universiti Kuala Lumpur Malaysia France Institute, Malaysia

Abstract

This study is aimed to investigate the use of optical flow techniques in detecting changes in facial expressions. Facial expression detection is very important as we can identify a person’s emotional and mental state. When communicating with people, we can identify someone’s expression accurately compared to a computer. Although many approaches are applied on this topic, there are still several drawbacks and limitations. In order to yield better results, we have applied optical flow technique to detect facial expression. In addition, we have carried out our experiment using Horn-Schunck method to optimize the results. Based on the experiment conducted, the average value represented by every facial expression can be identified, and the values are significant for future research that focuses on facial expression classification.

1. Introduction

The ability to detect changes in facial expressions is very important as we can identify expression and mental conditions of a person. A study by Mehrabian psychologist showed that only 7% of information are expressed verbally, 38% from an extended language, such as speech, tone and body language while 55% of information can be identified through facial expressions. Therefore, a lot of important information can be derived through facial expression detection, especially when a person is emotionally and mentally disturbed.

Human facial expression detection is divided into six simple categories, namely sad, happy, angry, surprise and hate. Various studies were conducted using these six categories, especially in image processing and pattern recognition in an effort to improve the present technology. Therefore this technology is widely applied in human-computer interface, computing, intelligent system, security cameras, pattern recognition, psychological analysis, social life, machine vision and social entertainment. With the advent of this technology, users accord with more comfortable life. In designing a system as such, some important aspects should be taken into consideration. Among other aspects that should be looked into are cost of equipment, maintenance cost, database, security level and also its user-friendliness. With those aspects taken into account, the system could have the necessary specifications which will meet consumers’ expectations.

Early research by Charles Darwin in his book entitled "The Expression of the Emotion in Man and Animals" stated that emotion or expression defines humans as a species not a culture. In 1969, "Ekman and Friesen" classified emotion in a universal cultural difference which contains six expressions, namely happiness, grief, angry, hatred, shock and fear. Facial expression factor has a high stake value as it can easily be identified. When communicating with people, human can identify someone’s expression accurately compared to a computer. Hence, intelligent system component exists in detecting human expression. The computer will be able to achieve accuracy in classifying expressions as good as humans if the algorithm that is used is as if human thought. Although many approaches are tackled on this topic, there are still several drawbacks and limitations. So, to yield better results, researchers use an optical flow technique to detect facial expression during one’s disturbed emotions. If this technique is effective in identifying facial expression, it can facilitate and increase the performance of optical flow as one of the techniques that can be used to detect facial expressions. As a result, it can be utilised as the algorithm in future applications involving facial expressions.

Previous research said that one of the popular methods for recognition of facial expressions such as happiness, sadness and surprise is based on deformation of facial features. Motion vectors which show these deformations can be specified by the optical flow [1]. With this method, for detecting expressions, the resulted set of motion vectors are compared with the standard deformation template that caused by facial expressions. In this paper, a new method is introduced to compute the quantity of likeness in order to make decisions based on the importance of obtaining vectors from an optical flow approach. For finding the vectors, one of the efficient
optical flow methods developed by Horn and Schunck is used.

2. Literature Review

The earliest work on face recognition can be traced back at least to the 1950s in psychology [Bruner and Tagiuri 1954] and to the 1960s in the engineering literature [Bledsoe 1964]. Some of the earliest studies include work on facial expression of emotions by Darwin [1972], Ekman [1998] and on facial profile-based biometrics by Galton [1888]. But facial expression change detection (face expression recognition) has now expanded in many applications especially in security applications. Using a face image as basic image in facial expression change detection has many benefits. One of the benefits is such detection does not need facial change to be drawn or photographed.

There are many techniques to find dimensions from images which will be processed, and optical flow is the technique being studied in this research. This technique is studied further to measure effectiveness level and capacity of security system and other intelligent applications. A variety of algorithms have been created or developed to fulfill introduction process of facial recognition. The image processing technique of this research is as discussed in the following section. Besides that, based on previous research [5] facial expressions are generated by contractions of official muscles, which results in temporally deformed facial features such as eyelids, eyebrows, nose, lips and skin texture, often revealed by wrinkles and bulges.

In addition, previous research in year 2003 [6] also shown that there are three types of feature extraction methods can be distinguished: (1) generic methods based on edges, lines, and curves; (2) feature-template-based methods that are used to detect facial features such as eyes; (3) structural matching methods that take into consideration geometrical constraints on the features.

To resolve problems regarding facial recognition, it should be prioritised because it involves important processes in determining accuracy and how effective is the technique in identifying certain facial expression. Therefore, selecting the appropriate and suitable technique will be discussed in deciding the best method for facial change recognition. Apart from this, back to the early 1970, a few researchers conducted studies on human facial expressions. Ekman and Friesen [5] postulated six primary emotions that possess each a distinctive content together with a unique facial expression. These prototypic emotional displays are also referred as basic emotions. They seem to be universal across human ethnicities and cultures and comprise happiness, sadness, fear, disgust, surprise and anger. Figure 1 shows all six universal facial expressions.

![Figure 1. Six images with six universal facial expressions adopted from A. R. Naghsh-nilchi and M. Roshanzamir, 2006 [1]](image)

Therefore, without a doubt, there are various techniques that are applicable to detect changes in facial expressions nowadays. And this research uses the Horn-Schunck technique as the major algorithm in detecting the changes in facial expressions. The Horn-Schunck algorithm has many advantages, namely its ability to launch current to images as explained in the book Determining optical flow 'Artificial Intelligence' (1981) in pages 185-203[2]. This ability will minimise distortions of the optic current itself and further enhance the image results being produced. Horn-Schunck algorithm will be explained more below.

There has been much work on face recognition and facial expression recognition in computer vision. Comprehensive literature surveys are provided in [5] and [6] on face recognition and facial expression recognition, respectively. Most facial expression recognition techniques use the feature deformation model based on optical flow [7] to realize person-independent recognition. Most face recognition techniques usually focus on specific parts of the face to deal with expression-independent face recognition. However, to the knowledge of the authors, none of the previous work has investigated simultaneous face and facial expression recognition.

As we all know that facial expression mapping has been widely studied in the field of computer vision and graphics. According to the previous research that have been done.
The paper is organized as follows. In Section 2, we give a brief overview of the proposed algorithm, and the methods used for facial expression synthesis. Section 3 presents a methodology that used in the experiments we have conducted and evaluates the performance of the approach presented. Results and discussions of the experiment will be discussed in Section 4. Finally, we present conclusions and future research directions in Section 5.

2.1. Surveillance Camera

A surveillance camera is one amongst the most extensively used technologies for corporeal security [8]. It also has become an important tool in security and has become a necessity to keep proper check [9]. As the number of surveillance cameras being installed in various fields increased, computational vision based object detection has become vital worldwide. In computer vision, this is the task of finding a given object in an image or video sequence. Several image processing techniques are developed for the detection of different objects from images and video sequences [9]. A video collection device is known as surveillance cameras.

Surveillance cameras accomplishments have become intensified most contemporary. The technology has evolved in such a way that, it strives to accomplish automatically by operating via facial expression and emotion recognition using the information in the form of images obtained from a surveillance system. This is known as a group of communication impedimenta devices that cluster image data from a surveillance camera that is installed at specific locations. This transmits the captured images through a wired or wireless communication channel, so that the concerned person can retrieve it [8].

2.2. Horn-Schunck Optical Flow Algorithm

According to [10], optical flow reflects the image changes due to motion during a time intervals. Optical flow calculation is a prerequisite for higher level processing which can solve motion problems. Previous research has also tested these algorithms on several standard image sequences [1].

Optical flow method [3] is an important method of motion image analysis and has a widely application in machine vision and image processing. Motion information of facial expression can be analyzed by the motion field of optical flow estimated. Horn-Schunck optical flow algorithm is a dense optical flow field based on the assumptions of grayscale consistency, the optical flow calculations would be inaccurate when the brightness is not invariance or a motion object exists deformation. Facial motion is typical of non-rigid motion. Using the traditional optical flow method will inevitably lead to inaccurate of optical flow field and affect the facial expression recognition rate.

In this paper, on the basis of the extended optical flow constraint equation, a novel approach for estimating facial expression change detection based on Horn and Schunck optical flow is presented. The approach is used to calculate optical flow field of facial expression sequences. The experiment results show that the performance of this approach is better than normal method [11].

On top of that, we propose an optical flow method for moving images because the characteristics of optical flow are quite robust to abrupt movement [5,6]. Previous research has used optical flow because it enables the extraction of velocity and angle. Hence this technique will calculate the displacement and dense of intensity in each frame in the movement video [13]. The Horn-Schunck algorithm (HS) is one of the classical algorithms in an optical flow due to its reasonable performance and simplicity of the algorithm. This algorithm is based on a differential technique computed by using a gradient constraint (brightness constancy) with a global smoothness to obtain an estimated velocity field [7,8]. In conventional predictive methods for motion estimation, the difference between the current frame and the predicted frame, based on a previous frame (motion vector, or MV), is coded and transmitted; then it is used to reconstruct a higher resolution still image or video sequence from a sequence of low resolution images in achieving super-resolution [15].

Optical flow presents an apparent change of moving object’s location or deformation between frames. Optical flow estimation yields a two-dimensional vector field, i.e., motion field, that represents velocities and directions of each point of an image sequence [10,11]. Previous research [18] has compared differential optical flow fields from Horn-Schunck, Lucas Kanade and Brox’s warping techniques. The Horn-Schunck algorithm aims for better smoothing effect by providing denser fields compared to others. Within the large range of object displacements, it provides consistent fields of optical flow. However, the fields are very sensitive to errors derived from the variety of their neighboring points [18]. Therefore, we use Horn and Schunck approach to estimate the optical flow needed in our algorithm.

2.3. Neural Network

Neural networks are trained to predict “before” and “after” pixel values for a sequence of images. These networks are then used to predict expected values for the same images used in training. Substantial differences between the expected and actual values represent an unusual change. Instead, neural networks are used to determine what an expected change is, and highlight the changes that do
not meet these expectations. The goal is not to understand the change, but to highlight areas deserving further analysis. Change detection, thus serves as a “first pass” in image analysis, weeding out significant visual differences that are not likely to be of interest [19,20].

Based on previous research [19], the images are thus represented as a set of point vectors; an artificial neural network (ANN) is then trained to predict the new values from the old values. The ANN is then used to predict the new values on the same images; the points where the prediction is significantly in error are deemed unusual changes. Note that there is no “training data” as such; the training is done against the actual test data. It is the ability of neural networks to generalize, rather than their ability to learn against a manually-tagged corpus, that is key [19].

Apart from that, previous research found that artificial neural networks have been applied to the change detection problem [3, 15], specifically using images and land use category “training data” to identify changes in land cover. These techniques still require considerable manual effort to define what is or is not interesting. In addition, this leaves the possibility that a preconceived notion of what is interesting may be wrong.

Moreover, the detection of change in an overhead imaginary has triggered an attraction to detect abnormalities over a big area, albeit under the influences of illumination. Chris Clifton [19,20] has developed a predictive modeling to identify unusual changes in images. It is done using neural network training method. A research by P Reisman. O. Mano et.al uses the x-axis of the graph corresponds to the position on the scan line by using Hough transform while the y-axis is the probability of that pixel that makes the detection of human and vehicles in the crowd is possible [4].

3. Methodology

This section will explain the different steps required for the facial expression change detection process. The following describes which pictures were used and how they were pre-processed. This will be followed by an explanation about the developed facial expression and interpretation process. The flow chart in Figure 2 shows the process by which the facial feature extraction process is performed. The format used in processing the raw video file is “.avi” towards behavior classification. The video file chosen being extracted into image frames and the Horn-Schunck optical flow is implemented to two different image frames of facial expression in order to find the velocity of moving objects.

![Flow chart of Facial Expression Change Detection](image)

Figure 2. Flow charts of Facial Expression Change Detection

Besides that, based on previous research [19,20], the pattern recognition and image processing are divided into four phases; (i) data collection (ii) pre-processing (iii) features extraction and (iv) classification.

3.1. Data Collection

There are three videos have been taken using a visual camera with various facial expressions of persons for data collection. The facial expressions that used during the experiments are normal expression, happiness, shocked and sad expression. The expressions have been chosen according to the previous research [1], [11], [23] – [26].

3.2. Image pre-processing

The facial image pre-processing includes: video of the facial expressions collection and image frame extraction. Colored images of the facial expressions were then transformed into grayscale images. All recorded images were focused on the face area.

During the preprocessing process, the facial expression video sequence will extract into facial image frames that are suitable for further processing. First, the given input facial expression of video sequence is converted into frames, which are RGB (colored) frames. Colored frames of the data set were then transformed into grayscale images. The resultant five feature regions (both eyebrows, both eyes and mouth), formed the input data for the extraction process.

3.3. Facial feature extraction

After the video sequence being extracted into facial image frames, the Horn-Schunck optical flow is implemented to facial expression frames chosen in
order to find the velocity of moving objects. The idea of motion detection is based on finding amount of difference in two consequent frames of a video sequence.

A set of comparative experiments was prepared utilizing Horn-Schunck facial expression video sequences. Four facial expressions such as normal, happy, shocked and sad were sampled to be recognise in this work. From the facial expression video file, we have collected variable length consecutive frames from each video sequence. The collected frames are then realigned with the size of 480 x 272 pixels. To evaluate the performance of the proposed system, we applied a total of 30 to 90 image sequences per expression for training and testing each expression respectively.

3.4. Facial expression classification

After the feature points were tracked and the velocity of the points calculated (facial expressions feature data). The next step of the classification process identifies the facial expressions of each sequence into the four categories: normal, happy, shocked and sadness category. To classify the different facial expression region neural network is used as classifier that is just one of many possible choices for learning the class separation boundaries. The classification process compared the optical flow plot of selected image frames from the video file which were calculated for each of the four expressions categories: normal, happy, shocked and sad.

In computer science and relative fields, artificial neural networks are computational models inspired by animal central nervous system (in particular the brain) that are capable of machine learning and pattern recognition. They are usually presented as systems of interconnected neurons that can compute values from inputs by feeding information through the network. The feature vector is given as input the neural network. Neural network involves a series of algorithms that attempt to identify underlying relationships in a set of data by using a process that mimics the way the human brain operates.

4. Result and Discussion

The experiments were conducted here using 100 different persons, to find a solution of the problem of facial expression of face recognition. For this study, feature extraction algorithm called Horn-Schunck optical flow algorithm has been used. For these experiments 100 different face images are of varying expressions with a total of 4 different poses which are normal, happy, shocked and sad. This work has been simulated using MATLAB (R2010a) [15] in a machine of configuration. Horn-Schunck optical flow showed its superiority over all the feature extraction methods by means of achieving the highest recognition rate.

Other than that, flow vectors are the collection of vectors which shows motion and deformation expressed in the face due to emotion representation. To obtain flow vectors, a series of facial images is chosen which its motion vectors show facial expression correctly similar to Horn and Schunck optical flow algorithm [2,3]. For example, when we want to save flow vectors of happiness, more fibrous vectors appear in the flow field. Table 2 shows an example of flow vectors which used for detection of normal face, happiness, shocked and sad. In this case, four frames of a different faces' image are programmed as the input image sequences. This procedure is run for all of basic emotions.
Moreover, to make results for flow vectors in the flow field more clearly, the \( t \) value (1) is created which is value of maximum optical flow minus the minimum value optical flow for all image sequences from each video of experiments [27]. Hence, the \( t \) value for every frame in all experiments was calculated and the highest value of the frames was selected as the reference image.

\[
\text{\( t \) value} = \text{maximum optical flow} - \text{minimum optical flow} \quad (1)
\]

Referring to the Table 1, we can identify the average value represented by every facial expression that is derived from (1). Through this average value, we can make a rough analysis on the optic flow change which occurred with the existence of two frames of facial expression images with different rates. Table 1 shows that Expression 3 has an optical flow value that is the highest compared to Expression 1 and Expression 2. This proved that Expression 3 has much movement in the facial expression, compared to Expression 1 and Expression 2. This is also proven in most of the study [1], [11], [23], [25], [26], [28] including this experiment that, when much movement in the face occurred, it will show the highest optical flow value.

<table>
<thead>
<tr>
<th>Info</th>
<th>Experiment 1</th>
<th>Experiment 2</th>
<th>Experiment 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum Of</td>
<td>-1.013</td>
<td>-1.012</td>
<td>-1.130</td>
</tr>
<tr>
<td>Maximum Off</td>
<td>+1.105</td>
<td>+1.140</td>
<td>+1.175</td>
</tr>
<tr>
<td>( t ) Value (Minimum Off - Maximum Off)</td>
<td>+2.135</td>
<td>+2.261</td>
<td>+2.311</td>
</tr>
</tbody>
</table>

Table 1. Results of average value

In conjunction with values in Table 1, Table 2 shows that Expression 3 has an optical flow value that is the highest compared to Expressions 1, 2 and 4. This proves that Expression 3 has much movement in facial expression, compared to Expressions 1, 2 and 4. Meanwhile Experiment 1 has an optical value stream that is the lowest compared to Experiment 2 and Experiment 3. This showed that (Experiment 1 < Experiment 2 < Experiment 3).

5. Conclusion

In this paper, we introduced an efficient algorithm for facial expression detection based on Horn and Schunck optical flow technique to extract the necessary motion vectors, and consequently, determining the faces of varying expressions like normal, happy, shocked and surprise.

Recently, facial expression change detection has been a challenging issue in the pattern recognition field. At present there are still many problems need to be further addressed. In conjunction with that issue, we have proposed an approach based on the extended optical flow constraint and the approach has been applied in facial expression recognition, and the experimental results show that the performance of this approach is better than the normal method.

We have proposed a video-based using Horn and Schunck optical flows for facial expression change detection. We have illustrated the performance of our proposed method applied to sequential image frames for facial expression recognition problems. The experimental results show that optical flow improves the feature extraction task. Furthermore, optical flow processed sequential facial expression images can provide superior recognition rate over the other feature extraction approaches [26].

For classification of these features neural network is used as classifier which provides better results and few false detections are observed in the presence of significant noise. Also training the neural network to separate the given input data into classes is little difficult.

Based on the experiments that has been conducted, the average values of the facial expressions were identified. Through this average value, researchers can make a preliminary analysis of the optic flow changes which occurred with the existence of two frames of facial expression images that are derived from two different rates.

In the future, we will consider classifying the expressions so that we can determine human behavior by applying Horn-Schunck algorithm and will be integrated with thermal sequences images data illumination with different facial action units in order to produce accurate results.
7. Acknowledgements

N. Zainudin would like to thank Research Management Office and Faculty of Defence Science and Technology, National Defence University Malaysia, and Ministry of Higher Education Malaysia for the support during the making of this paper. The research has been supported by the grant from RACE (RACE/F3/SG5/UPNM/3), January 2015 - January 2017 dated 29/03/2011.

8. References


