INTERNATIONAL JOURNAL ON ORANGE TECHNOLOGY

https://journals.researchparks.org/index.php/IJOT e-ISSN: 2615-8140 | p-ISSN: 2615-7071

Volume: 4 Issue: 6 |Jun 2022

Enhancement of Viola-Jones Algorithm using Local Binary Pattern Applied to Face Detection

Dominique Kyle Gonzaga, Jasmine Anetin, Vanessa Mae Castillon, Vivien A. Agustin

Computer Science Department, College of Engineering and Technology, Pamantasan ng Lungsod ng Maynila (University of the City of Manila),General Luna corner Muralla Street, Intramuros, Manila, Philippines, 1002

dkbgonzaga2018@plm.edu.ph
jsanetin2018@plm.edu.ph
vmgcastillon2018@plm.edu.ph
agustin.vivien0406@gmail.com

---***---

Annotation: Face detection is a type of computer image processing technology that can detect faces in digital images. In real-time applications such as CCTV surveillance and video tracking, automatic face detection and recognition is the most difficult and rapidly increasing study topic. One of the most well-known and often used methods for detecting human faces is the Viola-Jones Algorithm. The difficulty associated with the algorithm can be attributed to many variations in the angles of a person's face. In this paper, an enhanced Viola-Jones algorithm with a local binary pattern (LBP) is used to recognize numerous and tilted faces with excellent accuracy.

Keywords: Viola-Jones Algorithm; Local Binary Pattern; Face Detection.

1. Introduction

During the last few decades, one of the highest principal subjects regarding computer vision and pattern recognition is face detection. Several publications provide various approaches for face detection. Viola-Jones algorithm is an object detection framework presented by Paul Viola as well as Michael Jones in 2001 as the first and earliest to deliver a competitive object detection result in real-time. Although it may be taught to recognize a wide range of object classes, its primary motivation was the challenge of face identification. Viola-Jones is a prominent approach for detecting faces. This algorithm is well-known since it was the first to be able to perform face detection in real-time. Despite being a reliable real-time face detector, this algorithm has some drawbacks. Viola-Jones algorithm was initially intended for faces facing the front; hence, it detects faces facing the front better than the side, upward, or downward faces, Bokade, 2021. In this paper, local binary patterns are used alongside the Viola-Jones algorithm to identify faces that are tilted or angled with outstanding precision.

Local binary patterns were originally meant to describe conventional textures, but because a face may be understood as a composition of micro-textures depending on the local environment, they are equally valuable for face description. The local binary pattern descriptor consists of a local and global texture representation obtained by dividing the image into several blocks and then computing each texture histogram. The first gives particular and thorough face information that may be utilized not only to choose faces but also to provide face information for recognition, Lopez, 2010

2. Existing Viola-Jones algorithm

2.1. Overview

The Viola-Jones is an object-recognition framework that enables real-time image feature detection. This algorithm takes a lot of time to train, but it can detect faces in real-time at an impressive speed. Before face detection, the image would first be converted to grayscale. Viola-Jones Algorithm will then look at numerous subregions and

INTERNATIONAL JOURNAL ON ORANGE TECHNOLOGY
https://journals.researchparks.org/index.php/IJOT e-ISSN: 2615-8140 | p-ISSN: 2615-7071
Volume: 4 Issue: 6 |Jun 2022

will try to locate a face by looking for specific characteristics in each region. This algorithm used Haar-like features to detect faces. Haar-like features were considered a digital image feature that is used in object recognition, and with its help, the interpretation of different parts of a face became possible.

One Hungarian mathematician namely, Alfred Haar, provided concepts regarding Haar wavelets in the 19th century. Haar-like features contain a dark side and a light side. Through those sides, the machine can determine what the feature is. Some universal properties are evident in all human faces. For instance, the area of the nose is brighter than the area of the eyes, and the area of the eyes is much darker than the adjacent pixels. Haar-like features have three types first is the edge features, second is the line features, and lastly, the four-sided features (shown in Figure 1). Line and edge features can be applied to the line and edge detection, and four-sided features can be applied to diagonal feature detection.

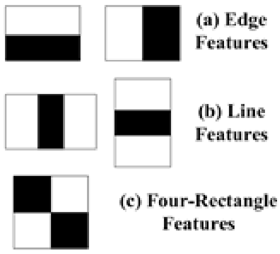

Fig. 1. Haar-like features

The difference between the number of pixel values in the dark area and the number of pixels in the light area gives a HaarLike feature value:

$$F\ (Haar) = \sum Fdark - \sum Fbright \quad (1)$$

The overall feature value and the feature value on the white area $\sum Fbright$, as well as the feature value on the black area $\sum Fdark$, was represented by $F(Haar)$. Haar-like features consist of two or three rectangles. The candidate's face is scanned for haar features related to the current stage.

Rectangle features may be calculated quickly utilizing an intermediate picture representation known as integral image. Integral picture at x, and y includes total of all the pixels positioned at the left, and above of x, and y:

$$ii(x,y) = \sum_{x' \le x,\ y' \le y} i(x',y'), \quad (2)$$

As demonstrated in Figure 2a, an integral picture was a way to quickly determine the value of the feature by transforming each pixel’s value into a new representation of the image.

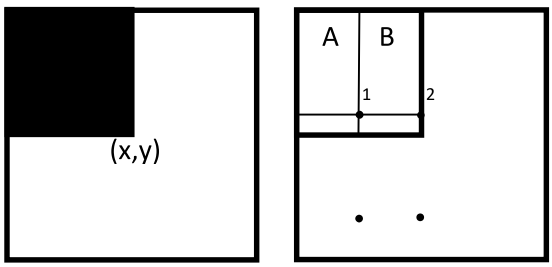

Fig. 2. (a) First illustration - integral image (x,y); (b) Second illustration - figure score count

INTERNATIONAL JOURNAL ON ORANGE TECHNOLOGY

https://journals.researchparks.org/index.php/IJOT e-ISSN: 2615-8140 | p-ISSN: 2615-7071

Volume: 4 Issue: 6 |Jun 2022

You can use the four array references to determine the sum of the pixels in rectangle D. The value of the integral picture in point 1 is none other than the sum of pixels in the rectangle A. Point 2 has the value of A + B, point 3 has the value of A + C, and point 4 has the value of A + B + C + D. Total within D is calculated as 4+1–(2+3), Damanik, et al., 2018. Any rectangular sum in four array references may be calculated using the integral picture (see Figure 2b).

The difference between the two rectangular sums can be calculated with eight references. The features of the above two rectangles include the sum of adjacent rectangles, so they can be calculated with six array references. Eight for three rectangular features and nine for four rectangular features., Viola, and Jones, 2001.

2.2 The problem with Viola-Jones Algorithm

One of the problems in the Viola-Jones Algorithm is that it has a limitation for multi-view face detection and lacks robustness in handling faces under extreme lighting conditions. Aashish, et al., 2017, mentioned that one of the drawbacks of the Viola-Jones Algorithm is that it is not much effective when it comes to detecting faces that are tilted or turned. In addition, there is also a sensitivity to lighting conditions, and overlapping sub-windows can detect the exact face differently. According to the study by Islam, et al., 2017, entitled "Comparison Between Viola-Jones and KLT Algorithms and Error Correction of Viola-Jones Algorithm", the Viola-Jones Algorithm needs to have a suitable and proper front view from the camera. And the faces must not face sideways because even though this algorithm can properly detect faces with a frontal view, it seems to be vulnerable when the face is bent at least 45 degrees or more. This is the main defect or error of this algorithm.

2.3 Pseudocode of Viola-Jones Algorithm

Input: test image

Output: image with detected frontal face drawn with rectangles

for $i \leftarrow 1$ to num of scales in pyramid of images do

Downsample image to create $image_i$

Compute integral image, $image_{ii}$

for $j \leftarrow 1$ to num of shift steps of sub-window do

for $k \leftarrow 1$ to num of stages in Haar cascade classifier do

for $l \leftarrow 1$ to num filters of stage k do

Filter detection sub-window

Accumulate filter outputs

end for

if accumulation fails per-stage threshold then

Reject sub-window as face

Break this k for loop

end if

end for

if sub-window passed all per-stage checks then

INTERNATIONAL JOURNAL ON ORANGE TECHNOLOGY
https://journals.researchparks.org/index.php/IJOT e-ISSN: 2615-8140 | p-ISSN: 2615-7071
Volume: 4 Issue: 6 |Jun 2022

Accept this sub-window as a face

end if

end for

end for

3. Enhanced Viola-Jones Algorithm

3.1 Enhancement of Viola-Jones Algorithm

The goal of this paper is to enhance the Viola-Jones Algorithm to be able to detect multi-viewed faces. This paper proposes to use a local binary pattern, also known as LBP features, for the angled face detection instead of using haar-like features.

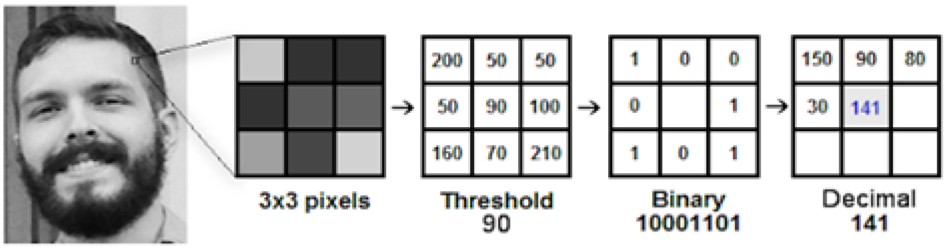

Fig. 3. Local binary pattern labeling

Figure 3 shows a grayscale picture of the sample face. Part of this image is saved as a 3x3 pixel window. Alternatively, the intensity of each pixel can be represented as a 3x3 matrix in the range 0-255. Then, the matrix's centre value was taken as the threshold. This obtained value will now be used to define new values that are derived from neighbours' 8 values. Then, a new binary value will be assigned to each of the neighbours of the centre value. Values more than or equal to the threshold are set to 1, and then the values that are lesser than the threshold are set to 0. Instead, the matrix will now contain the binary value, ignoring the central value. Each binary value from each of the points located in the matrix needs to be concatenated per line to a new binary value. This binary value is then converted into a decimal number and will be assigned to the matrix's centre value, which is the pixels from the original image. When this method is completed, a completely new image is obtained, which reflects the characteristics or properties of the original image better, Prado, 2017.

3.2 Pseudocode of Enhanced Viola-Jones Algorithm

Input: test image

Output: image with detected angled face drawn with rectangles

Convert image/frame to grayscale

Flip image/frame horizontally

Function angled face detector

for $i \leftarrow 1$ to num of scales in pyramid of images do

Downsample image to create $image_i$

Compute integral image, $image_{ii}$

RESEARCH PARKS

INTERNATIONAL JOURNAL ON ORANGE TECHNOLOGY

https://journals.researchparks.org/index.php/IJOT e-ISSN: 2615-8140 | p-ISSN: 2615-7071

Volume: 4 Issue: 6 |Jun 2022

for $j \leftarrow 1$ to num of shift steps of sub-window do

for $k \leftarrow 1$ to num of stages in LBP profile cascade classifier do

for $l \leftarrow 1$ to num filters of stage k do

Filter detection sub-window

Accumulate filter outputs

end for

if accumulation fails per-stage threshold then

Reject sub-window as face

Break this k for loop

end if

end for

if sub-window passed all per-stage checks then

Accept this sub-window as a face

end if

end for

end for

end function

Reflip image/frame horizontally

Repeat function angled face detector

Draw rectangle to detected angled faces

4. Methodology

4.1 Research Design

In this study, the researchers utilized experimental research with a quantitative approach in gathering and analyzing the data. The experimental method was adhered to during the testing and enhancement of the algorithm. While the quantitative approach follows to emphasizes the number of faces in the test image and video and the detected faces after the testing. Where it is used in presenting and constructing statistical models of the outcome result performance of the Viola-Jones algorithm compared to the enhanced algorithm.

4.2 Data Analysis and Presentation

The researchers used tables to present the data gathered throughout the experimentation and analyzed it by calculating the accuracy rate of both algorithms. This study is concerned with gathering the frequency of the total faces that appeared in the video and the images. It then counted the tilted and frontal faces separately, followed by the faces detected by both algorithms.

INTERNATIONAL JOURNAL ON ORANGE TECHNOLOGY

https://journals.researchparks.org/index.php/IJOT e-ISSN: 2615-8140 | p-ISSN: 2615-7071

Volume: 4 Issue: 6 |Jun 2022

4.3 LBP Cascade Training

With regards to the training of the LBP cascade classifier concerning the detection of angled faces, the researchers used 916 cropped positive images or images that contain angled faces from the CFPW dataset and 2032 non-face or negative images with the same aspect ratio as positive samples. The output of cascade training is an XML file that includes data regarding the objects to be detected. XML files are used to perform the detection.

4.4 Testing of the Algorithm (Data Gathering)

In the collection of data after the enhanced algorithm was formulated, the researchers first conducted a pretesting, which uses sets of test images from the same database used to train the cascade classifier containing multi-viewed cropped faces. This was done to find out if the algorithm was already in condition to detect multi-viewed faces without problems. When the researcher spotted errors in the enhanced part of the algorithm, debugging was done to improve it. After finalizing the enhanced algorithm, the researcher proceeds to the actual experimentation. The first experimentation was done using sets of real images that were taken randomly from the internet. The second experiment was done in CCTV footage with a duration of 3:35 minutes, which was set up at the entrance door of a shopping mall. For the frame selection, frames were selected every five seconds; if the frame does not contain faces, the selection skips for another five.

Table 1 shows an increase of 9% with the enhanced version of the Viola-Jones algorithm (VJA) when tested with images. Table 2 shows a huge gap with regards to the accuracy rate of face detection between the existing Viola-Jones algorithm and enhanced VJA with local binary pattern (LBP) when applied to a real-time application. With 55 faces detected by the proposed method and 33 faces identified by the existing algorithm, a difference of 34% is seen in the evaluated precision of face detection.

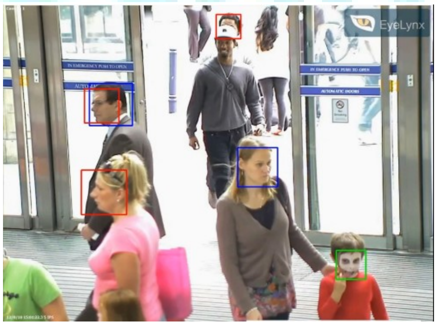

Fig. 4. (a) Left image - face detection using the existing viola-jones algorithm in a CCTV footage; (b) Right image - face detection using the enhanced viola-jones algorithm and local binary patterns in a CCTV footage

Figure 4a and Figure 4b present two different frames from the same CCTV footage. In Figure 4a, the existing Viola-Jones algorithm only detects one face that is frontal and fails to identify the others. In Figure 4b, the enhanced VJA with a local binary pattern identified five faces up to a 90° angle. (Green square indicates the detection from existing Viola-Jones; blue square: face detected is on the right side; red square for the left side.)

5. Conclusion

This study addresses two face detection approaches and algorithms, the existing Viola-Jones algorithm and VJA with Local Binary Pattern (LBP). The purpose is to detect faces that are tilted greater than an angle of 45 degrees. This was done by making use of the very same face databases in a comparative study of the two methods. The database includes multiple face photos and videos of diverse faces, such as frontal and angled. The results were assessed based on the preciseness and reliability of both techniques. Based on the total data, it can be stated that the enhanced algorithm has a combined 49% higher accuracy rate than the original algorithm. The proposed

https://journals.researchparks.org/index.php/IJOT e-ISSN: 2615-8140 | p-ISSN: 2615-7071
Volume: 4 Issue: 6 |Jun 2022

method also spotted 35 more faces than the pioneer algorithm. From this, the authors conclude that the enhanced Viola-Jones algorithm with local binary pattern outperforms the existing Viola-Jones algorithm in the matter of accuracy rate of detection in images as well as in real-time application.

Acknowledgements

The researchers would like to acknowledge and give thanks to God for His blessings and guidance throughout this research and its successful completion. The researchers would also like to express their gratitude to their advisers and to express their appreciation to them for sharing their facilitative comments and suggestions for further improvements. The researcher's gratitude also extends to their family and friends for their support and encouragement.

References

1. Aashish, K., Vijayalakshmi, A., 2017. Comparison of Viola-Jones and Kanade-Lucas- Tomasi Face Detection Algorithms. Oriental Journal of Computer Science and Technology. http://www.computerscijournal.org/vol10no1/comparison-of-viola-jones-and-kanade-lucas-tomasi-face-detection-algorithms/
2. Damanik, R. R., Sitanggang, D., Pasaribu, H., Siagian, H., Gulo, F., 2018. An application of viola jones method for face recognition for absence process efficiency. Journal of Physics: Conference Series. p. 3. https://iopscience.iop.org/article/10.1088/1742- 6596/1007/1/012013/pdf
3. Gupta, R. 2019. Breaking Down Facial Recognition: The Viola-Jones Algorithm. Towards Data Science. https://towardsdatascience.com/the-intuition-behind-facial-detection-the-viola-jones-algorithm-29d9106b6999
4. Islam, M. A., Naeem, A., Hasan, N., 2017. Comparison Between Viola-Jones and KLT Algorithms and Error Correction of Viola-Jones algorithm. IJRDO - Journal of Computer Science Engineering, p. 26. https://www.ijrdo.org/index.php/cse/article/view/499
5. Kadir, K. Kamaruddin, M. K., Nasir, H., Safie, S. I., Bakti, Z. A. K., 2014. A Comparative Study between LBP and Haar-like features for Face Detection Using OpenCV. 2014 4th International Conference on Engineering Technology and Technopreneurship (ICE2T). p.336-337. https://www.researchgate.net/publication/308836179_A_comparative_study_between_LBP_and_Haar-like_features_for_Face_Detection_using_OpenCV
6. Lopez, L. S., 2010. Local Binary Patterns applied to Face Detection and Recognition. Universitat Politecnica de Catalunya. p. 7. https://core.ac.uk/download/pdf/41809268.pdf
7. Prado, K.S., 2017. Face Recognition: Understanding LBPH Algorithm. Towards Data Science. https://towardsdatascience.com/face-recognition-how-lbph-works-90ec258c3d6b
8. Viola, P., Jones, M., 2001. Rapid Object Detection using a Boosted Cascade of Simple Features. Accepted Conference on Computer Vision and Pattern Recognition 2001. p. 2-3. https://www.cs.cmu.edu/~efros/courses/LBMV07/Papers/viola-cvpr-01.pdf