

## Face Detection Keypoints Using Dct And Clahe

Adhi Kusnadi, Liana Nathania, Ivransa Zuhdi Pane, Marlinda Vasty Overbeek

Universitas Multimedia Nusantara, Tangerang, Indonesia

**Article History:** Received: 11 January 2021; Revised: 12 February 2021; Accepted: 27 March 2021; Published online: 10 May 2021

### ABSTRACT

The research was used to overcome illuminations on facial images are continuing. This testing was related to the level of illumination variation, which was one of the important factors that given an effect on the accuracy of face detection keypoints. The dataset used for this study was taken from the front side of the object. The detection result quality of the datasets will be improved by handling the illumination variations using the DCT algorithm and the CLAHE algorithm. In the process of handling the illumination's problem, the accuracy result of five feature detectors in detecting facial keypoints were also evaluated. The research parameters to be calculated consist of recall, precision, and F-score. Testing is needed to prove that the F-score will increase after two images processing methods that applied. The test results proved that the combined application of DCT and CLAHE could increase the level of key point detection in the SURF algorithm.

### INTRODUCTION

Face detection keypoints research is an important research topic to study related to its benefits in technologies of face detection. The face detection technology holds the largest share in marketing technology today. Several techniques have been proposed to detect face points, especially for 2D image detection[1].The results of the detection accuracy are influenced by external factors such as the shape of the face, the face expressions, and the illumination found on the face image[2][3]. The illumination factor in an image is the most influential factor, because the variations caused by illumination are more significant than the physical features of the individual's own face[4][5]. Face detection will have worse results on images with poor illumination[2]. Therefore, a method is needed to overcome the illumination variations to get a high accuracy result of the face detection.

In previous studies[4][6][2][7][8][9][10][11][3],some methods have proposed to overcome variations in facial image illumination. There is also by cutting out some low frequency parts from DCT's logarithmic domain that are same as eliminating variations of the existing illumination. However, not all of the methods have been applied to the detection of keypoints that is an important factor in the face recognition process. In the research [12][13][14][15][16], the image keypoints of original images have been detected by some popular feature detectors. However, the detection accuracy obtained was not satisfactory because the image used was still original and the image processing method had not been applied at all. Furthermore, by combining the two methods from the two previous studies, this study will test the effect of eliminating some low-frequency variations in the DCT to improve the performance of face detection keypoints by feature detectors to increase the accuracy values. In this research five feature detectors are tested, Speeded Up Robust Features (SURF), Harris-Stephans, Features from Accelerated Segment Test (FAST), Binary Robust Invariant Scalable Keypoints (BRISK) and Minimum Eigen Value.

In dealing with illumination variations, first method that used to solve this problem is to implement the low frequency variations contained in the Discrete Cosine Transform (DCT) algorithm. DCT in image processing will transform the image domain from the spatial domain to the frequency domain. Due to variations in the illumination of an image found in the low frequency section, a number of coefficients at the low frequency DCT (25, 50 and 75) contained in the facial image discarded to compensate for variations in illumination. This method is also known as illumination normalization which aims to generalize illumination in the image[17].

Another method that used in handling variations of facial illuminations is the CLAHE algorithm. This algorithm will handle variations of facial illuminations by aligning the contrast levels in the image [18]. The image which being applied DCT will become a low contrast image that means the histogram will gather in one place. This causes the grey level to be uneven, so CLAHE needed to equalize the grey level. That is

because a good image has a histogram that fills the grey area in full with an even distribution on each pixel intensity value.

Both methods have been success tested on the image datasets from the 2D database that is Head-Pose Database. It comes out the F score value run into a significant increase when compared with the result in the original image which has not been applied any image processing. It is because both the low frequency variations in DCT algorithm and the CLAHE Algorithm have successfully handled variations of illumination that gives a considerable impact on the image, especially on the image feature that becomes more obvious to be detected.

The benefit of this research is to improve the accuracy of face detection keypoints by handling illumination variations and contrast differences in facial images. Along with the increased of the keypoints detection accuracy result, this study also provides benefits in improving the quality of feature detectors and enhancing the face detection system.

## LITERATURE STUDY

### 1.1. ILLUMINATION

The process of face recognition is greatly changed by changes in facial appearance caused by variations in illumination. Variation of illumination means the facial image consists of various areas of the face with different lighting. Meanwhile, to get a good face detection result, the illumination on the face image must evenly distribute across the surface of the image. Illumination can also consider as a complex factor on face images taken indoors and outdoors. Face recognition is more difficult for images with poor illumination variations. Shah et.al., has discussed about bad illumination which affects the accuracy of face detection results must be handled by handling variations of the illumination on the facial image[11].

### 1.2. YCbCr

The image conversion into YCbCr colour space from RGB colour space is done to get a compressed image file but still can store most of the information from the original image. After the conversion process, the channel that used in research is only the Y channel because most of the information received by the human eye is the luminance component on the Y channel not the chrominance on the Cb and Cr Channel. The image consists of digital pixels represented in the RGB format, where 0 and 255 respectively represent black and white so that the Y channel component can obtain according to the following equation.

$$[Y] = [16] + [0.257 \ 0.504 \ 0.098] \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (1)$$

### 1.3. KEYPOINTS

The platform called Kaggle has provided references regarding 15 facial keypoints that must be detected [19]. The more keypoints detected, the higher the accuracy of the detection. Conversely, the more points other than face keypoints are detected then the detection results will decrease in value. The 15 facial keypoints are shown in Table 1.

**Table 1. 15 Facial Keypoints**

Left Eye Centre	Right Eye Centre
Left Eye Inner Corner	Right Eye Inner Corner
Left Eye Outer Corner	Right Eye Outer Corner
Left Eyebrow Inner End	Right Eyebrow Inner End
Left Eyebrow Outer End	Right Eyebrow Outer End
Mouth Left Corner	Mouth Right Corner
Mouth Centre Bottom Lip	Mouth Centre Top Lip
Nose Tip	

### 1.4. DCT

A.Thamizharasi&Jayasudha has discussed about facial image detection which can be supported by applying the DCT algorithm[18]. DCT represents an image into the frequency domain consisting of 3 frequency bands namely low, medium and high frequency. In relation to variations of illumination, a number of DCT

coefficients at low frequencies will remove to minimize the variation of existing illuminations. By minimizing variation in illumination means handling the illumination variations on the entire surface of the image. It can be done only by setting DCT components that are at low frequencies to zero. This method is also known as illumination normalization using DCT algorithm. The normalization process is implemented through DCT coefficients of 25, 50 and 75 which shown in Figure 1.

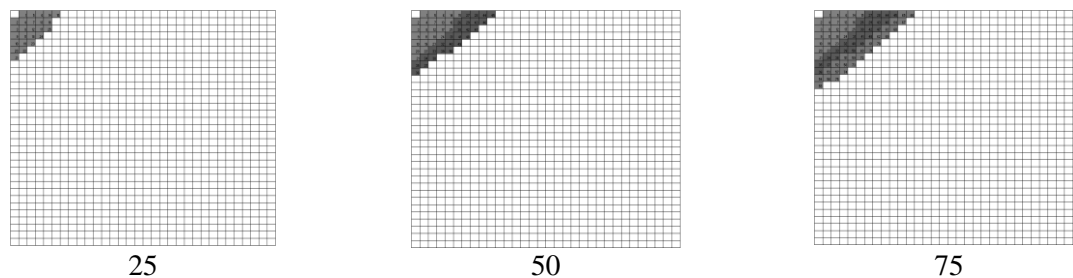


Figure 1. Frequency of DCT Domain with Variety Coefficient

### 1.5. CLAHE

Sonali et.al., has discussed about the CLAHE method which used to handle contrast problems that are not aligned to the image. CLAHE works by limiting the increase in contrast that usually done by Histogram Equalization. Thus, increased contrast in facial images can be limited as needed to handle illumination variations[20]. Other research also proposes CLAHE to be used in the face detection process[18]. The results obtained after applying the CLAHE method produce an accuracy value of 0.985. High accuracy results show that images processed with CLAHE can properly handle illumination variations.

### 1.6. FEATURE DETECTORS

#### 1.6.1. SURF

SURF consists of the keypoints detector and its descriptors. The detector places the keypoints in the image, while the descriptors describe the features of the keypoints and create feature vectors of the keypoints. The descriptors owned by SURF are to overcome rotation, brightness, and afterwards are reduced to units of contrast value length[21].

#### 1.6.2. HARRIS-STEPHANS

The Harris-Stephans Detector detects points based on variations in intensity of the existing pixel of an image. Small areas of the feature must show a large intensity difference when compared to other pixels through various directions. This is because the Harris detector is basically finding the difference in intensity for the displacement of points in all directions[22].

#### 1.6.3. BRISK

The BRISK detector detects strong keypoints for all images that can transform. The sampling pattern consists of points located in concentric circles applied to the area around the point to pick up the grey values. The BRISK oriented sampling pattern used to get the results collected in the descriptors. BRISK descriptors are binary, so keypoints are very efficiently matched[23].

#### 1.6.4. FAST

To detect the features of an image, the FAST detector operates on top of all image elements. This aims to identify whether the pixel  $p$  in the image can do intensity comparison with a predetermined thresholding value. From the results of the comparison it can be determined which be referred to as a keypoints feature or not. Fast works using the Brensenham circle method. Which means do the pixel point is the keypoints[24].

#### 1.6.5. MINIMUM EIGEN VALUE

The Minimum Eigen Value Detector was developed by Jianbo Shi and Carlo Tomasi and also called the Shi-Tomasi corner point detector. This method of detecting features stands out from the rest because it uses different features to monitor the quality of image features. The RMS residue from the first frames and currently quantifies changes in the appearance of features. This emphasizes the fact that this feature is categorized as a good feature if the keypoints tracker is functioning properly[22].

### 1.7. EVALUATION CRITERIA

To evaluate the performance of the detector features, number of detected keypoints will be calculated. The calculation of F-score using recall and precision. In the recall and precision equation, there are true positive

(tp), false positive (fp) and false negative (fn). True positive obtained from the number of detected keypoints. False positive is obtained from the number of points to be detected (in this case 15 points according to the facial keypoints table) reduced by the number of detected keypoints. False negatives are obtained from the number of points detected reduced by the number of detected keypoints.

Recall is calculated from the number of tp divided by the number of tp plus fn[25].

$$\text{recall} = \frac{tp}{tp+fn} \tag{2}$$

Precision is calculated from the number of tp divided by the number of tp plus fp[25].

$$\text{precision} = \frac{tp}{tp+fp} \tag{3}$$

F-score calculated from the number of recall and precision values obtained. This value will also be the value of the detection accuracy of keypoints at the end of this study [12].

$$\text{FScore} = 2 \times \frac{(\text{Recall} \times \text{Precision})}{(\text{Recall} + \text{Precision})} \tag{4}$$

**METHOD**

2D facial images accessed from the database. The images were chosen to be tested for handling illumination variations. These are the method used (Figure 2):

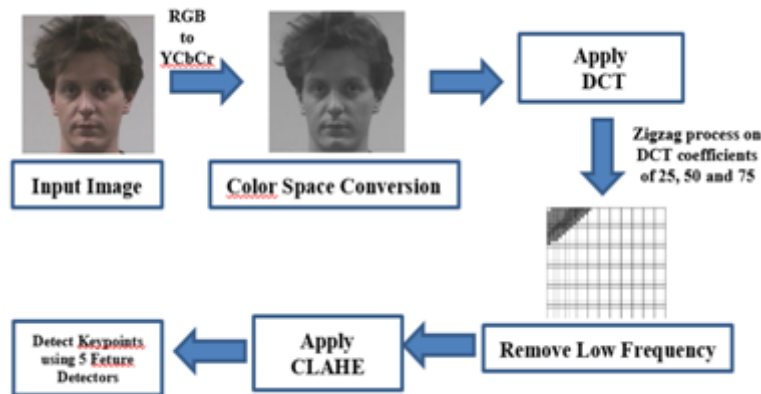


Figure 2. The Flow of Method

The procedure described as below:

1. Convert RGB colour space to YCbCr colour space and access the Y channel only because it has represents the colour scale that we need. The conversion calculation process is as in equation (1).
2. Apply the DCT algorithm to handle the illumination variations by using the DCT coefficients of 25, 50 and 75 and remove the low frequency.
3. Apply the CLAHE algorithm to handle the contrast alignment.
4. Detect facial images’ keypoints using five feature detectors.
5. Done.

**EXPERIMENTAL RESULT AND DISCUSSION**

Table 2 showed an example of calculation result using five feature detectors such as Harris, SURF, FAST, Minimum Eigen value, and BRISK.

**Table 2.Example of the keypoints calculations**

Feature Detectors		Harris S.	Surf	Fast	Brisk	Minimum E.V.
DCT Image	Point	15	15	15	15	15

	TP	1	6	0	2	1
FP		14	9	15	13	14
FN		14	9	15	13	14
Recall		0,067	0,4	0	0,13	0,067
Precision		0,067	0,4	0	0,13	0,067
F Score		0,067	0,4	-	0,13	0,067

The explanation of the calculation can be seen as below:

Define 15 points for each feature detector. It means that 15 points of facial image features detected by the feature detectors.

Count the keypoints detected on the image based on the facial keypoints listed in table 1 then set them as true positives (TP)

Based on the keypoints detected, count the false positive (FP)

Based on the points defined, count the false negative (FN)

Recall, precision and F-score counted based on the equation (2), (3), and (4).

1.8. Experiments on Head-Pose Database

In this test, Head-Pose database is used as many as 5 facial image datasets taken from 0's degree angle (the front side of the object's face). The five datasets applied to DCT algorithm with coefficients of 25, 50 and 75. After that, CLAHE algorithm will also applied. The entire dataset detected using five feature detectors. The datasets used are shown in Figure 3.



Figure 3. All Head-Pose's datasets used

Examples of dataset detected by the feature detectors on the original grey images are as follows in Figure 4.

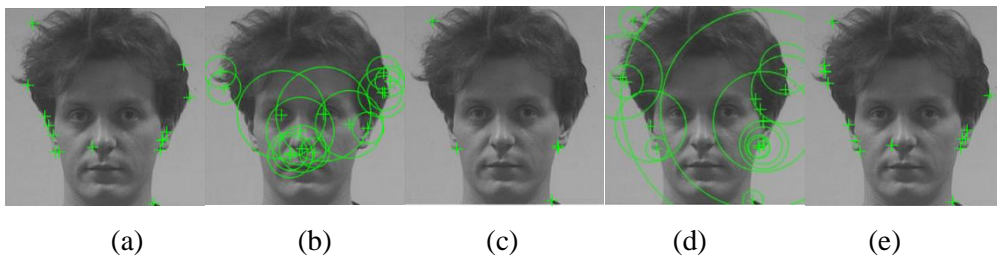


Figure 4. Example of keypoints detection by five feature detectors

The tables 3,4,5,6 and 7 showed the calculation result for each datasets. Calculations are performed on original grey datasets that have not been implemented with any image processing methods and calculations on datasets that have been implemented the DCT and CLAHE algorithms. The highest F-score test result for the original grey image on the five image bases is the SURF algorithm with average score 0,119. And the highest f-score average results after the DCT and CLAHE algorithm are implemented 0,23, increase 152% at DCT 25.

**Table 3. F-score result for First Dataset**

	Harris	SURF	FAST	BRISK	Min-Eigen
Original Grey	-	0,13	-	-	-
DCT 25	0,067	0,2	-	0,13	-

DCT 50	-	0,2	-	0,13	-
DCT 75	0,067	0,2	-	0,2	-

**Table 4. F-score result for Second Dataset**

	Harris	SURF	FAST	BRISK	Min-Eigen
Original Grey	-	0,13	-	-	-
DCT 25	0,067	0,4	-	0,13	0,067
DCT 50	-	0,33	-	0,2	-
DCT 75	0,067	0,13	-	-	-

**Table 5. F-score result for Third Dataset**

	Harris	SURF	FAST	BRISK	Min-Eigen
Original Grey	-	0,067	-	-	-
DCT 25	-	0,267	-	0,13	-
DCT 50	-	0,133	-	0,067	-
DCT 75	-	0,2	-	0,069	0,067

**Table 6. F-score result for Fourth Dataset**

	Harris	SURF	FAST	BRISK	Min-Eigen
Original Grey	-	0,067	-	0,067	-
DCT 25	-	0,267	-	0,207	0,069
DCT 50	-	0,13	-	0,2	-
DCT 75	0,067	0,33	-	0,13	-

**Table 7. F-score result for Fifth Dataset**

	Harris	SURF	FAST	BRISK	Min-Eigen
Original Grey	-	0,2	-	-	-
DCT 25	-	0,33	-	0,138	-
DCT 50	-	0,33	-	0,133	-
DCT 75	-	0,2	-	0,2	-

1.9. Analysis Results of the experiments on Head-Pose Database

Based on the experimental results that have been carried out, after implementing low frequency variations on the DCT and CLAHE algorithms, all the highest F-score results from each dataset have increased when compared to the F-score results from the original grey image before the two algorithms are implemented. From those result tables above, it can be concluded that the SURF detector has the best work performance by giving the highest F-score result with an increase of 152% at coefficient in DCT 25.

The results obtained by the image dataset are very varied and found many F-scores are still 0. It can be caused by various factors that exist. One of the strongest factors in influencing the results of detection of facial keypoints is the different shapes of humans' faces. The SURF detector can consistently give the highest F-score for each dataset because its performance is sensitive to differences in intensity. Through implementing image processing methods that handle variations of illumination (DCT and CLAHE), the two algorithms also produce image output that has clearer difference in intensity of the image. With the difference in intensity, the SURF detector becomes more accurate in distinguishing facial keypoints that must be detected and those that are not, so that the SURF detector's performance is more appropriate in detecting

keypoints. This also contributes to the fact that SURF detectors are superior when it comes to detecting keypoints when compared to other feature detectors.

## CONCLUSION

The implementation of low frequency variations on the DCT algorithm and CLAHE algorithm had been successfully implemented and succeeded in increasing the accuracy of face detection keypoints results. It can be seen from all the results of the highest F-score of each dataset used was successfully increased when compared with the results of the F-score of the original grey image that has not been implemented by the two algorithms. The accuracy of the final detection results increased by 152%. The highest F-score is produced by SURF detector as a detector with the best working performance. The DCT low frequency coefficient which is the most ideal in dealing with the problem of illumination variations on facial images is the DCT coefficient 25.

## ACKNOWLEDGMENTS

The authors gratefully acknowledge the support from Multimedia Nusantara University and Ristek-Dikti for grant PDUPT at the year of 2021.

## REFERENCES

- [1] M. Chihaoui, A. Elkefi, W. Bellil, and C. Ben Amar, "A survey of 2D face recognition techniques," *Computers*, vol. 5, no. 4, pp. 1–28, 2016.
- [2] A. Thamizharasi, "An Illumination Invariant Face Recognition by Selection of DCT Coefficients," 2016.
- [3] V. Štruc and N. P. C., "Photometric normalization techniques for illumination invariance Photometric Normalization Techniques for Illumination Invariance AVTOR: Vitomir Štruc INTERNAL REPORT : LUKS," no. November 2015, 2011.
- [4] V. P. Vishwakarma and T. Goel, "An efficient hybrid DWT-fuzzy filter in DCT domain based illumination normalization for face recognition," *Multimed. Tools Appl.*, vol. 78, no. 11, pp. 15213–15233, 2019.
- [5] L. N. V. Carneiro and G. Cámara-Chávez, "Logarithm Discrete Cosine Transform Domain and Discrimination Power Analysis for Illumination Invariant Face Recognition," *Proc. Int. Conf. Image Process. Comput. Vision, Pattern Recognit.*, pp. 1–7, 2012.
- [6] S. A. Khan, M. Ishtiaq, M. Nazir, and M. Shaheen, "Face recognition under varying expressions and illumination using particle swarm optimization," *J. Comput. Sci.*, vol. 28, pp. 94–100, Sep. 2018.
- [7] W. Chen, M. J. Er, and S. Wu, "Illumination compensation and normalization for robust face recognition using discrete cosine transform in logarithm domain," *IEEE Trans. Syst. Man, Cybern. Part B*, vol. 36, no. 2, pp. 458–466, 2006.
- [8] V. P. Vishwakarma, "Illumination normalization using fuzzy filter in DCT domain for face recognition," *Int. J. Mach. Learn. Cybern.*, vol. 6, no. 1, pp. 17–34, 2013.
- [9] S. Shan, W. Gao, B. Cao, and D. Zhao, "Illumination normalization for robust face recognition against varying lighting conditions," in *2003 IEEE International SOI Conference. Proceedings (Cat. No.03CH37443)*, 2003, pp. 157–164.
- [10] S. Du, M. Shehata, and W. Badawy, "A novel algorithm for illumination invariant DCT-based face recognition," in *2012 25th IEEE Canadian Conference on Electrical and Computer Engineering: Vision for a Greener Future, CCECE 2012*, 2012.
- [11] J. H. Shah, M. Sharif, M. Raza, M. Murtaza, and Saeed-Ur-rehman, "Robust face recognition technique under varying illumination," *J. Appl. Res. Technol.*, vol. 13, no. 1, pp. 97–105, 2015.
- [12] A. Kusnadi, Wella, R. Winantyo, and I. Z. Pane, "Evaluation of Feature Detectors on Repeatability Quality of Facial Keypoints in Triangulation Method," *2018 Int. Conf. Smart Comput. Electron. Enterp. ICSCEE 2018*, no. 1, pp. 1–4, 2018.
- [13] A. Kusnadi, R. Winantyo, and I. Z. Pane, "Removing DCT High Frequency on Feature Detector Repeatability Quality," in *2019 5th International Conference on New Media Studies (CONMEDIA)*,

- 2019, pp. 238–243.
- [14] A. Kusnadi, W. S. Kom, R. Winantyo, and I. Zuhdi Pane, “Enhancing the repeatability quality of feature detector in epipolar geometry,” *Utop. y Prax. Latinoam.*, vol. 24, no. Extra5, pp. 370–378, 2019.
- [15] N. H. Chan, K. Hasikin, and N. A. Kadri, “Evaluation of Feature Descriptor on D-Saddle Keypoint Detection in Retinal Image Registration,” *Proc. - 2019 IEEE 15th Int. Colloq. Signal Process. its Appl. CSPA 2019*, no. March, pp. 178–181, 2019.
- [16] A. Du, X. Huang, J. Zhang, L. Yao, and Q. Wu, “Kpsnet: Keypoint Detection and Feature Extraction for Point Cloud Registration,” in *Proceedings - International Conference on Image Processing, ICIP, 2019*, vol. 2019-Sept, pp. 2576–2580.
- [17] S. Dabbaghchian, A. Aghagolzadeh, and M. S. Moin, “Feature extraction using discrete cosine transform for face recognition,” *2007 9th Int. Symp. Signal Process. its Appl. ISSPA 2007, Proc.*, no. March, 2007.
- [18] A. Thamizharasi and D. Jayasudha, “An Illumination Invariant Face Recognition Using 2D Discrete Cosine Transform and Clahe,” *Int. J. Comput. Sci. Inf. Technol.*, vol. 8, no. 3, pp. 45–53, 2016.
- [19] Kaggle, “Facial Keypoints Detection,” 2016. .
- [20] Sonali, S. Sahu, A. K. Singh, S. P. Ghrera, and M. Elhoseny, “An approach for de-noising and contrast enhancement of retinal fundus image using CLAHE,” *Opt. Laser Technol.*, vol. 110, pp. 87–98, Feb. 2019.
- [21] R. C. Carro, J. M. A. Larios, E. B. Huerta, R. M. Caporal, and F. R. Cruz, “Face recognition using SURF,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 9225, pp. 316–326, 2015.
- [22] A. Vinay, N. Aklecha, Meghana, K. N. B. Murthy, and S. Natarajan, “On Detectors and Descriptors based Techniques for Face Recognition,” *Procedia Comput. Sci.*, vol. 132, pp. 908–917, 2018.
- [23] A. Vinay, A. R. Deshpande, B. S. Pranathi, H. Jha, K. N. B. Murthy, and S. Natarajan, “Effective Descriptors based Face Recognition Technique for Robotic Surveillance Systems,” *Procedia Comput. Sci.*, vol. 133, pp. 968–975, 2018.
- [24] A. Bhat, “Makeup Invariant Face Recognition using Features from Accelerated Segment Test and Eigen Vectors,” *Int. J. Image Graph.*, vol. 17, no. 01, p. 1750005, 2017.
- [25] C. D. Manning, H. Schütze, and G. Weikurn, “Foundations of Statistical Natural Language Processing,” *SIGMOD Rec.*, vol. 31, no. 3, pp. 37–38, 2002.