Open Access
Baghdad Science Journal **Vol.16 (4) Supplement 2019**

DOI: http://dx.doi.org/10.21123/bsj.2019.16.4(Suppl.).1087

Eye Detection using Helmholtz Principle

Hanaa M. Ahmed * ***Shrooq R. Hameed***

Received 20/10/2018, Accepted 11/6/2019, Published /12/2019

Abstract:

Eye Detection is used in many applications like pattern recognition, biometric, surveillance system and many other systems. In this paper, a new method is presented to detect and extract the overall shape of one eye from image depending on two principles Helmholtz & Gestalt. According to the principle of perception by Helmholz, any observed geometric shape is perceptually "meaningful" if its repetition number is very small in image with random distribution. To achieve this goal, Gestalt Principle states that humans see things either through grouping its similar elements or recognize patterns. In general, according to Gestalt Principle, humans see things through general description of these things. This paper utilizes these two principles to recognize and extract eye part from image. Java programming language and OpenCV library for image processing are used for this purpose. Good results are obtained from this proposed method, where 88.89% was obtained as a detection rate taking into account that the average execution time is about 0.23 in seconds.

Key words: Contour Extraction, Gestalt Theory, Helmholtz Principle, Skin Detection.

Introduction:

Eye detection is an important step in various human-computer interaction applications (1). Different approaches for eye detection have been proposed in the recent years. Eye detection is an essential step in many applications like: face detection/recognition, face expression analysis, gaze estimation, human interactions, pose estimation, biometric systems, character animation and surveillance systems (2, 3). In general, eye detection methods can be divided into four major categories, Shape-based, Appearance-based, structural-based, and Context-based approaches. The geometric shape of Iris is captured in Shape-based approach. Appearance-based approach, in turn, depends on color distribution and filter responses to get photometric appearance. Context-based approaches benefits from both the first and second approaches to find a relationship between eye center and other facial parts. Lastly, structural-based approach localizes eye through structural location information with respect to nose, mouth, etc.(4) In this paper, a different method based on Helmholtz principle, in which an object is detected whenever this object has large deviation from randomness, is used to detect eye in image with face being partially included.

Department of Applied Science, University of Technology, Baghdad, Iraq.
*Corresponding author: 110113@uotechnology.edu.iq

This method is based on two essential steps: in the first step, the skin is detected, the image is binarily divided into skin and non-skin regions, blurring and morphological operations applied to eliminate noise, and contours are extracted. In the second step, the non-skin regions are filtered, candidate eye region is detected using Helmholtz principle, finally actual eye region is cropped. 20 images are used for testing. The detection rate is about 88.89% obtained, and the average execution time is about 0.23 in second. Image must contain close up frontal eye, the distance between the camera and eye should not exceed five cm.

Related Work:

Nasiri used YCbCr color space then made two maps according to these components merging them to obtain new map. The result of the last step is that, candidates are generated. Finally threshold and geometric test are applied on these candidates. The detection rate of 50 people images was about 98.5%, and average execution time was 1.395 in second (5). Nobuo Funabiki, extracted eye contour from face image using deformable template matching, template parameters were optimized by using local search method in addition to hill climbing and tabu search to extract the contour of eye. The average execution time of eight face images was 25.08 seconds (6). Mingxin YU.

proposed a hybrid eye detection method including two steps. In the first, gray intensity variance filter was applied to eliminate non-eye region from image. In the second step, accurate two eye regions were determined through trained support vector machine classifier. This proposed method was tested on BioID face images with glasses and obtained detection rate was about 92.5% (7). Dodiya, Bhagirathi. proposed a method for detecting face, eye, and iris in an image sequence. For this purpose, thresholding applied, bilateral filter was used to remove noise and blob analysis was used to detect eye region and face region and then circular Hough transform was used to detect Iris. Dodiya concluded that blob analysis was a very fast technique to detect face and eye regions (8). In (9), fast and accurate method for extracting center of the pupil, iris radius, and the shape of the eye was proposed. The method consists of two stages. In the first stage, fast circular symmetry detector used to localize pupil center. And radial gradient projections used to compute the iris radius. While, in the second stage, based on shape and color information Monte Carlo sampling framework determined the shape of the eye. In (10), self-similarity scores were computed in each pixel of each image. Then, differential analysis was applied for the intensity level. The result of this step was normalized and integrated in a joint representation which depends on smoothing with Gaussian kernel in order to present circular regions, and then peaks were found which correspond to the eye center. In (11), eye location was evaluated from facial images. A group of candidate regions was proposed, and then set of convolutional neural network was applied to determine eye region. Finally, the center of the eye was located using another convolutional neural network.

Image Acquisition and Image Resize:

Image Acquisition:

The images of one eye were taken by mobile camera with resolution 3264×1836. The image is in Blue, Green, Red (BGR) color space. As shown in Fig.1.

Figure 1. Original Image.

Image Resize:

The image is resized from (3264×1836) to suitable size. The image was taken by mobile camera, when converted from bitmap to Matrix, it is down sampled.

Techniques used in First Step:

Skin Detection:

The captured image was converted from Blue, Green, and Red (BGR) color space to Hue, Saturation, Value (HSV) color space as shown in Fig.2, since it is more related to human color perception (12). HUE channel is extracted as shown in Fig.3, and threshold is applied to recognize skin region from non-skin region. In fact, skin region varies between 6 and 38. The skin region is converted to zero and non-skin region is converted to one resulting in binary image as shown in Fig.4.

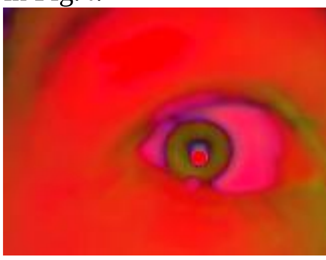

Figure 2. HSV Color Space.

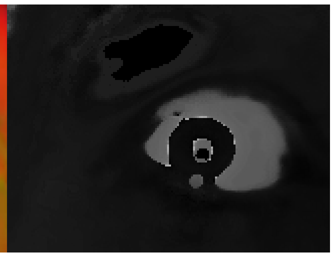

Figure 3. Hue Channel of Original Image.

Figure 4. Binary Image, in which white pixel represent eye and black pixel represent skin.

Median Blur:

Median blur is non-linear filter in which it takes the pixel value of gray level of the median in a window, and replaces the center of the window with that value. Median Filter algorithm has the following steps: (13)

Take a window of size $n \times n$ of image pixels.

Sort the pixel values of that window from smaller to larger.

Take the value of the median.

Put the value of the median in the pixel center of the window.

Move the window over all image area.

Repeat steps (1-5) until reach the end of the image.

Median filter is more efficient than the other types of linear filter since it can remove certain types of noise with less blurring of images with the same size (14, 15).

Dilation and Erosion:

They are fundamental morphological operations. The term morphology indicates the analysis of the shape of an object in digital images from a scientific point of view (16). They are operating on binary image. Dilation is the process of stretching objects by filling gaps among objects. So, it is possible to connect disconnected objects. Erosion, on the other hand, is the process of shrinking objects by eroding their boundaries. Both of erosion and dilation depend on structuring element (17, 18). Structuring element can be considered as a probe whose slides across the image test the spatial nature of the image at each pixel (19). Dilation operation requires the image to be dilated and have structuring element. The last one move across the image elements in a way like convolution. The result of dilating input binary image with structuring element is a set of all points of the input binary image such that the intersection of structuring element with the input binary image is not empty. Dilation is defined as set operation. Let M an image to be dilated by S structure element, written as$M \oplus S$, is defined as (18):

$M \oplus S = \left\{Z|\left(\hat{S}\right)_z \cap M \neq \emptyset\right\}$….. (1)

Where $\emptyset$ is an empty set, $\hat{S}$ is for the reflection of collection S.

Erosion like Dilation also requires two inputs binary image and structuring element, and the process between the two inputs similar to convolution. The result of Erosion input binary image with structuring element is a set of all points of the input binary image such that the structuring element is the subset of the input binary image. M is eroded by S, recorded as $M \ominus S$, and defined as(18):

$M \ominus S = \{z|(S)_z \cap M^c \neq \emptyset\}$…. (2)

$\emptyset$ is for the empty set, M^c is for the supplement of collection M.

Contour Extraction:

Contour is a piece of line whose length can be one or more pixels, and its wide is exactly one pixel. The boundary of the object can be represented by unbroken contour (20). There are many applications of contours; they disjoint many objects from each other and from background. By using contours, the size, area, and perimeter of the object can be calculated; they can also classify the shape of an object using contours. So, it is very important in the field of object representation and image recognition (21).

The Helmholtz principle:

The Helmholtz Principle explains how the perception of human works. It hypotheses that if we consider randomness as normal case, human eye should not observe meaningful events, and if these events were observed by human eye, they would be with a small probability. The small probability does not mean that these events are meaningful (22). Gestalt theory says that whenever a group of points have a common feature (like color, parallelism, alignment, closure, convexity, orientation or any other feature) this group of points can be grouped together to form an individual object. Grouping principle is recursive. For instance, if a group of points forms a line, recursively these lines can be grouped together depending on another gestalt law (like parallelism). Gestalt theory allows us to compute qualitative geometric primitives in any image. The Helmholtz principle formalizes these qualitative grouping principles. The grouping principles are stated as follows: when N atomic objects $O_1, O_2, \dots, O_n$ are present in an image, and K of them have a common feature. An important question arises here: Does a common feature occur by chance or is it a meaningful feature? To answer this question, we consider the quality to be spatial (position, orientation) distributed uniformly and randomly on all the Nth objects. This leads to another question: Is the observed repetition probable or not? The Helmholtz principle answers these questions such that if the observed configuration of Kth objects is very small, then, the grouping of these objects gives gestalt (23). In the cognitive system, the human eye percepts image in two sequential steps. In the first step, the human eye processes the overall information in an image quickly. In the second, the human eye concentrates only on one part of the available image information that is more interesting (24, 25). When a psychologist picks an object from the physical world and shows it to a person, the person says what he/she sees. A person's perception makes a gestalt in his/her mind and when he/she describes the object, he/she does not describe the visual stimulus; in fact, he/she describes the general description of the object (26).

In this paper, we follow the following procedure. In the first step, we distinguish between skin region and non-skin region. One of non-skin regions contains eye shape. So, in the Second step, we use the same color law of gestalt theory to detect the eye Sclera.

Methodology:

In this paper, a new method to detect human eye is introduced. For this purpose java programming language and opencv library for image processing are used. The image is resized to a suitable size. The image is taken by mobile camera in BGR color space then converted to HSV color space as shown in Fig.2. HUE channel is extracted to distinguish between skin region and non-skin region as shown in Fig.3. Pixel values of skin region in HUE channel vary between 6 and 38. The result of this step is a binary image in which the black represents skin region and the white represents non-skin region as shown in Fig.4. In the first step, we exclude the skin region. The non-skin region may contain eye region, and may or may not have eyebrow and hair. How to localize eye region among these regions? First of all, median filter and morphological operations are applied on a binary image resulting in an image as shown in Fig.5. Boundaries of each non-skin region are found by finding contour of each non-skin region as shown in Fig.6. Contours that lie nearby the four boundaries of the original image are excluded because there is a high probability that these contours are not an eye (may be hair, eye brow or any other thing that does not belong to face at all) as shown in Fig.7. Then the remaining non-skin regions that are nearby to each other are merged by 40 pixels in x-axis and five pixels in y-axis at maximum. The resulting image is shown in Fig.8. Finally, the only largest contour is selected as a candidate eye region.

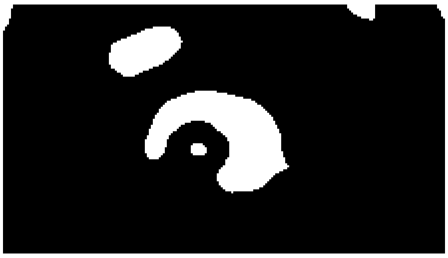

Figure 5. Image after applying Median blur, Erode, and Dilate Operations.

Figure 6. Contours of non-skin region.

Figure 7. Result of excluding contours that on the edges of the image.

Figure 8. Merge nearby contours.

After finding the candidate eye region, we apply Helmholtz principle to localize eye in order to ensure that the candidate eye region is actually the eye region itself. The grouping principle is applied on candidate eye region. Since sclera of the eye has a color near to white, it can be easily detected and distinguished using Hue channel. Pixel value of sclera in Hue channel is greater than 38, and there is no other thing in the image that has a pixel value near that value. For this reason the color of eye sclera is detected to extract the region of the overall eye. Helmholtz principle is used for this purpose, so, the number of false alarm (NFA) can be calculated as:

$$NFA = Ng \times [N \times \left(\frac{size\ of\ Sclera\ object}{size\ of\ entire\ image-1}\right)^K] \quad \ldots.(3)$$

Where N is the number of rows in the original image, while K is the number of image rows that contain eye Sclera pixels, and N_g is the number of all possible gestalts that have series equal or greater than the threshold (the threshold equal to 20). We estimate various gestalts, such as:

– Long enough series of Sclera pixels;
– Long enough series of non-Sclera pixels;
– Long enough series of alternate of non-Sclera and Sclera;
– Long enough series of alternate of Sclera and of non-Sclera;
– Long enough series of alternate pairs Sclera-Sclera- non_Sclera - non_Sclera;
– Long series of alternate triples;
– Long enough series alternating one Sclera and two non- Sclera;
– Long enough series alternating one non- Sclera and two Sclera.

The smaller the NFA is, the more meaningful the event is. Finally, eye region cropped from the original image as shown in Fig. 9.

Figure 9. Eye region cropped from Original image.

Experimental Results:

Twenty images taken by using mobile camera with resolution (3264×1836) are used for testing. We obtain 88.89% as a detection rate, and the average execution time is about 0.23 in second. On the other hand, the distance between eye and camera should not exceed five cm, and the light condition affects the result. The experimental result shows samples of images used for testing as shown in Fig.10. The comparison between our method and the state-of-the-art methods is presented in Table 1.

Figure 10. Image samples used to test the proposed method.

Table 1. The comparison between our method and the state-of-the-art methods.

Method	Detection Rate	Time (ms)
(3)	85.7 %	66
(4)	91.2 %	*
(7)	97.5 %	*
(9)	74.6594 %	*
(10)	85.29 %	*
(11)	85.6 %	13
our	88.89 %	230

Conclusion:

In this paper, a new eye detection method is presented, in which, eye Sclera is detected using Helmholtz principle. It is applied on up to 20 images, and it gives a good result. It can be used to detect eye in image where partial face is presented. The constraints of this paper are: First, image must contain close up frontal eye where the distance between the camera and eye should not exceed five cm. Second, it is affected by light condition as shown in Fig.8, in which some of the skin regions are detected as non-skin region because of the light condition (light is too bright). This problem is overcome by considering larger contour resulting from merging nearby contours which represent eye sclera as eye region. The merged contours contain the eye pupil. So, it becomes the largest contour in the image as shown in Fig.8.

Conflicts of Interest: None.

References:

1. Quan D, Zhao Y, Li J, Zheng H. An Effective Method and Device for Eye Detection with Physical Properties. In Proceedings of the 2nd International Conference on Computer Science and Electronics Engineering 2013 Mar 3. Atlantis Press.
2. Song F, Tan X, Chen S, Zhou ZH. A literature survey on robust and efficient eye localization in real-life scenarios. Pattern Recognition. 2013 Dec 1; 46(12):3157-73. Available from: http://dx.doi.org/10.1016/j.patcog.2013.05.009.
3. Markuš N, Frljak M, Pandžić IS, Ahlberg J. Eye pupil localization with an ensemble of randomized trees. Pattern recognition. 2014 Feb 1; 47(2):578-87. Available from: http://dx.doi.org/10.1016/j.patcog.2013.08.008.
4. Gou C, Wu Y, Wang K, Wang K, Wang FY, Ji Q. A joint cascaded framework for simultaneous eye detection and eye state estimation. Pattern Recognition. 2017 Jul 1; 67:23-31. Available from: http://dx.doi.org/10.1016/j.patcog.2017.01.023.
5. Nasiri JA, Khanchi S, Pourreza HR. Eye detection algorithm on facial color images. In2008 Second Asia International Conference on Modelling & Simulation (AMS) 2008 May 13 (pp. 344-349).

6. Funabiki N, Isogai M, Higashino T, Oda M. An eye-contour extraction algorithm from face image using deformable template matching. MEM.FAC.ENG.OKA.UNI. 2006; 40(1):78-82.
7. Mingxin YU, Yingzi LI, Xiangzhou WA. An efficient hybrid eye detection method. Turk J Elec Eng & Comp Sci. 2016 May 1; 24(3): 1586 - 1603. available from: http://journals.tubitak.gov.tr/elektrik/
8. Bhagirathi D, Malhan A, Jimmy P. Human face, eye and iris detection in real-time using image processing. Int J Eng Res Appl. 2014 May; 4(5):27-31.Available from: https://core.ac.uk/download/pdf/25928633.pdf.
9. Borza D, Darabant A, Danescu R. Real-time detection and measurement of eye features from color images. Sensors. 2016 Jul; 16(7):1105. Available from: www.mdpi.com/journal/sensors.
10. Leo M, Cazzato D, De Marco T, Distante C. Unsupervised eye pupil localization through differential geometry and local self-similarity matching. PloS one. 2014 Aug 14; 9(8):e102829. Available from: https://www.researchgate.net/publication/264787233.
11. Li B, Fu H. Real time eye detector with cascaded Convolutional Neural Networks. Applied Computational Intelligence and Soft Computing. 2018; 2018.Available from: https://doi.org/10.1155/2018/1439312.
12. Albiol A, Torres L, Delp EJ. Optimum color spaces for skin detection. In Proceedings 2001 International Conference on Image Processing (Cat. No. 01CH37205). 2001; 1, 122-124).
13. Odat A, Otair M, Shehadeh F. Image Denoising by Comprehensive Median Filter. IJAER. 2015; 10(15):36016-22.Available from: https://www.researchgate.net/publication/283099828.
14. Yang R, Yin L, Gabbouj M, Astola J. Optimal weighted median filters under structural constraints. In1993 IEEE International Symposium on Circuits and Systems 1993 May 3 ; 942-945.
15. Ng PE, Ma KK. A switching median filter with boundary discriminative noise detection for extremely corrupted images. IEEE Transactions on image processing. 2006 Jun; 15(6):1506-16.
16. Poshamallu G. Binary Image Processing Implementation on FPGA Using Morphological Dilation and Erosion Techniques. IJESC. 2016 April; 6(4).
17. Gil JY, Kimmel R. Efficient dilation, erosion, opening, and closing algorithms. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2002 Dec; 24(12):1606-17.
18. Raid AM, Khedr WM, El-Dosuky MA, Aoud M. Image restoration based on morphological operations. IJCSEIT. 2014 Jun;4(3):9-21.
19. Bartoo G, Kim Y, Haralick R, Nochlin D. Mathematical Morphology Techniques for Image Processing Applications in Biomedical Imaging. SPIE.1988; 914.
20. Seo J, Chae S, Shim J, Kim D, Cheong C, Han TD. Fast contour-tracing algorithm based on a pixel-following method for image sensors. Sensors. 2016 March 9; 16(3):353. Available from: www.mdpi.com/journal/sensors.
21. Deogade M, Pabale D, Jadhav S, Rani M. Dimension Detection of Object using Image Processing. IJIRT.2019 May; 5(12).
22. Gerogiannis DP, Nikou C, Likas A. Elimination of Outliers from 2-D Point Sets Using the Helmholtz Principle. Signal Processing Letters. 2015 Oct; 22(10):1638-42.
23. Desolneux A, Moisan L, Morel JM. Maximal meaningful events and applications to image analysis. The Annals of Statistics. 2003; 31(6):1822-51.
24. Wang Z, Li B. A two-stage approach to saliency detection in images. International Conference on Acoustics, Speech and Signal Processing. 2008 Mar 31.pp: 965-968.
25. van der Heijden AH. Two stages in visual information processing and visual perception. Visual Cognition. 1996 Dec 1;3(4):325-62. Available from: http://dx.doi.org/10.1080/135062896395625.
26. Guberman, S. On Gestalt theory principles. Gestalt Theory, 2015, 37(1): 25-44.

أستكشاف العين بأستخدام مبدأ Helmholtz

هناء محسن احمد **شروق رشيد حميد**

قسم علوم الحاسوب، الجامعة التكنلوجيا، بغداد، العراق.

الخلاصة:

كشف العين استخدم في تطبيقات متعددة مثل تمييز الأنماط ، البصمة، وأنظمة المراقبة والعديد من الأنظمة الأخرى. في هذه المقالة ،تم تقديم طريقة جديدة لتحديد العين واستخلاص الشكل الخارجي لعين واحدة من الصورة بالأعتماد على مبدئين هما Helmholtz و Gestalt. وفقا لمبدأ الأدراك لـ Helmholtz أنه أي شكل هندسي ملاحظ يكون ذو معنى أدراكيا أذا كان عدد مرات تكراره ضئيل جدا في صورة ذات توزيع عشوائي. لتحقيق هذا الهدف مبدأ Gestalt الذى ينص على أن الأنسان يلاحظ الأشياء أما عن طريق تجميع عناصره المتماثلة أو تمييز الأنماط . بصورة عامة وفقا لمبدأ Gestalt ان الانسان يدرك الأشياء من خلال الوصف العام لهذه الاشياء . في هذه المقالة تم الأستفادة من هذين المبدئين لتمييز وأستخلاص جزء العين من الصورة. اللغة البرمجية جافا مع مكتبة opencv المتخصصة في معالجة الصور تم استخدامهما معا لهذا الغرض. نتائج جيدة تم الحصول عليها من هذه الطريقة المقترحة ، حيث تم الحصول على 88.89% كنسبة الدقة أما بالنسبة لمعدل وقت التنفيذ يبلغ 0.23 من الثواني.

الكلمات المفتاحية: استخلاص الكنتور، نظرية كاستألت، مبدأ هيلمهولتز، الكشف عن الجلد.