РАЗРАБОТКА ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ И ИНЖЕНЕРИЯ ЗНАНИЙ

UDC 004.93

Laura Mukhamadiyeva, Aiman Moldagulova

International Information Technology University, Almaty, Kazakhstan

A COMBINED METHOD FOR FACE RECOGNITION

Abstract. Currently, face recognition is one of the most important achievements of biometrics, which allows using algorithms to determine the character of a person by his physiological characteristics. Research and analysis of existing methods for face recognition helps, to recognize the positive and negative aspects of each of the methods provided, on the basis of which an effective method of pattern recognition is created. The fundamental idea of this article is to improve the consolidated face recognition technique using the OpenCV specialized vision library and the Viola-Jones calculation, with the help of which a face from a video stream is correlated with certain 68 points of the face. Additionally, a calculation is considered to naturally alter the image difference in order to productively identify the structure of the face. In addition, in the work, it was possible to improve the quality of object recognition, reduce the degree of false confirmation, reduce the time for preparing the classifier and image processing.

Key words: face recognition, flexible contours, facial asymmetry, color correction.

Introduction

There are a few phases in the advancement of face acknowledgment strategies. The primary stage can be gotten back to the 70s of the only remaining century and identified with the measurable example acknowledgment approach. The following stage is related with the rise of strategies dependent on the Karunen - Loeve change or Principal Component Analysis (PCA). In the issue of face acknowledgment, it is utilized mostly to speak to a face picture with a vector of little measurement (head parts), which is then contrasted and the reference vectors put away in the information base. The fundamental objective of the chief segment examination is to essentially diminish the component of the element space so that it portrays as best as conceivable the "normal" pictures having a place with numerous people. Utilizing this technique, it is conceivable to distinguish different fluctuation in the preparation test of facial pictures and depict this changeability in the premise of a few symmetrical vectors, which are called eigenfaces.

Other significant achievements in face acknowledgment include: the Fisherface strategy, which applies direct discriminant examination (LDA) after the PCA stage to accomplish higher exactness; the utilization of nearby channels, for example, Gabor spouts, to improve facial highlights; and the engineering plan of a falling classifier dependent on learning AdaBoost for constant face discovery.

Face acknowledgment innovation has now progressed essentially since the Eigenface technique was proposed. In restricted circumstances, for example, lighting, act, position, wear and outward appearance can be controlled, programmed face acknowledgment can beat human acknowledgment, particularly when the information base (exhibition) contains countless countenances. Nonetheless, programmed face acknowledgment actually faces numerous difficulties when face pictures are gained in unhindered conditions.

As indicated by specialists [1] a face acknowledgment framework for the most part comprises of four modules as portrayed in figure 1: face identification, face arrangement, highlight extraction, and highlight coordinating.
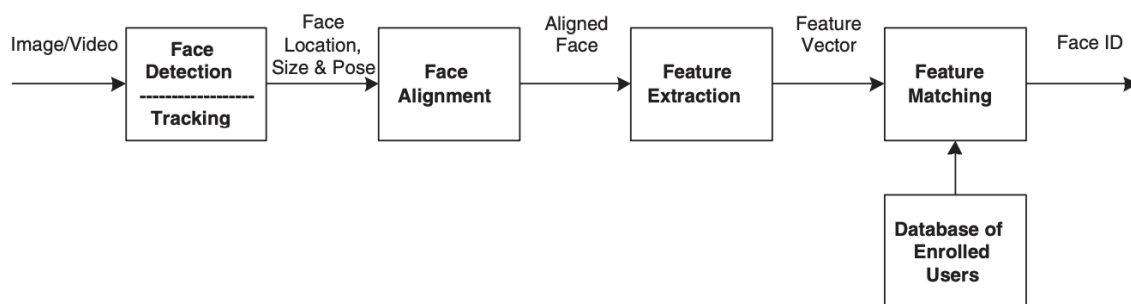
Figure 1 - Face recognition processing flow.

At the primary stage, the face in the picture is distinguished and restricted. At the second stage the face picture is adjusted (calculation and brilliance). At the third stage the highlights are determined and the acknowledgment itself is performed. At the fourth stage the processed highlights are contrasted and the principles put away in the information base. The fundamental distinction between existing calculations is the figuring of highlights and correlation of their assortments with one another.

Concerning programmed face acknowledgment frameworks they are applied in security frameworks with "face control" and observing of unapproved movement, toy robots and self-sufficient space tests, examination of reports and pictures, video information bases, self-learning PC frameworks, PC control utilizing signals, application program interfaces and characterized as one of the most encouraging exploration territories, which opens up new strategies for collaborating with a PC without utilizing traditional "human-PC" association approach. Agreeable and customized methods of speaking with a PC imply that the interfaces of another age ought to distinguish nature encompassing an individual and, at any rate, perceive himself, for example show a PC to distinguish an individual's character. There is a wide assortment of biometric ID techniques: voice, fingerprints, signature, acknowledgment of the retina and iris, human face; also, a considerable lot of them have gotten wide business application and are utilized in viable and business improvements [2]. The current requirements for making frameworks that execute the arrangement of such issues force extreme limitations on the speed of the calculations, which should work in close to continuous mode. To effectively take care of the issue of face acknowledgment, guaranteeing a fast of work ought to likewise be joined with few bogus acknowledgments. In frameworks that execute existing acknowledgment strategies, with an expansion in the degree of acknowledgment, a noteworthy increment in the quantity of bogus choices is watched, which makes their down to earth utilize troublesome [3].

Today, biometric distinguishing proof strategies are of incredible enthusiasm, permitting to decide an individual's character by its physiological qualities by acknowledgment by tests. An exemplary case of biometrics is unique mark examination, and the most recent innovations incorporate the acknowledgment of the retina and iris. An interface like "stop and pronounce yourself" is required for applications with high security necessities. For cutting edge brilliant conditions, face and voice acknowledgment innovations are most appropriate. They are unpretentious (acknowledgment happens a ways off, don't need an extraordinary degree of brightening), and don't limit the client in the opportunity of development [4].

Yet, the most significant thing, obviously, is that individuals ordinarily perceive each other by their countenances and voices, which implies they won't experience any burden with a framework dependent on comparable acknowledgment techniques. An exacerbation of the circumstance with global illegal intimidation has additionally given another driving force to the advancement of human face acknowledgment frameworks. The establishment of such frameworks in jam-packed spots (air terminals, train stations, huge malls) ought to encourage the early identification of needed people [5]. With all the wide range of calculations and picture acknowledgment techniques, an ordinary acknowledgment strategy comprises of three segments:

1. Change of the first picture to the underlying portrayal (may incorporate both pre-handling and numerical changes, for instance, the computation of the fundamental parts).

2. Featuring key attributes of the picture (for instance, the main n head parts or discrete cosine change coefficients are taken).

3. The characterization component: bunch model, metric, neural organization, and so on. Consider the most widely recognized numerical acknowledgment strategies utilized, in view of the mathematical attributes of the face.

Methods

Face acknowledgment is one of the most considered issues in such territories as advanced picture handling, PC vision, biometrics, video conferencing, the formation of smart security and access control frameworks, and so on. The face acknowledgment measure typically comprises of two phases: Search for a face region in the picture, and contrast the discovered face and the countenances in the information base. At present, the Viola – Jones strategy is the most famous technique for finding a face territory in a picture because of its fast and proficiency.

The Viola – Jones face finder depends on the principle thoughts: essential picture portrayal, the strategy for developing a classifier dependent on the versatile boosting calculation (AdaBoost), and the technique for joining classifiers into a course structure. These thoughts make it conceivable to construct a face indicator fit for working continuously. Head part strategy and wavelet change are utilized to get picture qualities. In the issue of face acknowledgment, they are effectively used to think about the segments that portray shading pictures with the parts that depict obscure pictures [6].

The point of this paper is to make another calculation dependent on a blend of the Viola – Jones technique, wavelet change, and head segment strategy (PCM) for face acknowledgment in advanced pictures and video arrangements progressively. Further, the paper portrays and investigates present day techniques for face acknowledgment. The possibility of the strategy is to speak to the pictures of countenances as a set (vector) of the principle parts of the pictures, called "own appearances". They, faces, have a helpful property: that the picture relating to each such vector has a face-like shape. The estimation of the key parts is diminished to the count of the eigenvectors and eigenvalues of the covariance lattice, which is determined from the picture. The entirety of the main segments increased by the relating eigenvectors is the picture reproduction. Table 1.1 gives data about existing techniques for the face acknowledgment.

Table 1 – Analysis of existing methods

Name of the methods	Accuracy	Speed of recognition	Speed of processing data	Sensitivity to lights	Sensitivity to glasses	Sensitivity to beard	Dependability to using techniques
Viola-Jones	92%	15 sh/s	15s	Yes	No	No	Yes
Active Appearance Models (AAM)	88%	5 sh/s	7s	Yes	Yes	No	Yes
Active Shape Models (ASM)	88%	4 sh/s	10s	Yes	Yes	No	Yes

Elastic graph matching	90%	7 sh/s	7s	Yes	Yes	Yes	Yes
Neural Network	94%	5 sh/s	1h	Yes	No	No	Yes
Principal component analysis	93%	6 sh/s	15min	Yes	No	Yes	Yes

For each face picture, its principle segments are determined. The acknowledgment cycle comprises in looking at the fundamental segments of an obscure picture with the parts of every known picture. It is expected that the pictures of faces comparing to one individual are gathered into bunches in their own space. Up-and-comer pictures with the littlest good ways from the info picture are chosen from the information base.

Oneself face strategy requires glorified conditions for its application, for example, uniform enlightenment boundaries, a nonpartisan outward appearance, and the nonappearance of obstruction, for example, glasses and facial hair. In the event that these conditions are not met, head segments won't reflect interclass variety. For instance, under different conditions light, the eigenvalue strategy is for all intents and purposes irrelevant, since the main head segments dominatingly reflect changes in enlightenment, and the correlation yields pictures with a comparative brightening level. On the off chance that the admired conditions are met, the acknowledgment exactness utilizing this strategy can arrive at values over 90%, which is an awesome outcome. Figuring a lot of eigenvectors is extremely work serious. One of the techniques is convolution of pictures in lines and segments - in this structure, the portrayal of the picture has a significant degree littler, estimations and acknowledgment are quicker, however it is not, at this point conceivable to reestablish the first picture [7].

The technique functions admirably and distinguishes facial highlights in any event, when seeing the subject from a slight point, up to around 30. The acknowledgment exactness utilizing this strategy can arrive at values over 90%, which is an excellent outcome. At tilt edges more prominent than 30, the probability of face recognition drops strongly. This element of the technique doesn't permit, in the standard usage, to identify a human face pivoted at a discretionary point, which enormously convolutes or makes it difficult to utilize the calculation in current creation frameworks, considering their developing needs.

Correlation of layout. The premise of this strategy is to choose regions of the face in a picture, and afterward think about these zones for two distinct pictures. Every territory coordinated expands the comparability score of the pictures. The easiest calculations like pixel-by-pixel examination are utilized to think about zones.

The weakness of this strategy is that it requires a great deal of assets both for putting away packages and for contrasting them. Because of the way that the easiest correlation calculation is utilized, the pictures must be shot under carefully settled conditions: no observable changes in point, lighting, enthusiastic articulation, and so on are permitted. The acknowledgment exactness utilizing this technique is about 80%, which is a decent outcome [8].

Strategy constraints: the remembered pictures ought not be fundamentally the same as - the picture ought not be uprooted or pivoted comparative with its unique state. To take out these hindrances, different alterations of the traditional Hopfield neural organization are thought of. The Hopfield network with symmetrical change permits to recuperate profoundly connected examples by changing their unique set to a double arrangement of vectors. Subsequently, a neural organization is gotten that can remember various vectors, and when taken care of to the contribution of any vector, it can figure out which of the retained it is generally comparative.

The acknowledgment exactness utilizing this technique is over 90%, and at times it even methodologies 100%, which is a practically superb outcome. For most current programmed face acknowledgment frameworks, the principle task is to contrast a given face picture and a lot of face pictures from an information base. The attributes of programmed face acknowledgment frameworks for this situation are surveyed by deciding the probabilities of incorrect refusal in acknowledgment (blunders of the principal kind) and mistaken acknowledgment (blunders of the subsequent kind). Notwithstanding mistake probabilities, an evaluation of flexibility to picture aggravations brought about by mix with complex foundations, inconstancy of lighting, evolving haircuts, and so forth is regularly used to assess the programmed face acknowledgment framework. Considering the abovementioned, it appears to be that it is promising to make cross breed strategies that utilization the preferences and leveling the weaknesses of the different specific methodologies talked about above [9].

The affirmation or request measure includes in building up a particular number of head sections for the information picture. Dependent upon their number, the affectability of the technique to disturbance and little differences in faces increases or reduces. Starting there forward, the decided "own appearances" are differentiated and the image portions of the planning set. As shown by a particular estimation, the great ways from the data picture to the model picture is looked and the best result is taken. For handling the code Visual Studio Code was utilizing as an assemblage place. Microsoft's VS Code is a lightweight, easy to understand code proofreader that is accessible on all stages and fantastically adaptable. This is an incredible decision for Python programming. The figure 2 shows the plan of VS code.
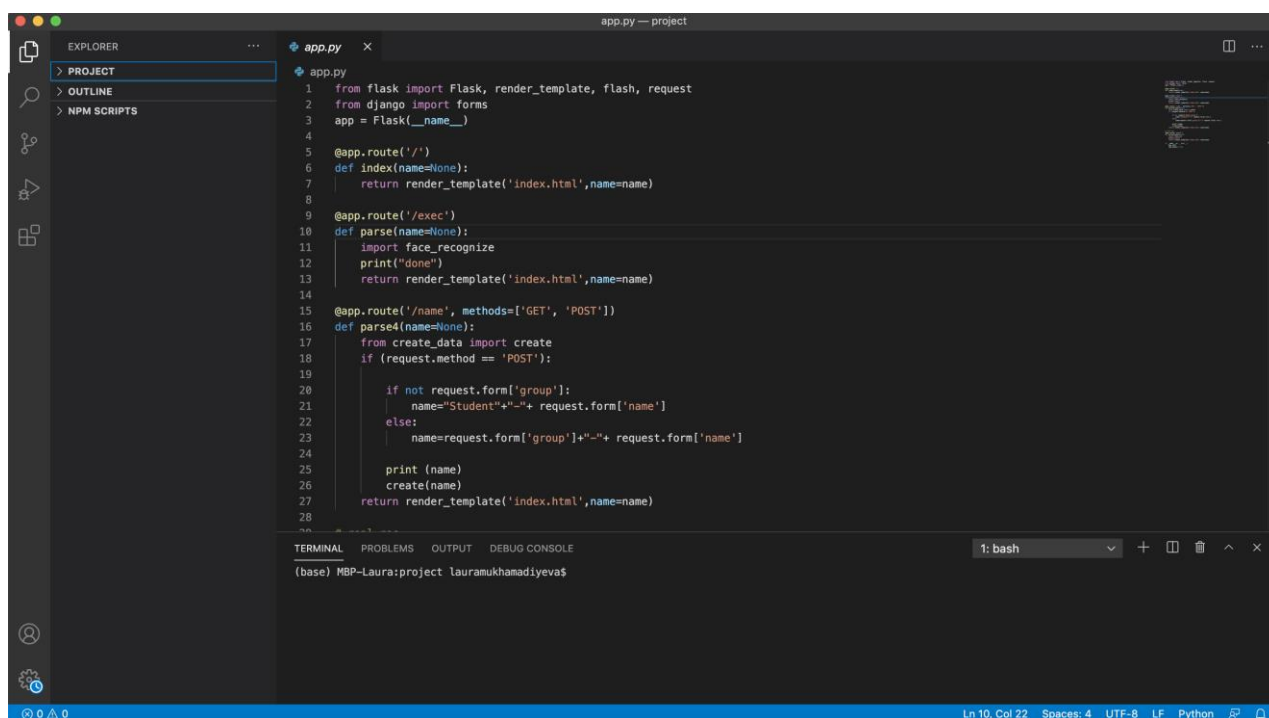

Figure 2 – Python in VS code

The Python programming language is a fundamental resource for making programs for a wide grouping of purposes, accessible regardless, for beginners. It will in general be used to deal with issues of various types. Right off the bat, we can include new individual, with his name and we can order him/her as an instructor or understudy. To start with, following strategies can be utilized here that utilization the movement history to foresee contrasts in the following casing. Furthermore, the correspondence issue can likewise be considered as the errand of evaluating the obvious develop-

ment in the picture (optical stream). The most significant aspect of the exploration are eyes of the individual. These days, the acknowledgment Asian individuals is very troublesome reason for the face shape and state of the eyes. The thesis paper considers such sort of things in more close manner and give the arrangement. From that point forward, the program prepared to remember you in area recognize me genuine. The primary contrast between capacities distinguish me and recognize me genuine is utilizing various techniques as an examination which one is successful [10].

The face acknowledgment issue is described by various elements, to be specific: high facial fluctuation because of the anatomical highlights of individuals; various degrees of enlightenment of articles, contingent upon the sort, amount and directional qualities of light sources; the need to recognize people with various spatial positions. In this progression framework can remember you by utilizing consolidated strategies.

This part concerns principle programing execution and considering the calculations how it functions. Each phase of the advancement bit by bit shows and thinks about all parts of the coding in more close manner and spotlight on the cycle of the turn of events. The primary capacities on the fundamental page initiates relies upon the calculation and participation among HTML and CSS. The figure 2 how the information added into the program to additional outcomes.

```python
@app.route('/name', methods=['GET', 'POST'])
def parse4(name=None):
    from create_data import create
    if (request.method == 'POST'):

        if not request.form['group']:
            name="Student"+"-"+ request.form['name']
        else:
            name=request.form['group']+"-"+ request.form['name']

        print (name)
        create(name)
    return render_template('index.html',name=name)
```

Figure 2 – The adding new person

The Python programming language is an incredible asset for making programs for a wide assortment of purposes, open in any event, for tenderfoots. It very well may be utilized to tackle issues of different sorts. Therefore, we get a rundown, toward the start of which are the most like the picture of the picture. The indexed lists for 200 photographs and 1500 pictures from the information base are as per the following: in 70% of cases, the picture comparing to the photographs is placed first in the rundown; in 95% of cases, the picture related with the photographs falls into the best twenty of the rundown.

The primary thought is Haar Cascade models to confront recognizable proof. The accompanying data shows how work the course framework. The introduction of each course, frontal face course distinguishes the essential state of face, eye course uses to effortlessly perceive the eyes, does nor make a difference wearing glasses or not. The accompanying screens of the code depict open or close the current substation of the eyes. The figure 3 gives data about the expectation of the eye status, it implies it is concerning open the eyes or not. This progression thoroughly shows is the individual is genuine or not.

```python
def predict(img, model):
    img = Image.fromarray(img, 'RGB').convert('L')
    img = imresize(img, (IMG_SIZE,IMG_SIZE)).astype('float32')
    img /= 255
    img = img.reshape(1,IMG_SIZE,IMG_SIZE,1)
    prediction = model.predict(img)
    if prediction < 0.1:
        prediction = 'closed'
    elif prediction > 0.9:
        prediction = 'open'
    else:
        prediction = 'idk'
    return prediction
```

Figure 3 – The prediction of the eyes status

The purpose of this work is to see faces with quick. To deal with the face affirmation issue, a computation is proposed reliant on the utilization of the Viola – Jones procedure, wavelet change, and the focal part method.

1. Suggestion portray another figuring reliant on the joint usage of the Viola – Jones method, wavelet change, and the central part procedure for seeing countenances in pictures and video progressions continuously.

2. A program was developed that realizes the proposed figuring for seeing countenances in pictures and video progressions dynamically in the article organized programming language Python.

3. An examination of the outcomes of tests licenses us to talk about the successful action of the made count and program when seeing countenances. Viola-Jones strategy. This strategy is exceptionally proficient for looking for objects in pictures and video groupings continuously. This finder has a very low likelihood of bogus face identification.

Results

The result of our work should be an automated complex for Express analysis of binary images, which implements a mathematical model for describing the spatial structure of the displayed fields with invariant statistical characteristics of brightness differences. As a result, the fundamental component of the work is a detailed study of the global structural area of noctilucent cloud as a marker of changes in the polar vortex caused by the climatic trends of our planet at high latitudes and the interaction of clouds with the structural elements of the earth's topography.

The aftereffect of our work ought to be a robotized complex for Express investigation of twofold pictures, which executes a numerical model for portraying the spatial structure of the showed fields with invariant factual attributes of splendor contrasts. Therefore, the principal part of the work is a nitty gritty investigation of the worldwide auxiliary region of noctilucent cloud as a marker of changes in the polar vortex brought about by the climatic patterns of our planet at high scopes and the communication of mists with the basic components of the world's geology.

For the affirmation of appearances in pictures and video progressions continuously in the language of article arranged programming Python using the Viola-Jones and Wavelet change strategies, writing computer programs was made.

The route toward securing the signs of eminent individuals is according to the accompanying: Convert video layout picture to grayscale. Application to the grayscale image of the Viola – Jones strategy for glancing through the face zone. Reduction the size of the face zone to 64×64 pixels. Application to the image of wavelet change got in a state of harmony 3 to remove facial features (wavelet coefficients). Saving the isolated characteristics in the information base. During the time

spent seeing a dark face steps are done, by then, considering the utilization of the focal part methodology, the amount of features is diminished and differentiated and the features set aside in the information base.

The purpose of the assessment is to filter for pictures of faces that contrast with demonstrated test plans. Mathematical tests were performed dependent on the base of facial pictures. The information base contains 2000 pictures of appearances of changed people, 40 photos of each face. Exactly when the information base was molded, the image size and shooting conditions were the equal.

At the arrangement stage, each individual is before the camera, performing translational and rotational advancements of his head. The program records 40 unmistakable photos of an individual and stores it in an information base with the name of that person. The path toward sparing the signs of well known individuals occurs as follows: Converting the image of a video diagram into a grayscale picture.

Saving the eliminated characteristics in the information base. During the time spent seeing a dark face, all methods are taken, by then, considering the utilization of the fundamental part strategy, the amount of features is diminished and differentiated and the features set aside in the information base. To check the correct action of the program, this methodology is reiterated, and the program dynamically makes sense of who is starting at now before the camcorder.

Test conditions for the facial acknowledgment programming bundle: An information base was made from the video transfer (faces with specks were removed) in 88 individuals. A test video test transfer has been made, comprising of 88 of similar people as in the information base, 48 people remembered for the rundown of 88 individuals, yet on one more day, and 40 people of outsiders excluded from the rundown of perceived individuals. Test consequences of the facial acknowledgment programming bundle:

1. Each of the 88 individuals are perceived from the information stream.
2. The program snapped the photo of the obscure individual.
3. The acknowledgment time per object was 480 milliseconds.

Conclusion

The proposed new calculation depends on the joint use of the Viola – Jones strategy, wavelet change for face acknowledgment in pictures and video groupings progressively. A program has been built up that executes the proposed calculation for face acknowledgment in pictures and video arrangements progressively in Python. Examination of the aftereffects of PC tests permits us to talk about the effective activity of the made calculation and program in face acknowledgment. This article gives a depiction and consequences of the created face acknowledgment program, which has the accompanying highlights:

1. A high level of distinguishing proof.
2. The absence of impact of components, for example, concealing the upper aspect of the face (haircut, headgear), eyes (glasses, except for sun glasses) and head tilt.
3. The ideal opportunity for perceiving an individual is (in the pre-owned programming bundle) 480 milliseconds for a base of 88 individuals.
4. To decide the focuses in the Viola – Jones calculation, the shading qualities of the face are not exceptionally basic.
5. The made programming bundle executes amendment of picture splendor and differentiation to build acknowledgment exactness.

Further regions of examination ought to be coordinated to:
- improved acknowledgment, considering the pivot and tilt of the head;
- compensation for the impacts of brilliance and difference;
- determination of the ideal number of key face boundaries.

Among all idea about affirmation procedures, the most abstract and promising are model strategies. Exhibiting license you to totally reflect reality and sensibly requires high resource costs and computational execution. Taking into account the material considered in this article, the going with closures can be drawn:

1) The Viola-Jones estimation is sensational for use dynamically structures due to the simplicity of the thing acknowledgment movement;

2) The idea of article acknowledgment vehemently depends upon the system and nature of classifier planning; in like way, for the correct portrayal it is basic to make the correct classifier.

REFERENCES

1. Jain A.K., Li S.Z. Handbook of face recognition. New York: springer; 2011.
2. Boll R. M. Guide to biometers. Face recognition methods by using main guide to biometers. M.: Tekhnosfera, 2007. – p. 385.
3. Sharma R. S., Kumar D.K., Puranik V.D. Performance Analysis of Human Face Recognition Techniques// Access mode: https://ieeexplore.ieee.org/abstract/document/8777610/, [29.07.2019]
4. Kolomiets V. T. The competative analysis of the face recognition// Access mode: http://habrahabr.ru/company/synesis/blog/238129/, free access [05.09.2016].
5. Pentland A., Chaudhary T. Face recognition for intelligent environments// Access mode: http://http://www.osp.ru/os/2000/03/177939/ free access [09.26.2016].
6. Adrian R.S., Anto S.N., Maulahikmah G.M. Implementation of face recognition algorithm for biometrics based time attendance system//Access mode: https://ieeexplore.ieee.org/document/7013165, [19.01.2015]
7. Neel R. M., Sonia K. S. Real-time implementation of face recognition system//Access:https://ieeexplore.ieee.org/abstract/document/8282685, [08.02.2018].
8. Chenggang Z. C., Yingmei S.Y. Research about human face recognition technology/ Access mode https://ieeexplore.ieee.org/document/5412901 [01.02.2010].
9. Naveen K.G., Er. N. K. A comparative analysis on various face recognition techniques//Access:https://ieeexplore.ieee.org/document/8250626, [11.01.2018].
10. Suad H. S., Asaf V.A., Real time face recognition system (RTFRS)//Access mode https://ieeexplore.ieee.org/document/7473527 [09.26.2016].

Лаура Мухамадиева, Айман Молдагулова
Тұлғаны танудың біріккен әдісі

Аңдатпа. Қазіргі кезде тұлғаны тану биометрияның маңызды жетістіктерінің бірі болып табылады, ол адамның физиологиялық сипаттамалары бойынша алгоритмдерді қолдануға мүмкіндік береді. Бетті танудың қолданыстағы әдістерін зерттеу және талдау ұсынылған әдістердің әрқайсысының жағымды және жағымсыз жақтарын тануға көмектеседі, соның негізінде үлгіні танудың тиімді әдісі жасалады. Осы мақаланың негізгі идеясы – OpenCV мамандандырылған көру кітапханасын және Виола-Джонсты есептеуді қолдана отырып, тұлғаны танудың шоғырландырылған техникасын жетілдіру, оның көмегімен бейне ағынының бет әлпетінің кейбір 68 нүктелерімен байланыста болады. Сонымен қатар бет құрылымын өнімді анықтау үшін кескін айырмашылығын табиғи түрде өзгерту үшін есептеулер қарастырылады. Аталған жұмыста объектіні тану сапасын жақсартуға, жалған растау дәрежесін төмендетуге, жіктеуішті дайындауға және кескінді өңдеуге уақытты қысқартуға мүмкіндік туды.

Түйінді сөздер: тұлғаны тану, икемді контурлар, тұлғалық асимметрия, түсті түзету

Лаура Мухамадиева, Айман Молдагулова

Комбинированный метод распознавания лица

Аннотация. В настоящее время распознавание лица - одно из важнейших достижений биометрии, которое позволяет с помощью алгоритмов определять характер человека по его физиологическим характеристикам. Исследование и анализ существующих методов для распознавания лиц помогает распознать положительные и отрицательные аспекты каждого из предоставленных методов, на основе чего создается эффективный метод распознавания образов. Фундаментальной идеей данной статьи является улучшение консолидированной техники распознавания лиц с использованием библиотеки специализированного зрения OpenCV и вычисления Виолы-Джонса, с помощью которого лицо из видео-потока соотносится с определенными 68 точками лица. Дополнительно рассматривается расчет для естественного изменения различия изображения, чтобы продуктивно идентифицировать структуру лица. Кроме того, в работе удалось повысить качество распознавания объектов, снизить степень ложного подтверждения, сократить время подготовки классификатора и обработки изображения.

Ключевые слова: распознавание лиц, гибкие контуры, асимметрия лица, коррекция цвета.

Сведения об авторах:

Лаура Мухамадиева, PhD студент, лектор кафедры компьютерной инженерии и телекоммуникаций Международного университета информационных технологий

Молдагулова Айман Николаевна, к.ф.-м.н., и.о. профессора кафедры информационных систем.

About authors:

Laura Mukhamadiyeva, PhD student, lector, Computer Engineering and Telecommunication Department, International Information Technology University

Aiman N. Moldagulova, cand. of physics and math. sci., acting professor, Information Systems Department, International Information Technology University.

Manabayev B.K.

International Information Technology University, Almaty, Kazakhstan

AN APPLICATION OF EXPLORATORY DATA ANALYSIS FOR EVALUATING FOOTBALL TEAM PERFORMANCE

Abstract. The following article provides information on the usefulness of exploratory data analysis in identifying similar patterns and anomalies of how a football team performs throughout the whole season. This paper presents in-depth look on popular, controversial hypothesis while also describing a hidden shift in play-style of certain football squads, and what results have those teams achieved in accordance with data gathered by sports analytics web-resource.

Key words: exploratory data analysis, football team performance, expected goals metric, football statistics

Introduction

Machine learning [1] is the science of computer algorithms which provide the capability to learn information from the given datasets without considering predetermined equations. Nowadays