Chapter 4

Diagnostic Decision Support System in Dysmorphology

Kaya Kuru and Yusuf Tunca

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/51118

1. Introduction

Dysmorphology is the aspect of clinical genetics concerned with syndrome diagnosis in patients who have a combination of congenital malformations and unusual facial features, often with delayed motor and cognitive development [1]. Making a diagnosis for a dysmorphic patient requires a high degree of experience and expertise [2] since many dysmorphic diseases are very rare. The human brain possesses a remarkable ability to recognize what is familiar or unfamiliar in a face. The general public can recognize the face shape of individuals with Down's syndrome and experienced clinical geneticists develop this innate skill to such a degree that they recognize subtle facial features associated with several hundred dysmorphic syndromes [3]. In some areas of the world, however, genetic diagnosis is generally performed by general practitioners, dermatologists or pediatricians, not particularly trained in dysmorphology, rather than trained geneticists due to the lack in the numbers of geneticists. Dysmorphologic diagnosis is usually performed by referring the images or terms standardized and specified in some limited number of catalogs and databases. However, not being very familiar with the terminology, especially for practitioners in rural areas, is an important handicap to dig into the right diagnosis. This can lead to diagnostic inaccuracy which in turn curtails both the right cure of patients that will best suit their particular needs and the right guiding of their parents who may be at risk of having a new dysmorphic child. The whole process of reaching a genetic diagnosis can be very lengthy and entail struggling with referrals to numerous medical professionals and waiting for appointments. Throughout this process, parents find themselves dragged from one medical professional to another until discovering the accurate diagnosis of their child. These parents often experience high levels of burdensome anxiety and frustration. Moreover, delay in diagnosis may also delay access to critical services such as clinical trials and a patient's referral to supportive services including early intervention, physical or occupational therapy. Syndrome recognition and diagnosis is of clinical importance for several reasons according to Smithson

[1]: first, it influences patient management because awareness of the pattern of anomalies associated with a particular syndrome highlights the investigations that need to be undertaken; second, diagnosis provides information about the long-term prognosis and may help to identify options for treatment: for example, bone marrow transplantation or enzyme replacement therapy can now be offered for some inborn errors of metabolism (e.g. in Fabry disease); third, the diagnosis determines what genetic advice can be given, including an estimation of genetic risks and possible means of prenatal diagnosis. Thus, delays in early treatment can have a miserable impact on the patient's health and can dramatically influence the chances of the child catching up to his/her peers and leading to a normal life. Therefore, reaching a thorough genetic diagnosis at an early stage is crucial. Not only does this affect the choice of cure for the patient, but it also enables the proper guidance, counseling and support for the parents to ensure an overall improvement in the patient's quality of life and care. Most physicians, neurologists or pediatricians are capable of noticing these early signs in a child, but are not equipped to perform a precise genetic diagnosis on their own especially for very rare diseases. Since there are thousands of possible genetic conditions to be taken into account and each condition is in itself very rare, a specialist's evaluation generally seems the best path to ensure a proper genetic diagnosis is reached though it is sometimes very difficult even for geneticists to diagnose these orphan diseases.

The face is widely recognized as an attribute which best distinguishes a person, even at first glance. Facial features give lots of clues about the identity, gender, age and ethnicity. The face develops under the influence of many genes and in many cases a face provides important information to diagnose a syndrome. Thus, facial appearance can be a significant hint in the initial identification of genetic anomalies generally associated with cognitive impairments. There are specific properties, especially for facial dysmorphology caused by genetic syndromes and these properties are used by geneticists to pre-diagnose even before a clinical examination and genotyping are undertaken. Analyzing of properties in faces is sometimes sufficient to diagnose for some cases, however, it is necessary to analyze other specific properties of the body such as the structure of the skeleton and the characteristics of speech produced for some other cases. Diagnosing of genotype-phenotype correlations correctly among many syndromes seems beyond the capability of human especially for very rare diseases. Most of the genetic diagnostic decision support systems (DDSS) have been applied in terms of the anthropometry and more specifically craniofacial anthropometry including stereophotogrammetry[1] so far. It might be possible to diagnose a good number of syndromes correctly by using computer-assisted face analysis DSS as asserted by some scientists such as Farkas [4], Loos [5], Boehringer [2], and Hammond [6, 7]. Farkas [4] pioneered techniques for studying facial morphology using direct anthropometry for nearly 40 years ago [4]. His approach, using a ruler, calipers, tape measure and protractor, has been applied widely in the analysis of facial dysmorphology [7]. Many clinicians undertake such a manual craniofacial assessment and compare a patient's phenotype to the norms of a control

1 Anthropometry is defined as the biological science of measuring the size, weight, and proportions of the human body [4]. Craniofacial anthropometry is performed on the basis of measures taken between landmarks defined on surface features of the head, face, and ears [12]. Stereophotogrammetry refers to combining multiple views of photos to form a 3D image [13].

http://dx.doi.org/10.5772/51118

population of comparable age and sex [7]. A most recent study of Hammond [7] has reviewed the surface-based image capture in 3D in syndrome delineation and discrimination, in the categorization of individual facial dysmorphology and in phenotype–genotype studies which is a very complex implementation of dysmorphological diagnosis process in a clinical environment. For dysmorphic syndromes with known genetic causes, molecular and/or cytogenetic analysis is the appropriate route of investigation in order to confirm a diagnosis. However, applying right analysis method throughout many probable analyses is very much dependent to the accurate diagnosis considered before genotyping is undertaken. The aim of this chapter is to propose a new methodology by observing the characteristic key components of facial dysmorphology associated with genetic disorder to indicate the right diagnostic criteria.

Douglas [8] has stated that classification of faces based on facial patterns in isolation is unlikely to be accepted by dysmorphologists unless the mathematical features extracted and identified by feature selection algorithms to be discriminatory can be related to facial appearance. Principal component analysis (PCA), independent component analysis (ICA), kernel principal component analysis (KPCA), local feature analysis (LFA), probability density estimation (PDA), multi-linear analysis (MLA), elastic graph matching (EGM), kernel discriminant analysis (KDA), support vector machine (SVM), Gabor wavelet and Fischer's linear discriminant analysis (LDA) exist to analyze the features of a face. Among them, the methods of PCA using eigenface, elastic graph matching, Gabor wavelet and Fischer's LDA are popular for face recognition. PCA can be computed as an optimal compression scheme that minimizes the mean squared error between an image and its reconstruction [9] as well as it may achieve good results of up to 96% recognition under ideal conditions [10] which can be provided with either capturing images in ideal environments by having a good illumination and using a good capturing device or preprocessing of images with several image processing techniques before extracting features. PCA using eigenfaces have been identified as computational efficient on even reduced hardware [11]. Additionally it is used for dimension reduction as in our study from 2D to 1D to ease and speed up the calculations. That is to say, the dimensionality of the required space for all images can be reduced to the number of input images instead of the pixel count that all images have in total.

City block distance, Euclidean distance, sub/space method, multiple similarity method, Bayes decision method and Mahalanobis distance are known typical distance functions [9]. Kapoor [9] highlights that the Mahalanobis distance is the most effective of the typical evaluation distance-based approaches that calculate the distance from a point to a particular point in the data set. We have tested the Mahalanobis distance and Euclidean distance in our study to find the genotype-phenotype correlations. The well/known Mahalanobis distance classifier[2] is based on the assumption that the underlying probability distributions are Gaussian [9]. Kapoor [9] indicates the difference such that Mahalanobis distance is a distance measure based on correlations between variables by which different patterns can be identified and analyzed. It is a useful way of determining similarity of an unknown sample

2 Interested reader may found detailed information about Mahalanobis distance from Gul's thesis [14].

set to a known one. It differs from Euclidean distance in that it takes into account the correlations of the data set and is scale-invariant, i.e. not dependent on the scale of measurements. In our study, the well-known Euclidean distance matching process outperformed the process of Mahalanobis distance. Thus, we chose this matching technique in our study.

Hammond [7] pointed out that in terms of future technological support, two (2D) or three-dimensional (3D) models of facial morphology are showing potential in syndrome delineation and discrimination, in analyzing individual dysmorphology, and in contributing to multi-disciplinary and multi-species studies of genotype–phenotype correlations. Our study is an example of substantiating this potential. We have developed a real-time computer system that can locate and track a patient's head, and then recognize the patient by comparing characteristics of the face to those of trained individuals with classified dysmorphic diseases. In this study, in terms of both the feature extraction and helping non experienced practitioners in diagnosis process as well as to support experts in their decisions, we established an application to ease the process and we refer to our method as "Facial Genotype-Phenotype Diagnostic Decision Support System (FaceGP DDSS) in Dysmorphology". Up to date, no complete solution has been proposed that allow the automatic diagnosis of dysmorphic diseases from the raw data (live camera, video or frontal photographs) without human intervention. The FaceGP DDSS aims not only to ease the required on-site expertise, but also to eliminate the time consuming catalog search of practitioners and geneticists to diagnose facial dysmorphic diseases through approximately 4.700 known dysmorphic diseases[3] automatically, no intervention from the user such as preprocessing of images. The FaceGP DDSS methodology can be implemented on any site easily. In the methodology, reference images or reference patients on live subjects having the specific dysmorphic diseases are used as a guide for identifying the facial phenotypes (the outward physical manifestation of the genotypes) to train the system. Digital facial image processing methods are employed to reveal facial features with disorders indicating dysmorphic genotype-phenotype interrelation. A great number of genetic disorders indicating a characteristic pattern of facial anomalies can be classified by analyzing specific features (eigenfaces) with the aid of facial image processing methods such as PCA. Distance algorithms such as Euclidean, Mahalanobis are used to construct the correlation of the input image with the trained images in matching. Some image enhancement methods such as histogram equalization and median filter are implemented on detected images to capture better features and compensate for lighting differences. This study proposes a novel and robust composite computer-assisted and cost-effective method by merging several methods in the characterization of the facial dysmorphic phenotype associated with genotype, in particular a method relying primary on face image capture (acquisition from either camera, video or frontal face images) and manipulation to help medical professionals to diagnose syndromes efficiently.

3 Many new dysmorphic diseases are described each year. London Dysmorphology Database (http://www.lmdatabases.com).

http://dx.doi.org/10.5772/51118

2. Methods

The FaceGP DDSS methodology has been established in C++ programming language. We benefited OpenCV[4] library. The application can function on any present computer, not requiring much CPU and memory while processing thanks to the easy implementation of PCA. The methodology comprises several main modules, namely *face detection and image acquisition, image processing, training and diagnosis/recognition module,* and these main modules are divided into several sub modules as illustrated in Figure1. Functions of these modules are explained in following subsections in detail.

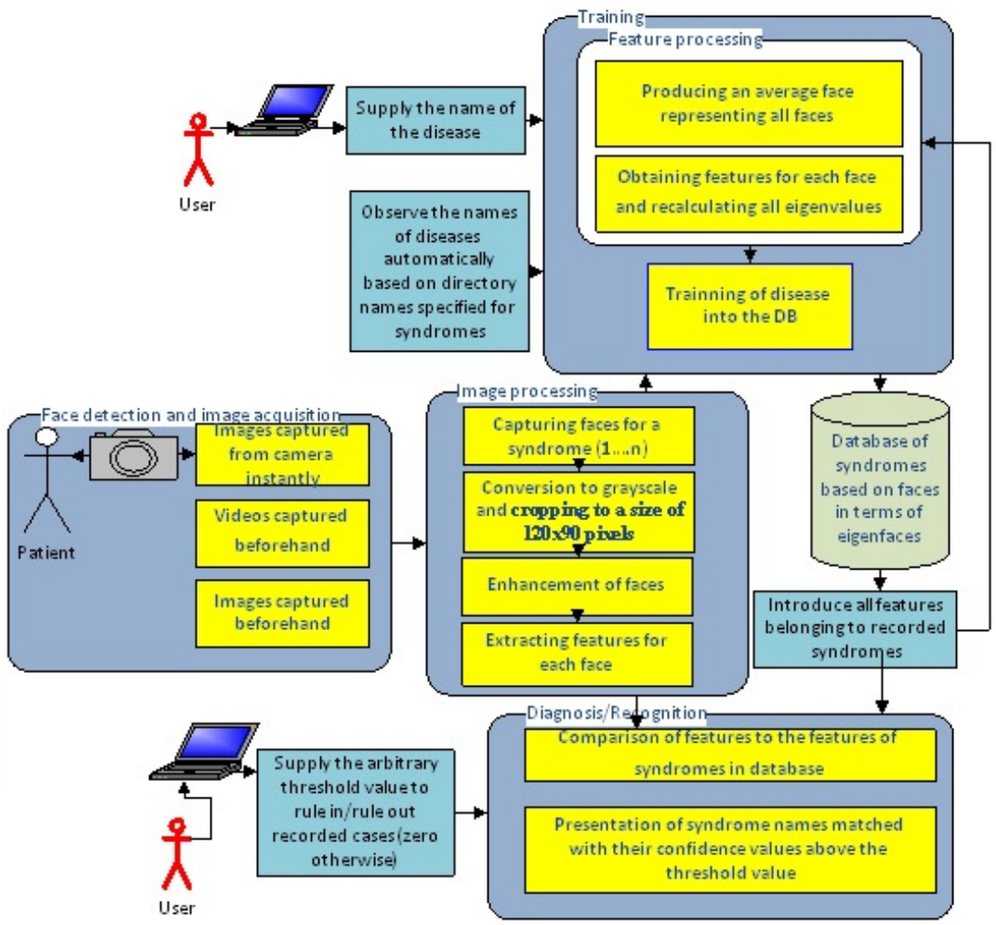

Figure 1. 1. Overall architecture of the methodology: the system consists of four main modules; face detection and image acquisition, image processing, training and diagnosis/recognition module. These modules are divided into several sub modules that are delineated in the specified sections of the modules.

4 OpenCV (Open Source Computer Vision) library can be reached from the site, http://opencv.willowgarage.com/wiki/.

2.1. Face Detection and Image acquisition

The patient images can be captured from several different environments: they can be captured from softcopy images stored in a disk, it can be observed from hardcopy pictures or real-time images which can be detected simultaneously from patients by a camera attached to the computer. The images can be automatically captured on an ordinary high-resolution digital camera (e.g. 8 megapixels) mounted across the patient (no matter what the distance is), provided that a frontal face is detected (to prevent motion artifacts). Natural movements of the body at rest (e.g., breathing), and a person's inability to sit motionless result in no problem as the camera detects and acquire images in a short period of time. Time of capture per image is 0.2s which is triggered by the system automatically excluding time for not acquiring frontal faces. Capturing images is so instantaneous that it may be more suitable for imaging children with mental retardation who are unable to hold a pose for long and potentially uncooperative. Images not including a proper frontal face are not captured and the application is idle during this time, especially when the head turning to sides, up and down. Frontal faces are the essential components of our data preparation and model building. In other words, the process of capturing images is just performed if the application detects an acceptable frontal face. The application is able to capture sufficient number of frontal images in real-time, which is required to train the system for subsequent analyses.

2.2. Image Processing

2.2.1. Capturing faces for a specific case

Images of patients are acquired for a specific syndrome either to train or to diagnose in terms of the function triggered by the user. The functions of the implementation are explained in the next section. The system activates the diagnosis function as it detects the images of patients if this function is triggered. Otherwise, it expects user interference to activate the training function in order to make sure that the procedure of capturing all required frontal face images for a specific syndrome is finished.

2.2.2. Conversion to grayscale and cropping of images

After images of patient are acquired, they are converted to grayscale and cropped to include just faces by relying upon the face outline, forehead, two eyes, cheek, chin and mouth. This stage is finalized by normalizing the image with a size of 120x90 pixels. Processing of this phase as depicted in Figure 2 ensures that all face images are exactly positioned the same regarding a proper rotation. A background removal algorithm is not implemented due to the cropped face in which there is almost no background and the rest of the image is removed before feature extraction.

http://dx.doi.org/10.5772/51118

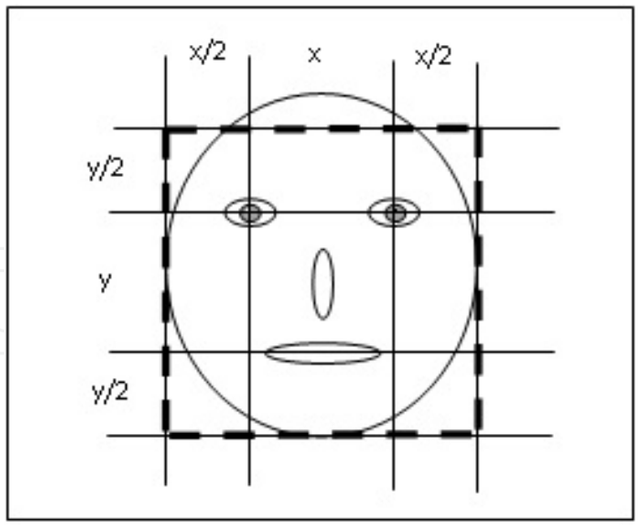

Figure 2. Delineation of cropping a dysmorphic face.

2.2.3. Enhancement of face images

The face images are standardized by employing two essential image enhancement methods, namely histogram equalization and median filtering to remove illumination variations in order to capture better features. Histogram equalization employed on an image is known to provide better extraction of features [14]. Histogram equalization is applied on too dark or too bright face images in order to enhance image quality and to improve diagnostic performance, thus, facial features become more apparent by enhancing the contrast range of the image. Median filtering is very effective to enhance images especially obtained from a camera without losing information [15]. There is no preprocessing of frontal images manually.

2.2.4. Feature extraction

Extraction of facial features is implemented in this sub module. After automatic image preprocessing is employed on raw image data, the feature extraction is implemented on the normalized face image to reveal the key features that are used to classify and recognize dysmorphic diseases. The FaceGP DDSS methodology is designed for modeling and analyzing large sets of face images. Whichever method is used, the most important problem in face recognition is reducing the dimensionality in terms of easing computational complexity. We haven't used craniofacial landmarks or extracting feature graphs[5]. Instead; we performed principal component analysis (PCA) which is a very effective method at classifying face images in the sense of reducing computational complexity. PCA which evaluates entire face is

5 Extracting feature graphs is based on interpolation of the basic parts of a face such as eyes, nose, mouth, and chin. In the method, with the help of deformable templates and extensive mathematics, key information from the basic parts of a face is gathered and then converted into a feature vector. L. Yullie and S. Cohen played a great role in adapting deformable templates to contour extraction of face images[16].

implemented to extract the features of dysmorphic faces in terms of its simplicity, learning capability and speed. PCA may be defined as the eigenvectors of the covariance matrix of the set of face images, meaning an image as a vector in a high dimensional space. It classifies the face images by projecting them to the 2-D space which is composed of eigenvectors obtained by the variance of the face images. These eigenvectors are a set of features that characterize the variation between face images. In other words, the features of the images are obtained by looking for the maximum deviation of each image from the mean image. Each image location contributes more or less to each eigenvector, so that it is possible to display these eigenvectors as a sort of ghostly face image which is called an "eigenface". Eigenfaces can be viewed as a sort of map of the variations between faces. PCA reduces the dimensionality of the dataset. Thus a face of 8-bit intensity values can be represented by an ordered sequences numbers in a vector, one dimensionally with PCA. This is a huge data compaction, as in our case reducing the representation of a face surface from 10800 dimensionality space (120x90 2D points each with x and y coordinates) down to 1 per image. Each face can be regenerated from eigenvalues using a linear weighted sum of the PCA modes in return. The most relevant information contained in a dysmorphic face image is extracted by PCA[6]. Briefly, after capturing sample dysmorphic images and normalizing them following some automatic image processing methods, eigenfaces are generated and saved. An eigenface keep values that make a dysmorphic image unique.

2.3. Training

The eigenfaces, eigenvalues and average image generated by PCA are stored in XML file as Haar-like features[7] together with their labeled diagnosed names. The methodology has the ability to store and read data in XML format. The number of generated eigenfaces (n-1) is almost equal to the number of input images (n) in our study. Classifiers are trained with those features extracted by feature extraction module. As new dysmorphic diseases are trained, the eigenfaces and eigenvalues are recalculated.

Users can easily add new dysmorphic diseases by using their archives in which there are several sample images representing other diseases not defined in the system one by one as well as more than one disease once by the help of the utility provided by the system.

2.4. Diagnosing/Recognition

The trained classifiers are employed for prediction in this sub module. Diagnosing/recognition process is a pattern recognition task. The prediction of the diagnosis of a patient requires the detection of frontal face from a camera or a file, normalizing and processing of face with techniques mentioned in Section 2.2 and extraction of facial features for comparing the trained classifiers. A dysmorphic face image captured and processed is transformed into

6 Details, especially formulas about PCA can be found in Akalin's study [15] and Calvo's article [17].

7 Haar-like features encode the existence of oriented contrasts between regions in the image. A set of these features can be used to encode the contrasts exhibited by a human face and their spacial relationships. Haar-like are computed similar to the coefficients in Haar wavelet transforms. Interested reader can refer to Viola's study for more information about Haar-like features [18].

http://dx.doi.org/10.5772/51118

its eigenface components and then these components are compared to those of predefined labeled classes. There are a number of algorithms in the literature that can compare faces to look for a match. A simple and intuitively appealing way to compare an individual face with two sets of faces is to calculate how close for which a method is a nearest-neighbor classification, how close are they in terms of the Euclidean distance. Once a face has been detected and extracted from an image it is ready to be compared against known/trained syndromes to see if there is a similarity. The captured face image compared to all the syndromes trained in the database one by one for similarities to make sure all similar faces above the threshold value are found in terms of the Euclidean distance[8]. The role of Euclidean distance comparison for image recognition is not much different in that it tries to capture how similar or different a test object is from training objects in terms of the 50 or so mode values, that face surface is to the average face surfaces of each set. Whichever of the average faces is closest determines the classification of the individual. In other words, during the recognition stage, when a new image is input to the system, the mean image is subtracted and the result is projected into the face space. The best match is found for the identity that minimizes the Euclidean distance. For example, for a particular syndrome, the features obtained from the PCA is compared to the trained faces in the database using an Euclidean distance comparison to calculate how close it is to the features of faces in the Database. The similar faces above the threshold value supplied by the user are probable diagnoses displayed to the user. Confidence values can be defined as the resemblance or degree of proximity of eigenface values (how near they are) between two sets of eigenface values obtained from the values of the trained diagnostic images and the identified image to be diagnosed. These values are used for assessing the reliability of the proposed diagnostic inference by the system.

Diagnosis is where the system compares the given patient's face features to all the other trained face features in the database and gives a ranked list of possible matches with respect to the confidence values above the threshold value. In our system, probabilities of similarities in diagnostic process are revealed to the user in a decreasing order rather than either "known" or "unknown" outputs as in face recognition systems, after comparing an undefined dysmorphic face with recorded defined dysmorphic faces that are diagnostically classified. That is to say, our system is not a face recognition system; rather, it is how much a disease is similar to the diseases classified in the database as ruling in or ruling out diseases in terms of the threshold value supplied by the user. This cost effective diagnosis methodology could then help to determine subsequent investigations including more appropriate genetic testing, and possibly even avoiding or delaying the need to undertake some of the more expensive genetic tests.

2.5. Interface and the Functions of the Methodology

A screen shot of the implementation is presented in Figure 3. The main utilities of the implementation are explained in following sub-sections.

8 Interested reader may reach the Calva's article [17] for more information about the Euclidean distance formulations and implementation for the comparison of eigenvalues of a test image to the trained labeled images.

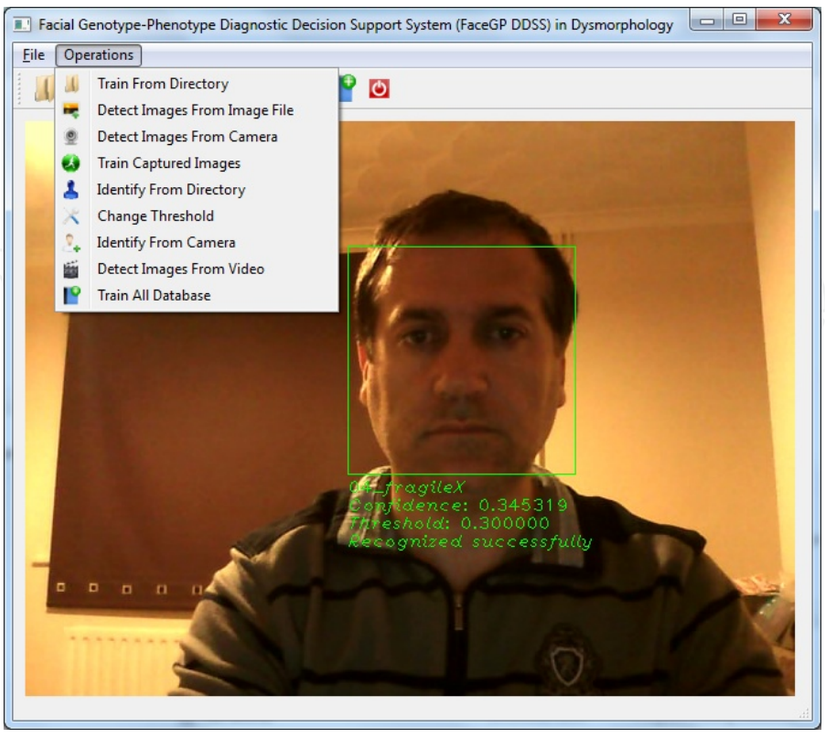

Figure 3. An example for the diagnosing process via camera: Four messages are displayed; the name of the probable diagnosis (e.g. Fragile X), the degree of proximity to that diagnosis (e.g. 0.345319), the threshold value entered by the user (e.g. 0.30), the message about whether a probable disease is found or not as "recognized successfully" or "unknown disease". The messages change if other diseases are found above the threshold value to reveal them on the screen. All the time required to search through 7 trained identified diagnoses that include 34 eigenfaces and find the nearest diagnoses in respect to the threshold value is few seconds.

2.5.1. Detect images from image file:

This utility detects faces from the images displayed on the screen where images are brought from the chosen directory automatically. Detection stops and training process begins when the user clicks the utility "train captured images". The name of the disease is entered by the user.

2.5.2. Detect images from camera:

This utility detects face appearances on live camera from a patient or patients who are diagnosed with same dysmorphic disease. Detection stops and training process begins when the user clicks the utility "train captured images". The name of the syndrome is supplied by the user.

2.5.3. Detect images from video:

Face images of a patient or patients who are diagnosed with same dysmorphic disease are detected from a video or several videos. Detection stops at the end of video and training process begins when the user clicks the utility "train captured images". The name of the syndrome is presented by the user.

http://dx.doi.org/10.5772/51118

2.5.4. Train captured images:

Detected face images from either a directory or a live camera are processed to be classified into the database.

2.5.5. Train from directory:

A directory in which there are several images of a specified disease is chosen and the name of that syndrome is entered by the user. All necessary algorithms are run through training of the diseases automatically. Diseases can be trained one by one by this utility.

2.5.6. Train all database:

All labeled diseases in a directory are processed by "face detection and image acquisition" and "image processing" modules respectively and automatically. Then, processed cropped frontal face images are trained into a database without user intervention by training module. In other words, all modules in Figure 3 are employed automatically. The system accepts the directory names where the dysmorphic images are as the names of the syndromes while training the datasets. The only thing expected from the user is to specify the folder where the datasets are.

2.5.7. Change threshold:

The user can specify the threshold value to rule in or rule out diseases during the identification process. The greater the threshold value is, the less the number of probable diagnoses is proposed, vice versa, the less the threshold value is entered, the greater the number of diagnoses is revealed to the user together with confidence (probability or proximity) values.

2.5.8. Identify from a directory:

Several images in a directory can be chosen at once and these images can be compared to the labeled syndromes stored in the database. A video that has dysmorphic patients can be chosen to diagnose probable diseases as well.

2.5.9. Identify from a camera:

A patient can be captured from a live camera mounted to the computer to diagnose a syndrome stored in the database.

2.6. Evaluation of the Methodology

A case study has been carried out to evaluate the methodology. We have analyzed 2D frontal face pictures of patients, each being affected with one of the syndromes. The scope and the design of the case study are presented in the following subsections. The findings of the study are presented in the next section, results.

2.6.1. Study sample

The study sample was composed of 7 syndromes comprising 35 individual frontal faces (5 for each) of patients that are included from Boehringer's study [2]. These diseases depicted in Figure 4 are microdeletion 22q11.2, Cornelia de Lange, fragile X, Mucopolysaccharidosis III, Smith–Lemli–Opitz, Sotos and Williams–Beuren. The patients for each syndrome are more or less the same age, but, in different genders (10 females and 25 males) for 5 syndromes that are microdeletion 22q11.2 (3 F: Fmales, 2 M: Males), Cornelia de Lange (2F, 3M), Smith–Lemli–Opitz (2F, 3M), Sotos (2F, 3M) and Williams–Beuren (1F, 4M).

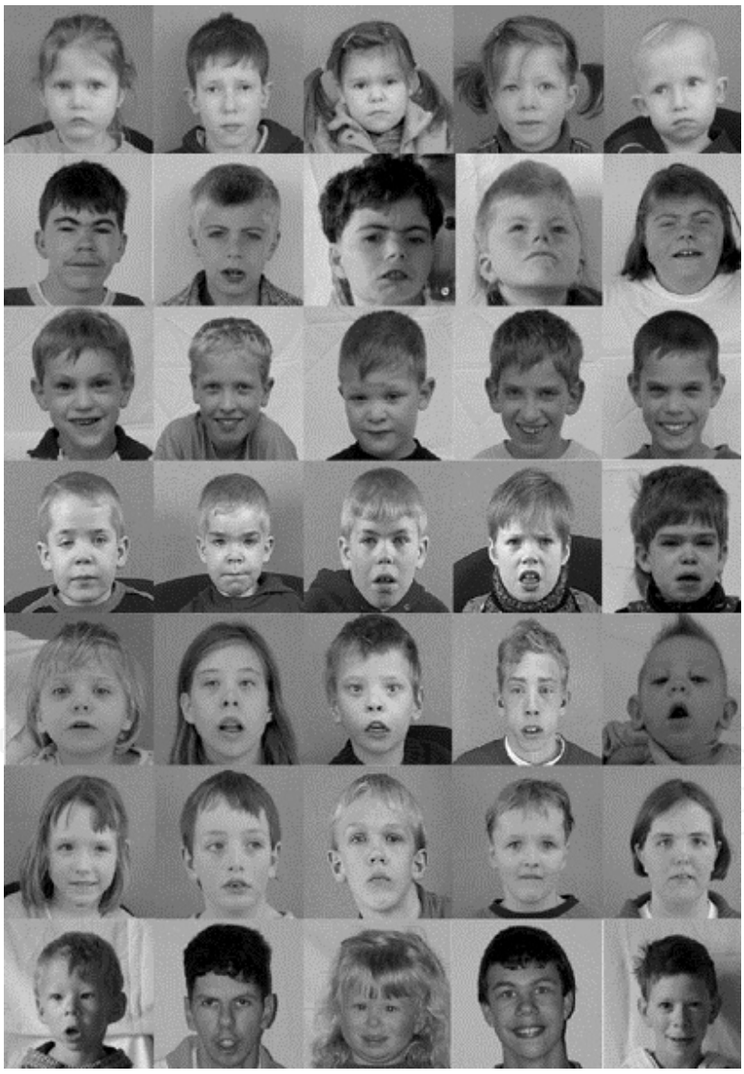

Figure 4. Seven syndromes: each row comprises one syndrome; microdeletion 22q11.2 (3 F: Females, 2 M: Males), Cornelia de Lange (2F, 3M), fragile X (5M), Mucopolysaccharidosis III (5M), Smith–Lemli–Opitz (2F, 3M), Sotos (2F, 3M) and Williams –Beuren (1F, 4M).

http://dx.doi.org/10.5772/51118

2.6.2. Study design

7 syndromes were trained by the utility named "Train all database" mentioned previously. The system could build a training set for 7 syndromes that have 34 eigenfaces for 35 faces less than one minute. The mean face and all 33 eigenfaces are displayed in Figure 5 and Figure 6 respectively. The first four images from the syndromes were used as a testing set to measure the recognition/diagnosis success. All these test images were put in a directory and each test image was compared to other trained 34 images in the database to measure how close it is on the vector space to others by using the utility of "identify from directory". All recognition process lasted about 1 minute in terms of pairwise comparison. This utility produces confidence values in a table and diagnose regarding these values above the threshold value entered by the user. In our case study we have aimed to find the probable diagnoses in the sense of rule-in 1, rule-in 2 and rule-in 3 diagnoses respectively by adjusting the threshold value for each person. Then, the success rates of these n-rule-in observations were obtained.

Figure 5. Mean face for 7 dysmorphic diseases generated by the system.

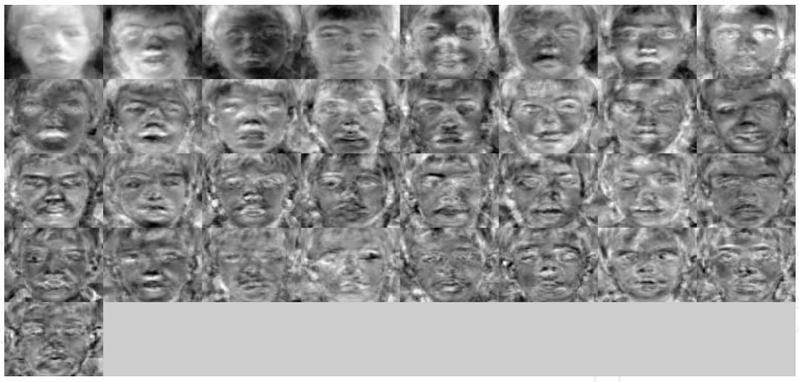

Figure 6. Eigenfaces of 7 dysmorphic diseases that comprise 35 frontal face images.

3. Results

The threshold value entered by the user increases or decreases the possible number of diagnoses, the less the threshold value is entered, the more diagnoses are proposed; vice versa,

the less diagnoses are included to be examined by the medical professionals. Pairwise analysis of 28 patients in 7 different syndromes is depicted in Table 1 regarding the confidence values. Each column corresponds to an individual patient's comparison to other patients one-to-one on the vector space of eigenfaces. The greater the value a column, the greater the probable proximity it has corresponding to that row whose syndrome indicates the probable diagnosis. The threshold value for each column is adjusted to rule in one, two and three diagnoses respectively. The results of three diagnoses for each tested patient are presented in Table 2 in which the gray cells correspond to the right diagnoses. Ruling in one, two and three diagnoses yields 21, 26 (extra 5 to rule in one) and 28 (extra 2 to rule in two) correct diagnoses among all 28 tested patients respectively. Ruling in one, two and three diagnoses designates the success rates of 75, 94 and 100 percent respectively as well. The diagnostic success rates are depicted in Figure 7.

		1	1	1	1	2	2	2	2	3	3	3	3	4	4	4	4	5	5	5	5	6	6	6	6	7	7	7	7
		1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
1	1	1.00	0.32	0.28	0.46	-0.36	-0.34	-0.25	-0.28	-0.07	-0.13	-0.21	-0.17	0.10	0.21	-0.03	0.16	0.14	0.29	0.31	0.06	0.36	0.22	0.06	0.24	-0.28	-0.09	-0.22	-0.35
1	2	0.32	1.00	0.29	0.29	-0.25	-0.28	-0.13	-0.13	-0.03	-0.22	-0.24	-0.22	0.07	0.07	-0.03	0.18	0.12	0.32	0.32	0.09	0.26	0.27	0.09	0.26	-0.24	-0.02	-0.16	-0.29
1	3	0.28	0.29	1.00	0.27	-0.20	-0.21	-0.11	-0.16	0.13	-0.16	-0.16	-0.09	0.05	0.09	-0.09	0.12	0.09	0.22	0.25	0.04	0.23	0.24	0.08	0.21	-0.16	0.03	-0.06	-0.19
1	4	0.46	0.29	0.27	1.00	-0.43	-0.32	-0.31	-0.29	-0.01	-0.12	-0.23	-0.20	0.20	0.31	-0.05	0.12	0.10	0.31	0.38	0.10	0.33	0.19	0.08	0.24	-0.30	-0.16	-0.30	-0.38
1	5	0.25	0.13	0.08	0.35	-0.52	-0.25	-0.35	-0.20	-0.16	-0.07	-0.27	-0.33	0.29	0.39	0.05	0.09	0.07	0.19	0.23	0.07	0.21	0.06	0.00	0.12	-0.30	-0.23	-0.44	-0.42
2	1	-0.36	-0.25	-0.20	-0.43	1.00	-0.09	0.00	-0.20	-0.06	-0.26	-0.06	0.00	-0.38	-0.45	-0.34	-0.30	-0.26	-0.36	-0.35	-0.31	-0.37	-0.24	-0.27	-0.27	-0.11	-0.01	0.09	0.05
2	2	-0.34	-0.28	-0.21	-0.32	-0.09	1.00	-0.11	0.02	0.03	-0.07	-0.03	-0.12	-0.12	-0.14	-0.08	-0.22	-0.16	-0.27	-0.23	-0.20	-0.29	-0.23	-0.15	-0.21	-0.01	-0.05	-0.20	-0.01
2	3	-0.25	-0.13	-0.11	-0.31	0.00	-0.11	1.00	0.12	0.00	-0.29	-0.12	-0.10	-0.18	-0.26	-0.25	-0.07	-0.09	-0.19	-0.20	-0.25	-0.24	-0.10	0.02	-0.08	0.13	0.02	0.00	0.21
2	4	-0.28	-0.13	-0.16	-0.29	-0.20	0.02	0.12	1.00	-0.12	-0.24	-0.22	-0.28	-0.06	-0.11	-0.09	-0.01	-0.10	-0.16	-0.18	-0.25	-0.20	-0.15	-0.08	-0.10	0.04	-0.08	-0.22	0.05
2	5	-0.05	0.05	-0.02	-0.14	-0.13	-0.27	0.16	-0.02	-0.15	-0.41	-0.29	-0.24	-0.16	-0.18	-0.20	0.07	0.05	-0.01	-0.11	-0.23	-0.10	0.00	0.03	0.00	-0.15	-0.16	-0.07	-0.09
3	1	-0.07	-0.03	0.13	-0.01	-0.06	0.03	0.00	-0.12	1.00	0.02	0.09	0.15	-0.06	-0.10	-0.18	-0.14	-0.06	-0.05	0.04	-0.06	-0.05	0.01	0.01	0.05	-0.03	0.13	0.02	0.05
3	2	-0.13	-0.22	-0.16	-0.12	-0.26	-0.07	-0.29	-0.24	0.02	1.00	0.04	-0.03	-0.03	-0.02	-0.12	-0.18	-0.17	-0.17	-0.10	-0.08	-0.10	-0.19	-0.24	-0.06	-0.18	-0.08	-0.27	-0.18
3	3	-0.21	-0.24	-0.16	-0.23	-0.06	-0.03	-0.12	-0.22	0.09	0.04	1.00	0.16	-0.23	-0.21	-0.14	-0.29	-0.12	-0.32	-0.19	-0.18	-0.28	-0.28	-0.27	-0.22	-0.10	-0.01	-0.01	-0.12
3	4	-0.17	-0.22	-0.09	-0.20	0.00	-0.12	-0.10	-0.28	0.15	-0.03	0.16	1.00	-0.27	-0.27	-0.29	-0.30	-0.21	-0.24	-0.15	-0.17	-0.25	-0.18	-0.24	-0.15	-0.12	0.08	0.10	-0.02
3	5	-0.25	-0.32	-0.20	-0.29	-0.02	-0.14	-0.25	-0.36	0.03	0.13	0.18	0.26	-0.29	-0.28	-0.28	-0.37	-0.28	-0.30	-0.26	-0.17	-0.31	-0.29	-0.33	-0.23	-0.20	-0.04	0.01	-0.09
4	1	0.10	0.07	0.05	0.20	-0.38	-0.12	-0.18	-0.06	-0.06	-0.03	-0.23	-0.27	1.00	0.38	0.12	0.19	0.03	0.12	0.15	-0.07	0.14	0.10	0.12	0.20	-0.10	-0.16	-0.37	-0.18
4	2	0.21	0.07	0.09	0.31	-0.45	-0.14	-0.26	-0.11	-0.10	-0.02	-0.21	-0.27	0.38	1.00	0.09	0.16	0.07	0.17	0.17	0.05	0.17	0.02	0.07	0.17	-0.23	-0.17	-0.39	-0.33
4	3	-0.03	-0.03	-0.09	-0.05	-0.34	-0.08	-0.25	-0.09	-0.18	-0.12	-0.14	-0.29	0.12	0.09	1.00	0.04	-0.02	-0.11	-0.05	-0.13	-0.04	-0.11	-0.15	-0.04	-0.26	-0.29	-0.40	-0.32
4	4	0.16	0.18	0.12	0.12	-0.30	-0.22	-0.07	-0.01	-0.14	-0.18	-0.29	-0.30	0.19	0.16	0.04	1.00	0.17	0.24	0.15	-0.11	0.15	0.10	0.11	0.18	-0.12	-0.17	-0.29	-0.19
4	5	-0.11	-0.02	0.06	-0.12	0.05	-0.01	-0.15	-0.10	0.02	-0.10	-0.11	-0.10	-0.14	-0.14	-0.02	-0.02	-0.04	-0.08	-0.03	-0.07	-0.09	-0.05	-0.16	-0.07	-0.26	-0.13	-0.16	-0.21
5	1	0.14	0.12	0.09	0.10	-0.26	-0.16	-0.09	-0.10	-0.06	-0.17	-0.12	-0.21	0.03	0.07	-0.02	0.17	1.00	0.16	0.15	-0.08	0.06	0.05	0.06	0.13	-0.16	-0.14	-0.21	-0.26
5	2	0.29	0.32	0.22	0.31	-0.36	-0.27	-0.19	-0.16	-0.05	-0.17	-0.32	-0.24	0.12	0.17	-0.11	0.24	0.16	1.00	0.40	0.16	0.25	0.27	0.16	0.21	-0.21	-0.05	-0.18	-0.27
5	3	0.31	0.32	0.25	0.38	-0.35	-0.23	-0.20	-0.18	0.04	-0.10	-0.19	-0.15	0.15	0.17	-0.05	0.15	0.15	0.40	1.00	0.14	0.23	0.26	0.11	0.26	-0.25	-0.08	-0.21	-0.31
5	4	0.06	0.09	0.04	0.10	-0.31	-0.20	-0.25	-0.25	-0.06	-0.08	-0.18	-0.17	-0.07	0.05	-0.13	-0.11	-0.08	0.16	0.14	1.00	-0.01	0.02	-0.03	-0.02	-0.31	-0.11	-0.21	-0.32
5	5	-0.13	-0.20	-0.22	-0.11	-0.35	-0.10	-0.39	-0.26	-0.25	0.09	-0.21	-0.30	-0.07	0.02	-0.10	-0.15	-0.09	-0.08	-0.11	-0.08	-0.11	-0.18	-0.19	-0.14	-0.29	-0.24	-0.39	-0.39
6	1	0.36	0.26	0.23	0.33	-0.37	-0.29	-0.24	-0.20	-0.05	-0.10	-0.28	-0.25	0.14	0.17	-0.04	0.15	0.06	0.25	0.23	-0.01	1.00	0.23	0.07	0.23	-0.26	-0.11	-0.27	-0.31
6	2	0.22	0.27	0.24	0.19	-0.24	-0.23	-0.10	-0.15	0.01	-0.19	-0.28	-0.18	0.10	0.02	-0.11	0.10	0.05	0.27	0.26	0.02	0.23	1.00	0.19	0.35	-0.20	-0.03	-0.15	-0.20
6	3	0.06	0.09	0.08	0.08	-0.27	-0.15	0.02	-0.08	0.01	-0.24	-0.27	-0.24	0.12	0.07	-0.15	0.11	0.06	0.16	0.11	-0.03	0.07	0.19	1.00	0.17	-0.02	-0.06	-0.18	-0.05
6	4	0.24	0.26	0.21	0.24	-0.27	-0.21	-0.08	-0.10	0.05	-0.06	-0.22	-0.15	0.20	0.17	-0.04	0.18	0.13	0.21	0.26	-0.02	0.23	0.35	0.17	1.00	-0.19	-0.04	-0.22	-0.19
6	5	0.13	0.15	0.13	0.16	-0.26	-0.29	-0.06	-0.14	-0.02	-0.18	-0.23	-0.11	0.10	0.10	-0.09	0.10	0.05	0.25	0.21	0.05	0.14	0.18	0.11	0.23	-0.19	-0.09	-0.17	-0.20
7	1	-0.28	-0.24	-0.16	-0.30	-0.11	-0.01	0.13	0.04	-0.03	-0.18	-0.10	-0.12	-0.10	-0.23	-0.26	-0.12	-0.16	-0.21	-0.25	-0.31	-0.26	-0.20	-0.02	-0.19	1.00	0.12	-0.01	0.38
7	2	-0.09	-0.02	0.03	-0.16	-0.01	-0.05	0.02	-0.08	0.13	-0.08	-0.01	0.08	-0.16	-0.17	-0.29	-0.17	-0.14	-0.05	-0.08	-0.11	-0.11	-0.03	-0.06	-0.04	0.12	1.00	0.16	0.13
7	3	-0.22	-0.16	-0.06	-0.30	0.09	-0.20	0.00	-0.22	0.02	-0.27	-0.01	0.10	-0.37	-0.39	-0.40	-0.29	-0.21	-0.18	-0.21	-0.21	-0.27	-0.15	-0.18	-0.22	-0.01	0.16	1.00	0.03
7	4	-0.35	-0.29	-0.19	-0.38	0.05	-0.01	0.21	0.05	0.05	-0.18	-0.12	-0.02	-0.18	-0.33	-0.32	-0.19	-0.26	-0.27	-0.31	-0.32	-0.31	-0.20	-0.05	-0.19	0.38	0.13	0.03	1.00

Table 1. Pairwise comparison: Each column corresponds to an individual patient's comparison to other patients one-to-one on the vector space of eigenfaces in terms of the proximity. The greater the value a column has, the greater the probable proximity it has corresponding to that row whose syndrome indicates the probable diagnosis. First column together with first row corresponds to the syndromes (e.g. 1 corresponds to microdeletion 22q11.2). Second column together with second row indicates the number of the images in the specified syndrome in the first column or first row.

		1	1	1	1	2	2	2	2	3	3	3	3	4	4	4	4	5	5	5	5	6	6	6	6	7	7	7	7
		1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
Rule-in	1	0.464	0.323	0.290	0.464	0.085	0.029	0.213	0.124	0.154	0.134	0.184	0.259	0.380	0.387	0.118	0.239	0.166	0.397	0.397	0.164	0.355	0.353	0.189	0.353	0.377	0.160	0.160	0.377
Rule-in	2	0.355	0.321	0.248	0.383	0.046	0.017	0.156	0.047	0.128	0.085	-0.012	0.095	0.288	0.380	0.048	0.186	0.163	0.315	0.383	0.103	0.246	0.266	0.161	0.264	0.129	0.128	0.095	0.213
Rule-in	3	0.307	0.267	0.243	0.331	-0.003	-0.005	0.017	-0.005	0.126	-0.015	-0.026	0.003	0.199	0.172	-0.024	0.183	0.125	0.266	0.258	0.050	0.231	0.265	0.118	0.256	-0.024	0.034	0.085	0.052

Table 2. Ruling in one, two and three diagnoses regarding the greatest values in columns of Table 1: the grey cells correspond to the right diagnosis. Ruling in one, two and three diagnoses yields 21, 26 and 28 correct diagnoses with success rates of 75, 94 and 100 percent among all 28 tested patients respectively. For instance, the value, 0.464, is observed from the cell where the syndrome name is coded as 1 and the frontal face is coded as 4. This indicates that

http://dx.doi.org/10.5772/51118

the frontal face which is coded as 1 in the syndrome list of 1 is the most proximate with a confidence value of 0.464 that is the greatest value in that column.

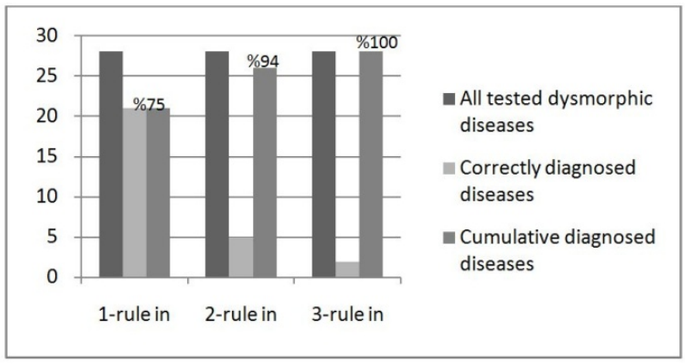

Figure 7. Graphical representation of success rates with respect to ruling in one, two and three diagnoses concerning the values in Table 2. Ruling in one, two and three diagnoses yields 21, 26 (including 5 more) and 28 (including 2 more) correct diagnoses among all 28 tested patients respectively. This designates the success rates of 75, 94 and 100 percent respectively as well.

4. Discussion

4.1. Main findings

FaceGP DDSS methodology has a success rate of 75%, 94% and 100% in terms of ruling in one, two and three diagnoses respectively. The results show that FaceGP DDSS methodology is able to make a biometric identification among syndromes successfully and efficiently based on the features of the patients' frontal faces, even though, the methodology has been tested by a limited number of 7 syndromes. Diagnosing syndromes correctly among many syndromes can be eased by the methodology provided that it is trained with those syndromes. One specific result of the study is that all test frontal faces of three syndromes that are microdeletion 22q11.2, fragile X and Williams–Beuren are correctly diagnosed in ruling in one diagnosis. Ruling in two diagnoses covers the 5 syndromes correctly whereas the ruling three does all seven syndromes. From these results we can conclude that the implemented methodology especially can guide medical professionals to employ correct cyto- and/or molecular genetic analysis that is the appropriate route of investigation in order to confirm a diagnosis with known genetic causes by ruling in probable diseases. Even with adequate knowledge, there remains the problem of reconciling sometimes imprecise descriptions of dysmorphic features in the literature with a personal and potentially subjective examination of an individual patient [7].

Preliminary results indicate that the application can be trained with many syndromes in several minutes and syndrome recognition can be established in few seconds either from an attached camera or from a file. We expect that the performance of the system doesn't degrade as the syndrome number grows owing to the computational efficiency of PCA. Moreover, multiple

diagnostic as well as multiple training of syndromes effectively and efficiently can be employed easily with the implementation which will attract medical professionals most. During training process, the larger the training dataset per syndrome, the better the success of syndrome recognition is, thanks to the pair wise comparison in our study, regarding different characteristics even in same syndromes such as Down syndrome that has three subtypes[9] and cri du chat syndrome that has several variations in the feature[10]. Our implementations may be used to assist in diagnosing and defining facial phenotypes associated with syndromes in different ethnic groups and in different age groups provided that these kinds of cases are included into the training process of a specific syndrome.

4.2. Comparison to other published studies

Hammond [7] has pointed out that there are well documented approaches to recording craniofacial dysmorphology in a more objective fashion. Moreover, he [7] has asserted that international experts in dysmorphology are currently developing standardized terminology to address issues of imprecision and inconsistency. Most genome-wide association studies to date have focused on a limited number of specific diseases or traits to test whether a computer can classify syndromes and then recognize them when compared to new cases. There have been successful discrimination studies using images of children with a limited number of dysmorphic syndromes [2, 5, 6, 16]. Some robust composite computer assisted decision support systems are needed to be established not only for practitioners but also geneticists to support their decisions through thousands of dysmorphic diseases. FaceGP DDSS methodology is aimed to serve several needs of medical professionals who work in dysmorphology. This study presenting the FaceGP DDSS methodology contributes to the medical environment in several aspects. One of which is the support of general practitioners or pediatricians in the rural areas rather than trained geneticists as well as making easy of works for geneticists throughout thousands of dysmorphic diseases defined in some catalogs and databases. The other one is that a limitless number of dysmorphic diseases can be trained and tackled with FaceGP DDSS methodology in diagnosing process simultaneously. On the other hand, other similar studies having a fixed number of syndromes up to 10 in number are implemented to reveal the potential rather than to be an application to serve medical professionals for their everyday needs. Right diagnosis and consequently right treatments can influence progression of disease, especially in term of removing the effects of environmental factors[11]. For instance, when you supplement a patient having a syndrome with hormone, the patient may get better. The FaceGP DDSS methodology is designed primarily for investigators who wish to diagnose their patients with dysmorphic diseases quickly, effectively and successfully. Furthermore, it aims to support scientists for their studies who do not have expertise in the particular domain of dysmorphology. A new understanding of rul-

9 Chromosome 21 can be affected in three main ways, leading to the three main sub-types of Down's syndrome. Full trisomy 21 Down's syndrome, Mosaicism Down's syndrome and Translocation Down's syndrome are the three subtypes of Down syndrome.

10 Wilkins's findings delineate the variation in the clinical and karyotypic features of cri du chat syndrome [20].

11 The complex interplay of genetic and environmental factors is a significant confounding factor in various human genetic approaches, including genome-wide and candidate gene association studies as well as linkage analysis [21].

http://dx.doi.org/10.5772/51118

ing in/ruling out diseases in our methodology can be very helpful for geneticists who wish to employ cyto- and/or molecular genetic analysis for their cases to confirm probable diagnoses. Moreover, the application is ready to be used with a user friendly interface when implemented at any site. Our methodology, noticeably different from others, doesn't require any manual intervention or preprocessing of images by users, rather, all necessary algorithms have been embedded into the methodology.

Medical professionals may construct their own databases in terms of their own dysmorphic facial findings, thus facilitating incorporation of their findings into the examination as well as may add their special cases into a formerly constructed database. In addition, these databases, later, can be shared in a web environment, can be easily used as plug-and-play separately and furthermore, valuable databases can be united in a unique database after their evaluation by an expert group to be served to the knowledge of the scientific community. This opens the door to cross-study analyses of not only the primary genotypes-phenotypes dysmorphological diseases (prevalence < 1/25.000) but also secondary genotypes-phenotypes syndromes (prevalence < 1/25.000, maybe 1/100.000). However, probable specified genotype-phenotype diseases should be trained before presenting the application to the use of the medical professionals to gain their support.

Users could benefit the methodology with a user friendly interface without any manual intervention which may cause the users to avoid the use of any system. These findings refute the notion presented in several studies that manual steps cannot be excluded entirely from any dysmorphic facial analysis software that intends to extract as much information as possible.

4.3. Limitations of the study

There are several concerns that we should keep in mind while implementing these kinds of studies. Some genetic dysmorphic diseases could not be brought into scientific literature for especially two reasons. One of which is that many cases are lost before birth and nobody doesn't have any incentive to investigate the reason and eventually no one knows the cause; the other and the most important one is that families facing nonspecific cases don't permit the geneticists to investigate the cases, even though the investigation would help the families greatly for their future decisions[12] and consequently babies or fetus are buried with their genetic secrets in the sense of the ethical rules. On the other hand, in practice, some genetic dysmorphic diseases that are very rare (e.g. prevalence 1/100.000) are not globally defined and they are stored in the local Electronic Medical Records (EMRs) or research databases across different medical institutions without any common data structure or representational format. These data elements are needed to be presented to the knowledge of the genetic communities on behalf of the current and future healthy population. Our ability to fully understand the genetic basis of common diseases is significantly hindered by the inability to precisely specify the phenotypes and in particular, identifying and extracting phenotypes at large varies greatly between different medical specialties and institution, and lacks the sys-

12 Such as having a dysmorphic baby (the recurrence risk: the possibility that the problem would happen again in another pregnancy) and more than that such as early diagnosis, disease prevention, patient management, or even adjunctive therapies to be developed.

tematization and throughput compared to large-scale genotyping efforts [22]. Beyond these clinical aspects, dysmorphology has contributed much to current understanding of the genetic basis of human development [1]. Moreover, imprecise and nonstandardized nomenclature, especially of facial features, places a major difficulty for the communication between clinical geneticists [2]. It has to be noted that neither 2D nor 3D methods have direct applicability in clinical practice yet, as the number of specified syndromes is still very small [2].

As Boehringer [2] emphasize, database support with respect to facial traits is limited at present to apply similar studies as we do to establish better applications. Distinctive dysmorphic frontal faces specific to dysmorphic genotype-phenotype diseases are needed to train the system. Currently, in our study, a very limited number of dysmorphic genetic diseases by using frontal faces have been trained for further recognition process. There are several genetic databases such as eMERGE (Electronic Medical Records and Genomics) and PhenX (Consensus Measures for Phenotypes and Exposures), Dysmorphology Database in Oxford Medical Databases (OMD) and OMIM. One of which named OMD is more appropriate for our study, because it is better prepared to reveal genotype-phenotype associations in terms of images and taxonomy of dysmorphology, although it has very limited number of frontal faces for syndromes. One of the reasons that we work on frontal 2D image analysis is that this prominent database (OMD) that we aim to include into the study contains 2D genetic dysmorphic images rather than 3D videos by which the number of frames are captured and recorded. Moreover, most of the geneticists studying on dysmorphic diseases usually keep 2D images of their patients in their databases. The main drawback of the majority of 3D face recognition approaches is that they need all the elements of the system to be well calibrated and synchronized to acquire accurate 3D data (texture and depth maps) [19]. That will make it easier for investigators to collect and analyze the 2D dysmorphic data associated with genotypes. Whereas capturing of 3D information results in a richer data set and allows for excellent visualization despite the difficulties in possessing the technology and in detecting 3D as mentioned by Kau [13][13], 2D analysis has several advantages in practical use: equipment is cheap and it is easy to handle [2]. Conventional and digital two-dimensional (2D) photography offer rapid and easy capture of facial images. 3D analysis of syndromes would sure give better results as depicted in Hommond's study [6]. However, the lack of available data in 3D invalidates any methodology implemented for the near future.

Of course, recognition of face shape does not imply a diagnosis. A diagnosis is made by an appropriately trained clinician backed up, whenever possible, by genetic testing. For some dysmorphic syndromes there is no definitive genetic test and a clinical diagnosis has to suffice. For others, for example Noonan syndrome, a number of important genes may have been identified, but mutations for those genes may not be found in some children for whom there is a compelling clinical diagnosis [3].

A masking is not applied to remove the background of the cropped faces in our study. By employing a background mask, which simply provides a face shaped region, the effect of

13 Due to inherent faults in technology and the distortion of light, none of the 3D imaging systems is accurate over the full field of view. Furthermore, all systems suffer from a potential for patient movement and alterations of facial expression between the multiple views needed to construct a 3D model of the face.

http://dx.doi.org/10.5772/51118

background change would be minimized and diagnosing success would be better which is going to be an improvement as a further study. Furthermore, we couldn't reach the raw material of 7 syndromes in which the photos we utilized in our study are not in good condition in terms of their appearances. Moreover the patients are not same ages and same sexes for each syndrome. The results would be better if better images were utilized and if the patients were in similar age group and sexes in the study. The more faces bearing the characteristics of any syndrome included in the training, the better the recognition of that syndrome will be.

5. Conclusion

In terms of future technological support, two (2D) or three-dimensional (3D) models of facial morphology are showing potential in syndrome delineation and discrimination, in analyzing individual dysmorphology, and in contributing to multi-disciplinary and multi-species studies of genotype–phenotype correlations [7]. Our study is an example of substantiating this potential. We describe a new approach to syndrome identification by merging several algorithms. The algorithms that we included in our study are not novel. They have been utilized in many studies so far. However we included most essential ones in a robust composite understanding in a way to serve the everyday needs of the medical professionals who work in dysmorphology.

The preliminary results indicate that computer based diagnostic decision support systems such as the one we have established might be very helpful to assist medical professionals in genotype-phenotype dysmorphic diagnosis. The study reveals that the differences between facial regions such as facial landmarks, eyebrows, hair, lips, and chins can give the possibility of predicting the diagnosis of syndromes. It may contribute to the medical professionals in several aspects. Some of these are:

- To support medical professionals who do not have expertise in the particular domain of dysmorphology such as general practitioners or pediatricians in rural areas,
- To support geneticists throughout thousands dysmorphic diseases, most of which are very rare and difficult to memorize,
- Generally to support investigators who wish to diagnose their patients with dysmorphic diseases quickly, effectively and successfully,
- To support investigators who strive to expand their studies by including their cases into the system with the implementation whose database is able to be broadened dynamically and easily, which provides them to keep and deal with their data more efficiently.
- To guide geneticists to employ correct cyto- and/or molecular genetic analysis that is the appropriate route of investigation in order to confirm a diagnosis with - known genetic causes by ruling in probable diseases.
- No preprocessing of data manually that may cause the users to avoid the utilization of any system is required.

- FaceGP DDSS methodology can provide genetic screening which is a preliminary process of applying standard analysis to large populations to pick up underlying symptoms of genetic disorders. Genetic screening is not a diagnosis, but can produce a differential diagnosis which would lead to a definitive diagnosis and hence to early intervention and treatment.

The hope is that the FaceGP DDSS methodology will be widely adopted by the scientific community, fostering a new era of cooperation and collaboration and facilitating cross-study. Based on user feedback, we expect to continue to update the functionality of the methodology. As the data gathering for age groups and ethnic groups becomes more standardized and evolved internationally in the sense of dysmorphology, general implementations valid for everybody will be more possible.

5.1. Future work

Further improvement in diagnosing/recognition seems to be possible by integrating a background cropping mask algorithm that simply provides a face shaped region and minimize the effect of background change [15].

The dysmorphic faces from main databases as well as from individual databases referring to the names of the diseases should be both categorized and trained by the application for further better diagnostic decision support. This is a huge time consuming and challenging process needed to be done in the near future for the easy acceptance of the methodology by the scientific community. We intend to extend this work to a wider environment by including domain experts from academic and government institutions by deploying the methodology at several sites including as possible as many syndromes. Furthermore, 3D image processing and fetus image analysis in dysmorphology is going to be the subject of our future study.

Acknowledgements

The authors are very grateful to TÜBİTAK (The Scientific and Technological Research Council of Turkey) for sponsoring the study.

Author details

Kaya Kuru[1*] and Yusuf Tunca[2]

*Address all correspondence to: kkuru@gata.edu.tr

1 IT Department, Gulhane Military Medical Academy (GATA), Ankara, Turkey

2 Department of Medical Genetics, Gulhane Military Medical Academy (GATA), Ankara, Turkey

http://dx.doi.org/10.5772/51118

References

[1] Smithson, S. F., & Winter, R. M. (2004). Diagnosis in Dysmorphology: Clues from the Skin. *The British Journal of Dermatology*, 151(5), 953-60.

[2] Boehringer, S., Vollmar, T., Tasse, C., & Wurtz, R. (2006). Syndrome identification based on 2D analysis software. *European Journal of Human Genetics*, 14(10), 1082-1089.

[3] Hammond, P. (2011). What's in a face? 31 March 2008 http://www.hospitalmanagement.net/features/feature1767/ (accessed 21 September 2011

[4] Farkas, L. G. (1998). Anthropometry of the head and face. New York, Raven Press.

[5] Loos, H. S., Wieczorek, D., Wurtz, R. P., et al. (2003). Computer-based recognition of dysmorphic faces. *Eur J Hum Genet*, 11(8), 555-60.

[6] Hammond, P., Hutton, T. J., Allanson, J. E., et al. (2005). Discriminating power of localized three-dimensional facial morphology. *Am J Hum Genet*, 77(6), 999-1010.

[7] Hammond, P. (2007). The use of 3D face shape modelling in dysmorphology. *Arch Dis Child*, 92(12), 1120-1126.

[8] Douglas, T. S., & Mutsvangwa, T. E. M. (2010). A review of facial image analysis for delineation of the facial phenotype associated with fetal alcohol syndrome. *Am J Med Genet Part A*, 152A(2), 528-536.

[9] Kapoor, S., Khanna, S., & Bhatia, R. (2010). Facial gesture recognition using correlation and mahalanobis distance. *International journal of computer science and information security*, 7(2), 267-272.

[10] Turk, M. A., & Pentland, A. P. (1991). Face recognition using eigenfaces. *IEEE Conf. on Computer Vision and Pattern Recognitio*, 586-591.

[11] Iancu, C., Corcoran, P., & Costache, G. (2007). A review of face recognition techniques for in-camera applications. In Signals, Circuits and Systems, 2007. ISSCS 2007. International Symposium on July (2007). , 1, 1-4.

[12] Aldridge, K. A., Boyadjiev, S. T., Capone, G., et al. (2005). Precision and Error of Three-Dimensional Phenotypic Measures Acquired From 3dMD Photogrammetric Images. *American Journal of Medical Genetics*, 138A(3), 247-253 .

[13] Kau, C. H., Rich, Mond. S., Incrapera, A., English, J., & Xia, J. J. (2007). Three-dimensional surface acquisition systems for the study of facial morphology and their application to maxillofacial surgery. *Int J Med Robotics Comput Assist Surg*, 3(2), 97-110.

[14] Gul, A. B. (2003). Holistic Face Recognition by Dimension Reduction. *Master Thesis. METU*.

[15] Akalin, V. (2003). Face Recognition Using Eigenfaces and Neural Networks. *Master Thesis. METU*.

[16] Tripathi, B. K., & Kalra, P. K. (2010). High dimensional neural networks and applications. In: Prathiar DK., Jain LJ. (ed) Intelligent autonomous systems: foundations and applications. Springer , 215-233.

[17] Calvo, A. R., Rurainsky, J., & Eisert, P. (2009). D-3D Mixed Face Recognition Schemes. In: Girgic M et al. Recent Advances in Multimedia Signal Processing and Communications. Springer , 121-144.

[18] Viola, P. A., & Jones, M. J. (2001). Rapid object detection using a boosted cascade of simple feature. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Kauai, USA: IEEE , 551-518.

[19] Dalal, A. B., & Phadke, S. R. (2007). Morphometric analysis of face in dysmorphology. *Comput Methods Programs Biomed*, 85(2), 165-72.

[20] Wilkins, L. E., Brown, J. A., Nance, W. E., & Wolf, B. (1983). Clinical heterogeneity in 80 home-reared children with cri du chat syndrome. *J Pediatr*, 102(4), 528-533.

[21] Mclnnes, R., & Michaud, J. L. (2004). Developmental Biology: Frontiers for Clinical Genetics. *Clin Genet*, 65(2), 163-176.

[22] Pathak, J., Pan, H., Wang, J., Kashyap, S., et al. (2011). Evaluating Phenotypic Data Elements for Genetics and Epidemiological Research: Experiences from the eMERGE and PhenX Network Projects. *AMIA Summits Transl Sci Proc.*, 41-45.

IntechOpen