Facial Image Classification of Mouse Embryos for the Animal Model Study of Fetal Alcohol Syndrome

Shiaofen Fang,
Computer Science Department, Indiana University Purdue, University Indianapolis, 723 W. Michigan St., SL 280, Indianapolis, IN, 46202, USA, Ph. 01-317-274-9727

Ying Liu,
Computer Science Department, Indiana University Purdue, University Indianapolis, 723 W. Michigan St., SL 280, Indianapolis, IN, 46202, USA, Ph. 01-317-274-9727

Jeffrey Huang,
Computer Science Department, Indiana University Purdue, University Indianapolis, 723 W. Michigan St., SL 280, Indianapolis, IN, 46202, USA, Ph. 01-317-274-9727

Sophia Vinci-Booher,
Dept. of Anatomy & Cell Biology, Indiana Univ. Medical Center, 635 Barnhill Dr., MS 508, Indianapolis, IN 46202, USA, Ph. 01-317-274-7359

Bruce Anthony, and
Dept. of Anatomy & Cell Biology, Indiana Univ. Medical Center, 635 Barnhill Dr., MS 508, Indianapolis, IN 46202, USA, Ph. 01-317-274-7359

Feng Zhou
Dept. of Anatomy & Cell Biology, Indiana Univ. Medical Center, 635 Barnhill Dr., MS 508, Indianapolis, IN 46202, USA, Ph. 01-317-274-7359

Shiaofen Fang: sfang@cs.iupui.edu; Ying Liu: liu53@cs.iupui.edu; Jeffrey Huang: bioinfo101@gmail.com; Sophia Vinci-Booher: svincobo@iupui.edu; Bruce Anthony: cbanthon@iupui.edu; Feng Zhou: icmc100@iupui.edu

Abstract

Fetal Alcohol Syndrome (FAS) is a developmental disorder caused by maternal drinking during pregnancy. Computerize imaging techniques have been applied to study human facial dysmorphology associated with FAS. This paper describes a new facial image analysis method based on a multi-angle image classification technique using micro-video images of mouse embryo. Images taken from several different angles are analyzed separately, and the results are combined for classifications that separate embryos with and without alcohol exposures. Analysis results from animal models provide critical references for the understanding of FAS and potential therapy solutions for human patients.
1. Introduction

Fetal Alcohol Syndrome (FAS) is a developmental disorder caused by maternal drinking during pregnancy. Children with FAS exhibit short stature, low IQ, learning and memory deficit, low executive function, and an array of cognitive, behavioral, and affective disorder [10]. Currently, the first as well as the representative diagnosis of FAS is the facial dysmorphology extracted from reported population with established maternal drinking history. The typical facial dysmorphology, including short palpebral fissures, flat upper lip, flattened philtrum and midface hypoplasia, only represent in a subpopulation of children from the pregnancy of alleged drinker. However, not all pregnancy drinking result in typical “FAS” face; at absence of typical facial dysmorphology wide range of neurodevelopmental deficits exist. A broader spectrum of series mentally and physically disabled children with or without “typical facial dysmorphology” born to pregnant mothers who drink is now adopted as fetal alcohol spectrum disorder (FASD). It is estimated that the FASD population is about 10 times of that in FAS. This makes the diagnosis difficult and challenging, and demands improvement of the sensitivity and effectiveness for this expanded set of conditions.

Animal FAS model allowing for defined alcohol exposure pattern and concentration at specific developmental stage will be a valuable tool to fine-gauge the facial dysmorphology in better understanding diagnosis of FAS. Furthermore, the association or dissociation of facial dysmorphology and neurodevelopment deficit are consequential to these variables. It is hope that by establishing a computational analysis of facial feature detection algorithm, it can be used to test the divergence of facial dysmorphology under the potential variable — dose, pattern, and stage of alcohol exposure, and the genetic diversity. The goal of this study is to test the feasibility of this approach.

We report here that using micro-video digitization, high resolution facial images can be acquired starting as early as embryonic stage for early diagnosis. The digitalized images were collected from a series of fixed camera angles. We introduce a new multi-angle image analysis method that takes advantage of the availability of high resolution facial images taken from multiple angles. The algorithm is able to provide accurate and reliable classifications of mouse embryo facial images that separate normal embryos from the ones with alcohol exposures.

2. Related Work

Using image processing and pattern recognition techniques for the classification of FAS faces has been previously reported. In [9], a 2D image analysis technique was proposed. The result was not very satisfactory as far as the classification rate is concerned. In [6], a 3D technique was applied using 3D facial images from laser scanners. The 3D result is significantly improved as it uses true geometric shape information. Unfortunately, 3D shape information for mouse embryos is very difficult to acquire as the embryo size is too small for any 3D scanner. Our strategy in this paper is to collect multiple 3D images from each subject and apply 2D image analysis on multiple image samples and then combined the results to generate a collective classification result. As discussed in [17], multiple images can often improve face recognition performance, we believe the same will hold true for mouse embryo image classification.
Most of the techniques used in this work are related to facial recognition. There are two major approaches for automated identification of human faces. The first approach, the abstractive one, extracts (and measures) discrete local features ‘indexes’ for retrieving and identifying faces, while standard statistical pattern recognition techniques are then employed for matching faces using these measurements. The other approach, the holistic one, conceptually related to template matching, attempts to identify faces using global representations. Characteristic of this approach are connectionist methods such as back-propagation neural networks (‘holons’), principal component analysis (PCA), and singular value decomposition (SVD). Some recent work on the early stages involved in face recognition would model the human head as an elliptical structure in order to segment it [14]. As the human face is an elastic object, Arad [1] suggest modeling facial expression and the resulting image warping using RBF learning so image normalization can be achieved.

The experiments performed by Brunelli and Poggio [5] suggest that the optimal strategy for face recognition is still holistic and corresponds to template matching. Although recognition by matching raw images has been successful under limited circumstances [2], it suffers from the usual shortcomings of straightforward correlation-based approaches, such as sensitivity to face orientation, size, variable lighting conditions, and noise. The reason for this vulnerability of direct matching methods lies in their attempt to carry out the required classification in a space of extremely high dimensionality. To overcome the curse of dimensionality, the connectionist equivalent of data compression methods is employed first. Principal Component Analysis (PCA) is a popular technique used to derive a starting set of features.

Turk and Pentland popularized the use of PCA for face recognition [16] and defined a subspace whose basis vectors, called eigenfaces, are principal components of face database. Propose a probabilistic method was proposed in [12] for face recognition based on eigenfaces. To improve the PCA stand alone performance, this representation criterion is combined with a discrimination criterion. A widely used discrimination criterion is Fisher Linear Discriminant (FLD) [3,7]. Swets and Weng [15] proposed the multidimensional discriminant analysis and linear projection to define Discriminant Karhunen Loeve (DKL) Projection.

3. Data Collection

2D digital images of the alcohol-treated and control (non-alcohol-treated) embryos are taken by high-resolution digital camera. The Micro-video imaging offers a high throughput image acquisition. Each sample image capture process takes less than 3 minutes. Micro-video captures 180° digital images under a stereomicroscope for high-resolution surface structure and conformation. Embryos (whole body by emersion) or postnatal samples (by perfusion) were fixed in 4% paraformaldehyde. Embryos were placed in a cylindrical plexi-glass cast with a central canal that accommodates the girth of the embryo body, holding it firmly in the cast so that the head and upper forearms were stably exposed outside the cast (Fig. 1). The plexi-glass cast was mounted to a 1RPM direct drive bi-rotational motor (Hurst model PA, Princeton, IN). Videos of embryo heads were taken using a Leica MZFLIII dissection scope at 10× magnification (Leica Microsystems, Bannockburn, IL). Images were captured using a Spot Insight Color Camera 3.2.0 at a capture speed of 1 capture/second with Spot Advanced software (Diagnostic instrument inc, Sterling Heights, MI). Videos capturing will start at -90° of full face and rotated clockwise to 90° giving a full 180° digital lateral-to-lateral image sequence.

4. Image Analysis For Mouse Embryo Classification

We aim to develop a computerized facial feature analysis technique that can accurately discriminate FAS subjects from control subjects. A machine learning approach is employed in our technique. A leave-one-out cross validation approach is used to validate the classification...
results. The algorithm works in four steps. (1) Face alignment. This is to ensure the proper alignment of facial features between different face images to facilitate feature comparisons and analysis. (2) Feature selection. It identifies and extracts the set of pixels on the face images that have the maximum discriminatory power in separating FAS and controls. (3) Feature analysis. In this step, pattern recognition techniques are applied to the detected features to generate an FAS classification function (with the features as variables), the classifier, that can distinguish FAS faces from controls. (4) Validation.

4.1 Image Alignment

Before alignment, all color images were first converted to grayscale images, as only intensity values were used in feature analysis. In order to properly compare different facial images, all images needed to be precisely aligned to each other so that the computer algorithm can accurately build correspondence between pixels that represent the same facial feature. The alignment process starts by first defining a template face, randomly picked from the samples, and then aligning each facial image with the template face. The template face is used here as only a reference face for alignment purpose without much of a significance otherwise. The alignment is achieved using a landmark-based image registration technique: local weighted mean method [13]. 24 landmark points were picked on each face to build a reference for correspondence. A transformation was then computed for each new face image such that all its landmark points are mapped precisely to the corresponding landmark points of the template face. The final step of the alignment is to trim each facial image into a uniform size of 300 × 350 pixels containing a standard area of the face.

4.2 Feature Selection

In image analysis, features represent information in (or computed from) the image that are relevant to the analysis goals. In 2D images, pixel values (intensities) at different pixel locations are typically used as features. The number of pixels in each image, however, is very large compared to the number of image samples available. In pattern recognition, more features do not necessarily lead to better classification results [4]. Thus, feature selection (or dimension reduction) are essential in feature analysis. Facial feature selection involves the derivation of salient features from the facial image data in order to reduce the amount of data used for classification and simultaneously to provide enhanced discriminatory power. If we consider each pixel intensity values as features, we are facing a 300 × 350 dimensional feature space. Feature reduction requires the projection of this high-dimensional data into a lower dimensional space.

Principal Components Analysis (PCA) [16] was applied in our method. PCA has been shown to be a statistically optimal solution to transform image features into much lower dimensions. It derives a new set of variables based on correlation of features (pixels) through learning a transformation function from a set of images. Given an m-dimensional vector representation of n images, PCA seeks to find a d-dimensional subspace (d ≪ m), whose basis vectors correspond to the maximum variance direction in the original image space. The principal components (maximum variance directions) are usually represented as a linear combinations of original pixels, and are calculated as the eigenvectors of the covariance matrix of the set of images. Figure 2 shows some of the PCA features.

Traditionally, eigenvectors are ordered based on their eigenvalues so that the top eigenvectors represent the set of features (pixels) that provide the maximum representational power. In our analysis, however, we are primarily interested in the type of features that are the most discriminatory in separating alcohol-treated and control subjects. Therefore the goal of feature selection is to find the optimal feature set that provides the most discriminatory power. A Correlation-based Best First search algorithm is employed in our approach. The Best First
search starts with an empty set of features and generates all possible single feature expansions. The subset with the highest evaluation is chosen and is expanded in the same manner by adding single features. If expanding a subset results in no improvement in discrimination of the two datasets, the search drops back to the next best unexpanded subset and continues from there. This is based on the hypothesis that “Good feature subsets contain features highly correlated with (predictive of) the class, yet uncorrelated with (not predictive of) each other” [8].

The evaluation of each feature subset is done using a Linear Discriminant Analysis (LDA) technique [11]. The aim is to find a projection matrix that maximizes the ratio of distances between classes and distances within each class in order to find the best separation boundary for different classes. The optimal subset is the one that provides the best LDA performance with the training set.

4.3 Feature Analysis for Classification

The selected feature vector is analyzed using a pattern classifier. There are many powerful machine learning based data classifiers. Radial Basis Function Networks classifier is used in our study.

We had also experimented with other classifiers (e.g. Support Vector Machine), and found that the results are generally similar. Radial Basis Function Networks (RBFN) is a special neural network classifier for supervised learning. It is a multilayer, feed-forward neural network that is well suited to applications such as pattern discrimination and classification.

The image analysis process described so far is applied to a set of images taken from the same angle. Since the micro-video images were taken from a sequence of pre-determined angles, we can apply this process multiple times, each time for one specific angle. As they all use the same group of mouse embryo subjects, it is natural to merge the results from all angles to come to a combined classification solution. In this experiment, we employed a simple voting system, in which the final classification for a given test subject is the majority of the classification results from all test angles.

4.4 Validation

A Leave-One-Out cross validation approach is used for testing in our study. In this approach, one face is selected as a test set from the entire sample of $n$ faces, and the rest of the $n-1$ faces are treated as the training set. This process will repeat $n$ times so that each face will be used as a test set exactly once. A more rigorous validation approach would be a 3-fold cross validation in which both the alcohol treated group and the control group were divided into three parts of equal sizes, and each time, one-third of the datasets are put aside as a test set and the rest of the datasets are used as training set. But this usually requires a very large sample to be reliable. We plan to implement the 3-fold cross validation once we collected a large enough sample.

5. Results

Micro-video images of a total 64 mouse embryos were collected. For each subject, images were taken from a sequence of camera angles. Our experiment used 5 angles: the frontal, plus and minus 13 degrees, and plus and minus 39 degrees. The final classification results were obtained by voting from the classification results of the 5 angles. Table 1 shows the prediction results for the 64 Leave-One-Out cross validation tests. The overall classification rate is 89%, which is excellent for this type of datasets. We also notice that the classification rate for the individual angles are significantly lower, as shown in Table 2. This indicates that the multi-angle approach can significantly improve the results from single angle 2D image analysis.
6. Conclusions

In the method presented in this paper, we used for the first time an animal model of computational algorithm for the facial imaging to examine the effects of the alcohol exposure to the facial development mice embryo. The study demonstrated that advanced computational methods such as machine learning, pattern recognition, and computational imaging, can be effectively applied to the diagnosis of animal facial dysmorphology. The multi-angle approach clearly demonstrates power in improving classification rates of the individual angles, and shows that 3D scanning or reconstruction may not be necessary in order to achieve high classification rate.

In the future, we will continue improving this algorithm with a larger set of embryo samples. One important aspect of animal study for FAS is the identification of common and unique facial features that are the most discriminatory for both human and mouse. We plan to develop an algorithm to reconstruct and back-project the facial features identified on 2D images onto a 3D face so that proper comparisons with human face can be made.

Acknowledgments

This work is supported in part by NIH-NIAAA grants 1U01AA017123-01. We would like to thank Tatiana Foroud, Rick Ward, Li Shen, and Elizabeth Moore for their valuable comments and discussions, and Jeff Rogers for technical support.

References

Figure 1. Capturing Micro-video embryo images on a turntable under stereomicroscope
Figure 2. PCA features
Table 1
Combined classification results (rate: 89%)

<table>
<thead>
<tr>
<th>Known Classes</th>
<th>Prediction</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Alcohol (+)</td>
</tr>
<tr>
<td>Alcohol Treated</td>
<td>31</td>
</tr>
<tr>
<td>Control</td>
<td>6</td>
</tr>
<tr>
<td>Angle (degrees)</td>
<td>Classification Rate</td>
</tr>
<tr>
<td>----------------</td>
<td>---------------------</td>
</tr>
<tr>
<td>-39</td>
<td>71.9%</td>
</tr>
<tr>
<td>-13</td>
<td>73.4%</td>
</tr>
<tr>
<td>0</td>
<td>74.1%</td>
</tr>
<tr>
<td>13</td>
<td>71.9%</td>
</tr>
<tr>
<td>39</td>
<td>71.9%</td>
</tr>
</tbody>
</table>

Table 2: Classification rates from all angles.