International Journal of
Soft Computing And Software Engineering (JSCSE)
ISSN: 2251-7545
Vol.1,No.1, 2011
DOI: 10.7321/jscse.v1.n1.4

PCA- Based Feature Extraction and k-NN algorithm for Early Jaundice Detection

[1] *Muhammad Naufal Mansor, [2] Sazali Yaacob, [3] Hariharan Muthusamy, [4] Shafriza Nisha Basah, [5] Shahrul Hi-fi Syam bin Ahmad Jamil, [6] Mohd Lutfi Mohd Khidir, [7] Muhammad Nazri Rejab, [8] Ku Mohd Yusri Ku Ibrahim, [9] Addzrull Hi-fi Syam bin Ahmad Jamil, [10] Jamaluddin Ahmad , [11] Ahmad Kadri Junoh

[1,2,3,4,11] *Intelligent Signal Processing Group (ISP), University Malaysia Perlis, no 70 &71, Blok B, Jalan Kangar- Alor setar, 02000, Kangar, Perlis, Malaysia., apairia@yahoo.com, sazali22@yahoo.com, wavelet.hari@gmail.com, shafriza@gmail.com, kadri@unimap.edu.my*

[5,6,7,8,9] *Fakulti Kejuruteraan Elekterik, Politeknik Tuanku Syed Sirajuddin, Km. 25, Lebuhraya Kuala Perlis Changlun, Ulu Pauh, 02600, Arau, Perlis., syahrull30@yahoo.com, mdlutfi@gmail.com, nazri_554@yahoo.com, kmyusri82@gmail.com, azrulhifisyam@yahoo.com*

[10] *Pediatric Department, Hospital Tuanku Fauziah, Jalan Kolam,01000, Kangar, drjamaluddin@pls.moh.gov.my*

Abstract. **Jaundice is a yellow discoloration of the skin and/or whites of the eyes that are often seen in newborn infants. The discoloration is caused by a yellow substance called bilirubin. Infants with high blood levels of bilirubin, called hyperbilirubinemia, develop the yellow color when bilirubin accumulates in the skin. The main symptom of jaundice is yellow coloring of the skin and conjunctiva of the eyes. Jaundice can also make babies sleepy which can lead to poor feeding. Poor feeding can make jaundice worse as the baby can become dehydrated. If a baby has conjugated jaundice, it may have white chalky stool (poo) and urine that is darker than normal. A PCA method was employed to study the behavior of the infant. The experimental results reveal that the proposed method can minimize the morbidity and mortality than the conventional method based on k-NN Algorithm.**

Keywords: *Jaundice Infant, Image Processing, PCA, k-NN*

1. Introduction

Infants whose palms and soles are yellow, have serum bilirubin level over 255 umol/l (15 mg/dL) (more serious level). Studies have shown that trained examiners assessment of levels of jaundice show moderate agreement with Icterometer bilirubin measurements [1]. In infants jaundice can be measured using invasive or non-invasive methods. In non invasive method Ingram Icterometer and Transcutaneous Bilirubinometer are used. However, those are only the methods to measure the level of jaundice. The methods to check the symptom are done manually.

In babies whose bilirubin blood levels reach hazardous levels, bilirubin may cross to the brain and cause reversible damage (called early acute bilirubin encephalopathy) or permanent damage (called kernicterus). Frequent monitoring and early treatment of infants at high risk for jaundice can help to prevent severe hyperbilirubinemia. Thus, the goals of this paper are to promote a new first aid of jaundice newborn monitoring. A PCA method was employed to study the behavior of the infant. The experimental results reveal that the proposed method can minimize the morbidity and mortality than the conventional method based on k-NN algorithm.

International Journal of
Soft Computing And Software Engineering (JSCSE)
ISSN: 2251-7545
Vol.1,No.1, 2011
DOI: 10.7321/jscse.v1.n1.4

2. System Overview

The long-term goal of this research is the development of a stand-alone automated system that could be used as a supplement in the NICU to provide 24-h/day noninvasive. The development of a jaundice-detection system involves the following three tasks: (a) the extraction of PCA feature information from video recordings of infants monitored for jaundice, (b) the selection of quantitative features that convey some unique behavioral characteristics of neonatal jaundice such as mean, standard deviation, skew ness, kurtosis, energy and entropy and (c) validation test to distinguish between jaundice and normal newborn infant. The complete block diagram is shown in Fig.1.

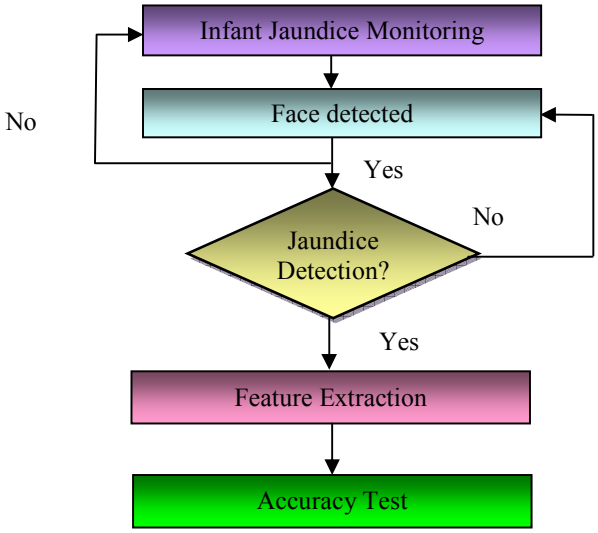

Figure 1. Early Jaundice Detection Overview System

3. Newborn under Phototherapy

This study relied on images selected from a random database developed by the http://www.google.infant monitoring.com. All the acquire images are differ from the illumination level, distance and the angle of the picture taken. This variety of images gives the further experiment more interesting in order to find a robust portion to monitor the infant jaundice. Infants with neonatal jaundice are treated with colored light called phototherapy is shown in Fig.2. The data collection computer configured with MATLAB R2009a and the Image Acquisition Toolbox

International Journal of
Soft Computing And Software Engineering (JSCSE)
ISSN: 2251-7545
Vol.1,No.1, 2011
DOI: 10.7321/jscse.v1.n1.4

Figure 2. Newborn are treated with colored light called phototherapy

4. Face Detection

The Intel Open CV cascade Classifier is based on a face detection algorithm developed by [2-3]. The Algorithm scans an image and returns a set of locations that are believed to be faces. The algorithm uses an Ada-Boost based Classifier that aggressively prunes the search space to quickly locate faces in the image. Figure 3 show an example of the Haar Classifier applied to an image, where the squares indicate face detections

The accuracy of this algorithm can be evaluated using a standard configuration that ships with the OpenCV source code. The accuracy is tested by detecting a set of faces in the Infant Face Dataset, which is designed for testing face detectors.

Because the Haar Classifier is part of the OpenCV library, the source code should be well optimized. OpenCV is a popular open source image library that was originally created by Intel. The data structures and algorithms in OpenCV have been carefully tuned to run efficiently on modern hardware. It is evident from the code of the Haar Classifier that hand optimization was used to improve the performance of that algorithm

Figure3. Face Detection results

5. Principal Component Analysis

Principal Component Analysis (PCA) [4] is a standard technique for dimensionality reduction and has been applied to a broad class of computer vision problems, including feature selection (*e.g.,* [5]), object recognition (*e.g.,* [6]) and face recognition (*e.g.,* [7]). While PCA suffers from a number of shortcomings [8, 9], such as its implicit assumption of Gaussian distributions and its restriction to orthogonal linear combinations, it remains popular due to its simplicity. The idea of applying PCA to image patches is not novel (*e.g.,* [10]). Our contribution lies in rigorously demonstrating that PCA is well-suited to representing key point patches (once they have been transformed into a canonical scale, position and orientation), and that this representation significantly improves jaundice detection performance as discuss in Chapter 7

6. k-NN Algorithm

The K nearest neighbor (kNN) classifier is an extension of the simple nearest neighbor (NN) classifier system. The nearest neighbor classifier works based on a simple nonparametric decision. Each query image I_q is examined based on the distance of its features from the features of other images in the training database. The nearest neighbor is the image which has the minimum distance from the

International Journal of
Soft Computing And Software Engineering (JSCSE)
ISSN: 2251-7545
Vol.1,No.1, 2011
DOI: 10.7321/jscse.v1.n1.4

query image in the feature space. The distance between two features can be measured based on one of the distance functions such as, city block distance d_1, and Euclidean distance d_2 or cosine distance d_{cos}

$$d_1(x,y) = \sum_{i=1}^{N} |x_i - y_i| \quad (1)$$

$$d_2(x,y) = \sqrt{\sum_{i=1}^{N} |x_i - y_i|} \quad (2)$$

$$d_1(x,y) = 1 - \frac{\vec{x}.\vec{y}}{|x|.|y|} \quad (3)$$

K nearest neighbor algorithm uses K closest samples to the query image. Each of these samples belongs to a known class Ci. The query image Iq is categorized to the class CM which has the majority of occurrences among the K samples. The performance of the kNN classifiers highly related to value of the k, the number of the samples and their topological distribution over the feature space. Many approaches are introduced to improve the performance of the kNN systems using wavelet techniques [11], Cluster-Based Trees [12] and Tolerant rough sets [13] and so on. In this paper we show ensemble PCA based techniques can be used to improve the performance of the system.

7. Results and Discussion

The proposed algorithm was evaluated on a hundred and twenty subjects with different race, gender and age. The average size of each image is 400-500 pixels. The entire subjects were tested for ten trials. Ten images were taken for each subject. Average accuracy for all subjects was shown in Table1.

Trial	*Accuracy Detection %*
1	90.12
2	91.23
3	92.22
4	92.33
5	92.45
6	92.45
7	93.44
8	94.55
9	95.65
10	

Table 1. Accuracy Detection Based on PCA and k-NN Algorithm

The confidence values for the recognition of a person is calculated using the Euclidean distance between the PCA projected values of the test image and PCA projected values of the train database. This value determines whether recognition of a face image using this method is dependable or not. When the confidence value is low recognition is not dependable. The confidence value obtained for different test images are tabulated in Table I.

8. Conclusion

DOI: 10.7321/jscse.v1.n1.4

The proposed jaundice detection algorithm localizes the face from the given input image using the Haar Classifier method where employed. The detected face image is projected using Eigen face analysis and classified using the K nearest neighborhood (KNN) classifier. This algorithm is efficient as it can be integrated with the output from multi-modal sensors and thus can be used as part of multi-sensor data fusion

9. References

[1] Madlon-Kay, Diane J. Recognition of the Presence and Severity of Newborn Jaundice by Parents, Nurses, Physicians, and Icterometer Pediatrics: pp 100-103, 1997.
[2] P.Viola and M.J. Jones, "Robust real time face detection,"Int. J. Comput. Vision, vol.57, no.2, pp.137–154, 2004.
[3] R. Lien hart and J. Maydt, "An extended set of haar-like features for rapid object detection, "in Proceedings of the 2002 International Conference on Image Processing (P.of the 2002 International Conference on Image Processing, ed),vol.1,pp. 900–903, 2002
[4] I. T. Joliffe. *Principal Component Analysis*. Springer-Verlag, 1986.
[5] K. Fukunaga and W. Koontz. Application of the Karhunen-Loeve expansion to feature selection and ordering. *IEEE Trans. Communications*, 19(4), 1970.
[6] H. Murase and S. Nayar. Detection of 3D objects in cluttered scenes using hierarchical eigenspace. *Pattern Recognition Letters*, 18(4), April 1997.
[7] M. Turk and A. Pentland. Face recognition using eigenfaces. In *Proceedings of Computer Vision and Pattern Recognition*, 1991.
[8] J. Karhunen and J. Joutsensalo. Generalization of principal component analysis, optimization problems and neural networks. *Neural Networks*, 8(4), 1995.
[9] D. Lee and S. Seung. Learning the parts of objects by non-negative matrix factorization. *Nature*, 401, 1999.
[10] R. Fergus, P. Perona, and A. Zisserman. Object class recognition by unsupervised scale-invariant learning. In *Proceedings of Computer Vision and Pattern Recognition*, June 2003.
[11] Hui Kong, Xuchun Li, Jian-Gang Wang and Chandra Kambhamettu, "Ensemble Learning for Face Recognition". The International Conference on Biometrics Authentification (ICBA), Hong Kong, 2006
[12]Yu-Long Qiao; Jeng-Shyang Pan; Sheng- He Sun, "Improved K nearest neighbor classification algorithm," Circuits and Systems, 2004. Proceedings. The 2004 IEEE Asia-Pacific Conference on , vol.2, no.pp. 1101- 1104 vol.2, 6-9 Dec. 2004
[13]Bin Zhang, Sargur N. Srihari, "Fast k- Nearest Neighbor Classification Using Cluster-Based Trees," IEEE Transactions on Pattern Analysis and Machine Intelligence ,vol. 26, no. 4, pp. 525-528, April, 2004.