Face Recognition under Variant Illumination Using PCA and Wavelets

Mong-Shu Lee*, Mu-Yen Chen, and Fu-Sen Lin

Department of Computer Science and Engineering,
National Taiwan Ocean University, Keelung, Taiwan
Tel.: 886-2-2462-2192; Fax: 886-2-2462-3249
{mslee,chenmy,fslin}@mail.ntou.edu.tw

Abstract. In this paper, an efficient wavelet subband representation method is proposed for face identification under varying illumination. In our presented method, prior to the traditional principal component analysis (PCA), we use wavelet transform to decompose the image into different frequency subbands, and a low-frequency subband with three secondary high-frequency subbands are used for PCA representations. Our aim is to compensate for the traditional wavelet-based methods by only selecting the most discriminating subband and neglecting the scattered characteristic of discriminating features. The proposed algorithm has been evaluated on the Yale Face Database B. Significant performance gains are attained.

Keywords: Face recognition, Principal component analysis, Wavelet transform, Illumination.

1 Introduction

Human face recognition has become a popular area of research in computer vision recently. It is applied to various different fields such as criminal identification, human-machine interaction, and scene surveillance. However, variable illumination is one of the most challenging problems with face recognition, due to variations in light conditions in practical applications. Of the existing face recognition methods, the principal component analysis (PCA) method takes all the pixels in the entire face image as a signal, and proceeds to extract a set of the most representative projection vectors (feature vectors) from the original samples for classification. First, Turk and Pentland [15] extracted noncorrelational features between face objects by PCA, and applied the neighborhood algorithm classification method to face recognition. Yet, the variations between the images of the same face due to illumination and view direction are always larger than the image variations due to a change in face identity [1]. Standard PCA-based methods cannot facilitate division of classes as feature vectors obtained from face image under varying lighting conditions. Hence, if only one upright frontal image per person, which is under severe light variations, is available for training, the performance of PCA will be seriously degraded.

* Corresponding author.

A.-B. Salberg, J.Y. Hardeberg, and R. Jenssen (Eds.): SCIA 2009, LNCS 5575, pp. 341–350, 2009.

Many methods have been presented to deal with the illumination problem. The first approach to handling the effect that results from illumination changes is constructing illumination model from several images acquired under different illumination condition. The representative method, the illumination cone model that can deal with shadow and multiple lighting sources, is introduced by [2, 10]. Although this approach achieved 100% recognition rates, it is not practical to require seven images of each person to obtain the shape and albedo of a face. Zhao and Chellappa [19] developed a shape-based face recognition system by means of an illumination-independent ratio image derived by applying a symmetrical shape-from-shading technique to face images. Shashua and Riklin-Raviv [14] used quotient images to solve the problem of class-based recognition and image synthesis under varying illumination. Xie and Lam [16] adopted a local normalization (LN) technique to images, which can effectively eliminate the effect of uneven illumination. Then the generated images with illumination variation insensitivity property are used for face recognition using different methods, such as PCA, ICA and Gabor wavelets. Discrete Wavelet transform (DWT) has been used successfully in image processing. An advantage of DWT is that with few wavelet coefficients it can capture most of the image energy and the image features. In addition, its ability to characterize local spatial-frequency information of image motivates us to use it for feature extraction. In [9], three-level wavelet transform is performed to decompose the original image into its subbands on which the PCA is applied. The experiments on Yale database show that the third level diagonal details attain the highest correct recognition rate. Later, wavelet face [4] only uses the low-frequency subbands to present the basic figure of an image, and ignore the efficacy of high-frequency subbands. Ekenel and Sankur [7] came up with a fusing scheme by collecting the information coming from the subbands that attain individually high correct recognition rates to improve the classification performance.

Although some studies have been conducted on the discriminatory potential of single frequency subband in DWT, little research has been done on the counterparts of the combination of frequency subbands. In this study, we propose a novel method to handle the problem of face recognition with varying illumination. In our approach, DWT is adopted first to decompose an image into different frequency components. To avoid neglecting the image features resulting from different lighting condition, a low-frequency and three midrange frequency subbands are selected for PCA representation. In the last step of the classification rule, it is the weighting combination of the individual discriminatory potential, applied to the PCA-based face recognition procedure. Experimental results demonstrated that applying PCA on four different DWT subbands, and then merging distinct subbands information with relative weights in classification achieve a rather excellent recognition performance.

2 Wavelet Transform and PCA

2.1 Multi-resolution Property of Wavelet Transform

Over the last decade or so, the wavelet transform (WT) has been successfully adopted to solve various problems of signal and image processing. The wavelet transform is fast, local in the time and the frequency domain, and provides multi-resolution analysis of real-world signals and images. Wavelets are collections of functions in L^2 constructed from a basic wavelet ψ using dilations and translations. Here we will only consider the families of wavelets using dilations by powers of 2 and integer translations:

$$\psi_{j,k}(x) = 2^{\frac{j}{2}} \psi(2^j x - k),\ j,k \in Z .$$

We can see that the time and frequency localization of the wavelet basis functions are adjusted by both scale index j and position index k.

Multi-resolution Analysis is generally an important method for constructing orthonormal wavelet bases for L^2. In multi-resolution schemes, wavelets have corresponding scaling function φ, whose analogously defined dilations and translation $\varphi_{j,k}(x)$ span a nested sequence of multi-resolution space V_j, $j \in$ Z. Wavelets $\{\psi_{j,k}(x) : j,k \in Z\}$ form orthonormal bases for the orthogonal complements $W_j = V_j - V_{j-1}$ and for all of L^2. Therefore, the wavelet transform decomposes a function into a set of orthogonal components describing the signal variations across scales [5]. For one-dimensional wavelet transform, a signal f, is represented by its wavelet expansion as:

$$f(x) = \sum_{k \in Z} c_I(k) \varphi_{I,k}(x) + \sum_{j \ge I} \sum_{k \in Z} d_j(k) \psi_{j,k}(x), \tag{1}$$

where the expansion coefficients $c_I(k)$ and $d_j(k)$ in (1) are obtained by an inner product, for example:

$$d_j(k) = < f, \psi_{j,k} > = \int f(x) 2^{\frac{j}{2}} \psi(2^j x - k) dx .$$

In practice, we usually apply the DWT algorithm corresponding to (1) with finite decomposition levels to obtain the coefficients. Here, the wavelet coefficients of a 1-D signal is calculated by splitting it into two parts, with a low-pass filter (corresponding to the scaling function ϕ) and high-pass filter (corresponding to the wavelet function ψ), respectively. The low frequency part is split again into two parts of high and low frequencies, and the original signal can be reconstructed from the DWT coefficients.

The two-dimensional DWT is performed by consecutively applying one-dimensional DWT to the rows and columns of the two-dimensional data. Two-dimensional DWT decomposes an image into “subbands” that are localized in time and frequency domains. The DWT is created by passing the image through a series of filter bank stages. The high-pass filter and low-pass filter are finite impulse response filters. In other words, the output at each point depends only on a finite portion of the input image. The filtered outputs are then sub-sampled by 2 in the row direction.

These signals are then each filtered by the same filter pair in the column direction. As a result, we have a decomposition of the image into 4 subbands denoted HH, HL, LH, and LL. Each of these subbands can be regarded as a smaller version of the image representing different image contents. The Low-Low (LL) frequency subband preserves the basic content of the image (coarse approximation) and the other three high frequency subbands HH, HL, and LH characterize image variations along diagonal, vertical, and horizontal directions, respectively. Second level decomposition can then be conducted on the LL subband. Such iteration process is continued until the specified number of desired decomposition level is achieved. The multi-resolution decomposition strategy is very useful for the effective feature extraction. Fig. 1 shows the subbands of three-level discrete wavelet decomposition. Fig. 2 displays an example of image Box with its corresponding subbands LL_3, LH_3, HL_3 and HH_3 in Fig. 1.

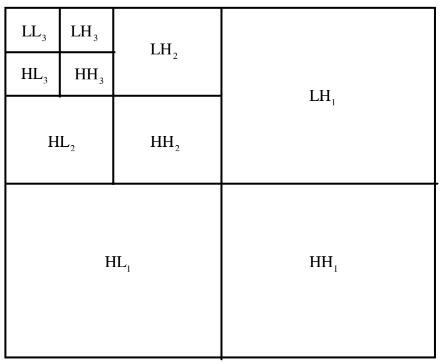

Fig. 1. Different frequency subbands of a three-level DWT

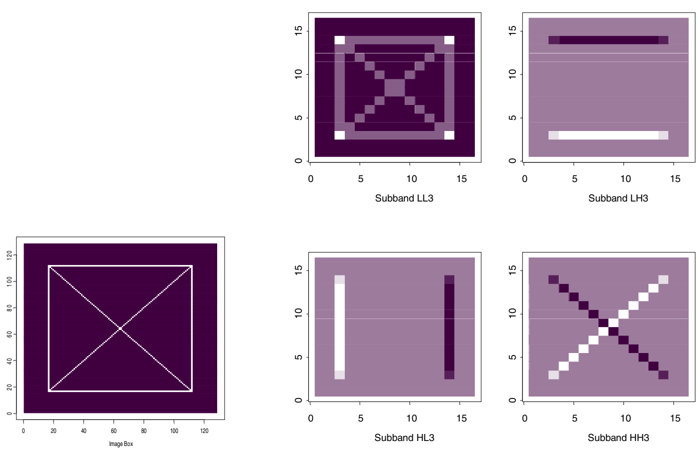

Fig. 2. Original image Box (left) and its subbands of LL_3, LH_3, HL_3 and HH_3 in a three-level DWT

2.2 PCA and Face Eigenspace

Principal component analysis (PCA) is a dimensionality reduction technique based on extracting the desired number of principal components of the multidimensional data. Given an $N-$dimensional vector representation of each face in a training set of M images, PCA tends to find a $t-$dimensional subspace whose basis vectors correspond to the maximum variance direction in the original image space. This new subspace is normally a smaller dimension $(t << N)$. These new basis vectors can be calculated in the following way. Let X be the $N \times M$ data matrix whose columns $x_1, x_2, ..., x_M$ are observations of a signal embedded in $\mathbb{R}^N$; in the context of face recognition, M is the available training images, and $N = m \times n$ is the number of pixels in an image. The PCA basis Ψ is obtained by solving the eigenvalue problem $\Lambda = \Psi^T \mathrm{E} \Psi$, where E is the covariance matrix of the data

$$\mathrm{E} = \frac{1}{M}\sum_{i=1}^{M}(x_i - \overline{x})(x_i - \overline{x})^T, \text{ where } \overline{x} \text{ is the mean of } x_i.$$

$\Psi = [\psi_1, ..., \psi_m]^T$ is the eigenvector matrix of E, and Λ is the diagonal matrix with eigenvalues $\lambda_1 \geq .... \geq \lambda_N$ of E on its main diagonal, so ψ_j is the eigenvector corresponding to the jth largest eigenvalue. Thus, to perform PCA and extract t principal components of the data, one must project the data onto Ψ_t, the first t columns of the PCA basis Ψ, which correspond to the t highest eigenvalue of E. This can be regarded as a linear projection $\mathbb{R}^N \rightarrow \mathbb{R}^t$, which retains the maximum energy (i.e., variance) of the signal. This new subspace $\mathbb{R}^t$ defines a subspace of face images called face space. Since the basis vectors constructed by PCA had the same dimension as the input face images, they are named "eigenfaces" by Turk and Pentland [15].

Combined with the effectiveness of capturing image features of DWT and the accuracy of data representation of PCA, we are motivated to develop an efficient scheme for the face recognition in the next section.

3 The Proposed Method

The study is aimed to enhance the recognition rate of the face image under varying lighting conditions by the standard PCA-based methods. In the literature, the DWT was applied in texture classification [3] and image compression [6] due to its powerful capability in multi-resolution decomposition analysis. The wavelet decomposition technique was also used to extract the intrinsic features for face recognition [8]. In [11], a 2D Gabor wavelet representation was sampled on the grid and combined into a labeled graph vector for elastic graph matching of face images. Similar to [9], we apply the multilevel two-dimensional DWT to extract the facial features. In order to reduce the effect of illumination, the pre-processing of training and unknown images may choose to employ histogram equalization before taking DWT.

The whole block diagram of the face recognition system including training stage and recognition stage is as in Fig. 3. A three-level DWT, using the Daubechies' S8 wavelet, is applied to decompose the training image, as illustrated in Fig. 1. Generally, the low frequency subband LL_3 represents and preserves the coarser approximation of an image, and the other three sub-high frequency subbands characterize the details of the image texture in three different directions. Earlier studies concluded that the information in the low spatial frequency bands play a dominant role in face recognition. Naster *et al.* [13] have found that facial expression and small occlusions affect the intensity manifold locally. Under frequency-based representation, only the high frequency spectrum is affected. Moreover, changes in illumination affect the intensity manifold globally, in which only the low frequency spectrum is affected. When there is a change in human face, all frequency components will be affected. Based on these observations, we select the HH_3, LH_3, HL_3 and LL_3 subbands in the third level to employ the PCA procedure in this study. All these frequency components have played their parts with different weights in discriminating face identity.

In the recognition step, distance measurement between the unknown image and the training images in the library is performed to determine whether the input of an

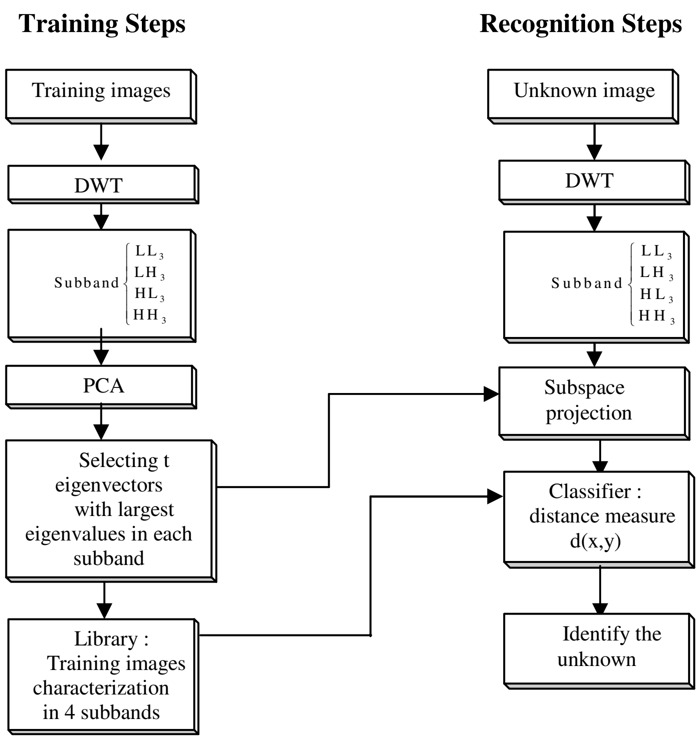

Fig. 3. Block diagram of the proposed recognition system

unknown image matches any of the images in the library. In terms of classifying the criterion, the traditional Euclidean distance cannot measure the similarity very well when there illumination variations on the facial images exist. Yambor [17] reported that a standard PCA classifier performed better when the Mahalanobis distance was used. Therefore, the Mahalanobis distance is also selected as the distance measure in the recognition step of our experiments. The Mahalanobis distance is formally defined in [12], and Yambor [17] gives a simplification, which is adopted here as follows:

$$d_{Mah}(x, y) = -\sum_{i=1}^{t} \frac{1}{\sqrt{\lambda_i}} x_i y_i$$

where x and y are the two face images to be compared and λ_i is the ith eigenvalue corresponding to the ith eigenvector of the covariance matrix E.

Finally, the distance between the unknown image and the training image is a linear combination of their discriminating ability of four wavelet subbands, and is defined as follows:

$$\begin{aligned} d(x, y) = 0.4 d_{Mah}^{HH_3}(x, y) + 0.3 d_{Mah}^{LH_3}(x, y) \\ + 0.2 d_{Mah}^{HL_3}(x, y) + 0.1 d_{Mah}^{LL_3}(x, y) \end{aligned} \quad (2)$$

where $d_{Mah}^{HH_3}(x, y)$, $d_{Mah}^{LH_3}(x, y)$, $d_{Mah}^{HL_3}(x, y)$ and $d_{Mah}^{LL_3}(x, y)$ are the Mahalanobis distance measured on the subbands of HH_3, LH_3, HL_3, and LL_3 respectively. The weighting coefficients put in front of each subband in equation (2) were selected on the basis of their recognition performance in the single-band experiment with Subset 3 images of Yale Face Database B. The average recognition accuracy of the four different subbands using Subset 3 images (with and without histogram equalization) is recorded in Table 1. It can be shown that the HH_3 subband gives the best result, and thus the weighting coefficient of subband HH_3 deserves the largest value 0.4 in the classifier equation (2). The weighting coefficients of the other three subbands LH_3, HL_3, and LL_3 are in decreasing order according to their decline in average recognition rate in Table 1.

Table 1. The average recognition performance (with and without histogram equalization) using Subset 3 images of Yale Face Database B on different DWT subbands

DWT Subband	Average recognition accuracy
HH_3	89.2%
LH_3	86.4%
HL_3	81.4%
LL_3	78.6%
Average	83.9%

4 Experimental Results

The performance of our algorithm is evaluated using the popular Yale Face Database B that contains images of 10 persons under 45 different lighting conditions, and the test is performed on all of the 450 images. All the face images are cropped and normalized to a size of 128x128. The images of this database are divided into four subsets according to the lighting angle between the direction of the light source and the camera axis. The first subset (Subset 1) covers the angular range up to 12°, the second subset (Subset 2) covers 12° to 25°, the third subset (Subset 3) covers 25° to 50°, and the fourth subset (Subset 4) covers 50° to 75°. One example images of these four subsets are illustrated in Fig. 4.

For each individual in the Subset 1 and 2, two of their images were used for training (total 20 training images for each set), and the remaining images were used for testing. As a method to overcome left and right face illumination variation that appeared in Subset 3 and Subset 4, we computed the difference between the average pixel value of the left and right face, where the left and right face were divided on the vertical-axis center of the input image. We selected two images with the left and right face difference greater than the threshold value 30 (experimental value) per person from Subset 3 and Subset 4 to form the training image set, and the rest of the images

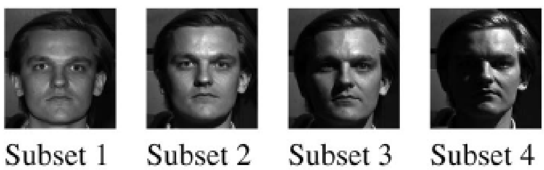

Fig. 4. Sample images of one individual in the Yale Face Database B under the four subsets of lighting

Table 2. Comparison of recognition methods with Yale Face Database B

(The entries with indicated citation were taken from published papers)

Method	Similarity measure	Size of training sample	The number of eigenfaces	Recognition rate
WT(Fusing six subbands into one-single band) +PCA [7]	Correlation coefficient	2	80	77.1%
WT(subband HH3) + PCA [9]	Correlation coefficient	2	11	84.5%
The proposed method	Mahalanobis distance	2	36	99.3%
LN(local normalization) +HE + PCA [16]	Mahalanobis distance	1	200	99.7%

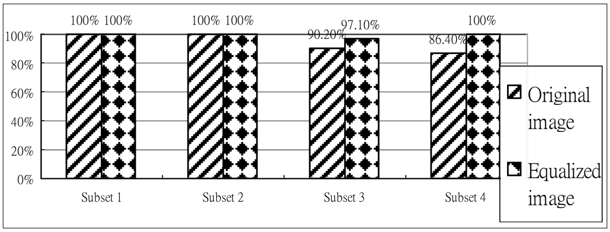

Fig. 5. The recognition performance of the algorithm when applied to the Yale Face Database B

were used as test images. The proposed method was tested on the image database as follows: the existing PCA with the first two eigenvectors excluded, and PCA with histogram equalized images. Fig. 5 tabulates the recognition rates using the images on the database and PCA approaches, where nine eigenvectors in each subbands (total 36 eigenvectors) calculated from the training images were used for face recognition. The result of the PCA application to original images on Subset 1, 2, 3 and 4 with the first two eigenvectors excluded shows high recognition performance of 100%, 100%, 90.2% and 86.4% respectively. Moreover, the result of the PCA application after histogram equalization (HE) on Subset 1, 2, 3 and 4 was recognition performance of 100%, 100%, 97.1% and 100% respectively (with average 99.3%). The PCA-based recognition performance may be influenced by several factors, such as the size of training sample, the number of eigenfaces, and similarity measure. Under similar influence factors, we compare the performance between the proposed method and other PCA-based face recognition methods in Table 2. The local normalization (LN) approach achieved the highest recognition rate 99.7% in Table 2, but they use 200 eigenfaces. Obviously, our recognition rate is comparable to the LN approach and significantly improves the traditional PCA-based face recognition methods.

5 Conclusions

In this study, a novel wavelet-based PCA method for human face recognition under varying lighting condition is proposed. The advantages of our method are summarized as follows:

1. Wavelet PCA offers a method through which we can improve the performance of normal PCA by using low frequency and sub-high frequency components, which lowers the computation cost while keeping the essential feature information needed for face recognition.
2. We carefully design the classification rule, which is a linear combination of four subband contents according to their individual recognition rates in a single-band test. Therefore, the weights for each subband used in the distance function are highly meaningful.

The experimental result shows that the proposed method demonstrates very efficient performance with the histogram-equalized images. The future work includes the evaluation of the other image data with illumination variation, such as CMU PIE database.

References

1. Adini, Y., Moses, Y., Ullman, S.: Face recognition: The problem of compensating for changes in illumination direction. IEEE Transaction on Pattern Analysis and Machine Intelligence 19, 721–732 (1997)
2. Belhumeur, P.N.: Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection. IEEE Transaction on Pattern Analysis and Machine Intelligence 19, 711 (1997)
3. Chang, T., Kuo, C.: Texture analysis and classification with tree-structured wavelet transform. IEEE Tran. on Image Processing 2, 429 (1993)
4. Chien, J.T., Wu, C.C.: Discriminant waveletfaces and nearest feature classifiers for face recognition. IEEE Transaction on Pattern Analysis and Machine Intelligence 24(12), 1644–1649 (2002)
5. Daubechies, I.: Ten Lectures on Wavelets. In: SIAM. CBMB Regional Conference in Applied Mathematics Series, vol. 61 (1993)
6. DeVore, R., Jawerth, B., Lucier, B.: Image compression through wavelet transform coding. IEEE Trans. on Information Theory 38, 719–746 (1992)
7. Ekenel, H.K., Sanker, B.: Multiresolution face recognition. Image and Vision Computing (23), 469–477 (2005)
8. Etemad, K., Chellappa, R.: Face recognition using Discreminant eigenvectors. In: Proceeding IEEE Int'l. Conf. Acoustic, Speech, and Signal Processing, pp. 2148–2151 (1996)
9. Feng, G.C., Yuen, P.C.: Human face recognition using PCA on wavelet subband. Journal of Eectronic Imaging (9), 226–233 (2000)
10. Georghiades, A., Kriegman, D., Belhumeur, P.: Illumination cones for recognition under variable lighting: faces. In: Proceeding IEEE C CVPR SANT B (1998)
11. Lyons, M.J., Budynek, J., Akamatsu, S.: Automatic classification of single facial image. IEEE Transaction on Pattern Analysis and Machine Intelligence 21(12), 1357–1362 (1999)
12. Moon, H., Phillips, J.: Analysis of PCA-based face recognition algorithms. In: Boyer, K., Phillips, J. (eds.) Empirical Evaluation Methods in Computer Vision. World Scientific Press, MD (1998)
13. Nastar, C., Ayach, N.: Frequency-based nongrid motion analysis. IEEE Transaction on Pattern Analysis and Machine Intelligence 18, 1067–1079 (1996)
14. Shashua, A.: The quotient image: Class-based re-rendering and recognition with varying illuminations. IEEE Transaction on Pattern Analysis and Machine Intelligence 23(2), 129–139 (2001)
15. Turk, M., Pentland, A.: EIigenfaces for Recognition. Journal of Cognitive Neuroscience 3, 71 (1991)
16. Xie, X., Lam, K.: An efficient illumination normalization method for face recognition. Pattern Recognition Letters 27(6), 609–617 (2006)
17. Yambor, W., Draper, B., Beveridge, R.: Analyzing PCA-based face recognition algorithms: eigenvector selection and distance measures. In: Christensen, H., Phillips, J. (eds.) Empirical Evaluation Methods in Computer Vision. World Scientific Press, Singapore (2002)
18. Zhao, J., Su, Y., Wang, D., Luo, S.: Illumination ratio image: synthesizing and recognition with varying illuminations. Pattern Recognition Letters (24) (2003)
19. Zhao, J., Chellappa, R.: Illumination-insensitive face recognition using symmetric shape-from-shading. In: Proceeding IEEE conf. CVPR Hilton Head (2000)