UNIVERSIDADE DE SÃO PAULO

Instituto de Ciências Matemáticas e de Computação

Techniques for Image Compression: a Comparative Analysis

Patricia R. Oliveira
Roseli F. Romero
Luis G. Nonato
Josmar Mazucheli

Nº 49

NOTAS

São Carlos - SP

UNIVERSIDADE DE SÃO PAULO
Instituto de Ciências Matemáticas e de Computação
ISSN 0103-2577

Techniques for Image Compression: a Comparative Analysis

Patricia R. Oliveira
Roseli F. Romero
Luis G. Nonato
Josmar Mazucheli

Nº 49

NOTAS

Série Computação

São Carlos – SP
Out./2000

Techniques for Image Compression: a Comparative Analysis

Patricia R. Oliveira, Roseli F. Romero, Luis G. Nonato
Departamento de Computação e Estatística
Instituto de Ciências Matemáticas e de Computação
Universidade de São Paulo – Campus de São Carlos
Caixa Postal 668
13560–970 São Carlos, SP
{rufino, rafrance, gnonato}@icmc.sc.usp.br

Josmar Mazucheli
DEs - UEM
Av. Colombo, 5790, Maringá - PR, Brazil
josmar@uem.br

Resumo

Nos dias atuais, compressão de imagens é uma tarefa que tem se tornado cada vez mais necessária. Algumas técnicas para compressão de imagens são investigadas neste trabalho. A primeira dessas técnicas é o padrão JPEG, que é o método mais utilizado para compressão de imagens. A segunda é a Análise de Componentes Principais (PCA), também conhecida como transformação de Karhunen-Loève, que é um método estatístico aplicado para análise de dados multivariados e extração de características. Neste trabalho, uma abordagem que utiliza o método estatístico clássico e outra baseada em redes neurais artificiais são consideradas para a realização da PCA. Em um estudo comparativo, os resultados obtidos pela rede neural PCA para a compressão de imagens médicas são analisados juntamente com outros obtidos através da utilização do método estatístico clássico e do padrão de compressão JPEG.

Abstract

Nowadays image compression is a task that has been more necessary than ever. Some techniques for image compression are investigated in this article. The first one is the well known JPEG that is the most widely used technique for image compression. The second is Principal Component Analysis (PCA), also called Karhunen-Loeve transform, that is a statistical method applied for multivariate data analysis and feature extraction. In the latter,two approaches are being considered. The first approach uses the classical statistical method and the other one, artificial neural networks. In a comparative study, the results obtained by PCA neural network for compressing medical images are analyzed together with those obtained by using the classical statistical method and the JPEG compression standard technique.

1 Introduction

In the last decades, the information technology has led to enormous amounts of data (measure in gigabytes and terabytes) being stored on computer systems. In this way, techniques to reduce the dimension of data are more necessary than ever. In this paper, we are concerned with three techniques for performing image compression. The first one is JPEG compression standard [4] which is a standard technique of comparison in judging new compression algorithms. The two other techniques are Principal Component Analysis (PCA) based methods.

PCA has been widely used as an efficient way to provide an informative, low-dimensional representation of data from which important features can be extracted. This technique can be considered as a useful tool for linear data compression in which the truncation errors are minimized. Because it is optimal in the mean-square error sense among all linear orthogonal transforms, PCA has been used in signal and image processing, data analysis and pattern recognition [5]. In the image processing domain PCA has been used in many applications such as symmetry analysis, data compression and human face recognition [3].

The first approach that can be used for performing PCA is the classical statistical method and the other one is an artificial neural network approach. In this work, we are interested in the latter, since neural network implementations [2] can overcome some problems inherent in the classical algorithms. Moreover, learning algorithms and neural networks are adaptive in nature, thus, they are well suited for many real environments where adaptive systems are required.

In this work, the performance of a PCA neural network proposed by Rubner et al [9] is evaluated. This network is used to compress some medical images and, next, the results obtained by the PCA network are compared with some results obtained by applying the classical PCA approach and JPEG standard. This paper is organized as follows. Section 2 briefly presents the JPEG standard for compression. Section 3 presents some considerations about the PCA classical statistical method. Section 4 presents two neural networks that perform PCA. In section 5, results for some medical image compression applications are shown. Finally, conclusions and future work are presented in section 6.

2 The JPEG Compression Standard

JPEG is an international standard for color image compression and it is an acronym for Joint Photographic Experts Group [7]. This group was created by two major standard organizations, the ISO (International Organization for Standardization) and the CCITT (International Telegraph and Telephone Consultative Committee).

JPEG uses a mathematical transformation known as Discrete Cosine Transform (DCT) to convert an image to the frequency domain using a quality factor of 0 to 100. The DCT is performed on a 8x8 matrix of pixel values, and it yields a 8x8 square matrix of frequency coefficients. In this new matrix, the coefficients found in row and column 0, known as DC components, carry more notable information about the image than the other coefficients.

As the elements are farther from the DC coefficient, they tend to become lower in magnitude. It means that by performing the DCT on the input data, the representation of the image has been concentrated in the upper left coefficients of the calculated matrix, with the coefficients containing less useful information in the lower right of the same matrix. These less important components can be ignored without seriously damaging the quality of the picture.

In its quantization stage, the JPEG algorithm uses a quantization matrix in which every element in the DCT matrix is represented by corresponding to quantum value. The quantum value can be related with the mathematical error found between an input and output image after its decompression or with the effect of the decompression on the human eye.

The last stage in JPEG process is the coding of the quantized image. The first step of this phase changes the DC coefficient at the position (0,0) to a relative value. The DC elements are converted as being the difference between them and the DC element of the last block. Next, the elements of the image are arranged in a zigzag sequence. After this, they are encoded using two different mechanisms.

The first is run-length encoding of zero values. The other mechanism is called Entropy Coding. This involves sending out the coefficient code using Huffman or arithmetic coding.

As many other transformations of the same class of mathematical operations, such as the well-known Fast Fourier Transform (FFT), there is an Inverse DCT (IDCT) function that converts the frequency representation of the signal back to the spatial domain.

3 Principal Component Analysis (PCA)

Principal component analysis is a linear orthogonal transform from a p-dimensional space to an m-dimensional space, $m \leq p$, so that the coordinates of the data in the new space are uncorrelated and the greatest amount of variance of the original data is kept by only a few coordinates.

Let p variables $X_1, X_2, \ldots, X_p$. The goal of this analysis is to find linear combinations of them in order to produce indices $Z_1, Z_2, ..., Z_p$ that are uncorrelated and orthogonal. These indices are ordered in such a way that Z_1 retains the largest amount of variation, Z_2 retains

the second largest amount of variation and so on. That is,

$$var(Z_1) \geq var(Z_2) \geq \ldots \geq var(Z_p)$$

where $var(Z_i)$ represents the variance of the indice Z_i in the data set being considered. These indices Z_i are called principal components.

When performing a principal component analysis, one expects that most of the indices will display so low variances as to be dismissed. If this happens, the variation in data set can be appropriately described by the few Z indices that retain the largest variances. In this case, the dimensionality of the data set can be considerably reduced.

In order to avoid that some variables have an undue influence on the analysis, it is common to code the variables $X_1, X_2, \ldots, X_p$ to have means of zero and,in some PCA applications, variances of one. This standardization of the measurements ensures that they all have equal weight in the analysis. In image compression applications, the means are removed from the variables, but the variances are not normalized to one.

In fact, a principal component analysis just involves finding the eigenvalues and the corresponding eigenvectors of the data set covariance matrix. The corresponding eigenvalues of this matrix give an indication of the amount of information the respective principal components represent.

More formally, assuming that the eigenvalues are ordered as $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_p \geq 0$, then λ_i corresponds to ith principal component

$$Z_i = a_{i1}X_1 + a_{i2}X_2 + a_{ip}X_p$$

In particular, $var(Z_i) = \lambda_i$ and the constants $a_{i1}, a_{i2}, \ldots, a_{ip}$ are the elements of the corresponding eigenvector.

Thus, transforming the original data set according the principal components is related to the maximization of the projection variance, which is equivalent to minimization of the mean square reconstruction error. In this way, PCA can be seen equivalently as either a variance maximization technique or a least-mean-square technique.

In the next section, we establish a connection between the classical PCA and neural networks.

4 PCA Neural Networks

The PCA networks presented in this section use an unsupervised learning process that is based in variations of the Hebbian learning rule.

The first neural network [10] considered in this work is a feedforward network composed of a single layer of linear neurons operating under the modified Hebbian learning rule

$$w_i(n+1) = w_i(n) + \eta y(n)\left[x_i(n) - y(n)w_i(n)\right]$$

In this case, the synaptic weight vector $w(n)$ of a self-organized neuron converges with a probability 1 to a vector of unit Euclidean length, which lies in the maximal eigenvector direction of the correlation matrix of the input vector [8].

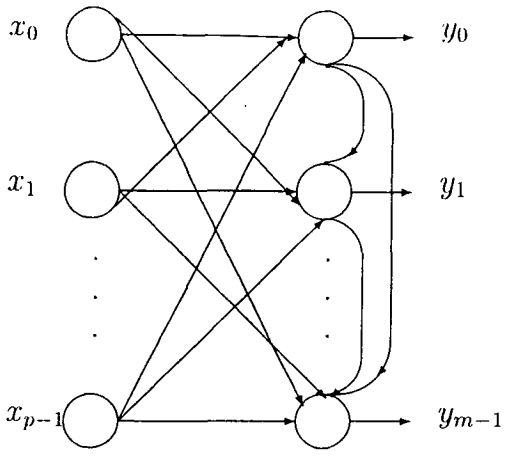

Figure 1: Adaptive PCA neural network

The output $y_j(n)$ of neuron j at time n produced in response to the set of inputs $\{x_i(n);\ i = 0, 1, \ldots, p-1\}$, is given by:

$$y_j(n) = \Sigma_{i=0}^{p-1} w_{ji}(n) x_i(n), \qquad j = 0, 1, \ldots, m-1$$

The synaptic weight $w_{ji}(n)$ is adapted using a generalized form of Hebbian learning

$$\Delta w_{ji}(n) = \eta \left[y_j(n) x_i(n) - y_j(n) \Sigma_{k=0}^{j} w_{ki}(n) y_k(n) \right]$$

where $i = 0, 1, \ldots, p-1$; $\quad j = 0, 1, \ldots, m-1$ and $\Delta w_{ji}(n)$ is the change applied to the synaptic weight $w_{ji}(n)$ at time n and η is the learning-rate parameter.

For this neural network, according to the convergence theorem proposed and proved by Sanger [10], in the limit:

$$\Delta w_j(n) \to 0 \text{ and } w_j \to q_j, \qquad j = 0, 1, \ldots, m-1$$

so that $||w_j(n)|| = 1$ for all j. The values $q_0, q_1, \ldots, q_{m-1}$ represent the normalized eigenvectors assigned to the largest m eigenvalues of the correlation matrix R of the input vector $x(n)$, ordered by decreasing values.

This neural network model has been implemented and tested for performing compression in small data sets. For these cases, we have noted that this model has slowly converged. So, it has not been used to compress medical images.

Another PCA neural network proposed by Rubner et al [9], shown in Figure 1, includes in its structure anti-Hebbian feedback connections [1].

The network output units are organized in such a way that the output unit i is connected to the output unit j with connection strength u_{ij} if and only if $i < j$.

The output $y_j(n)$ of neuron j at time n produced in response to the set of inputs $\{x_i(n)\ ;\ i = 0, 1, \ldots, p\text{-}1\}$, is given by:

$$y_j(n) = \Sigma_{i=0}^{p-1} w_{ij}(n) x_i(n) + \Sigma_{l<j} u_{lj}(n) y_l(n),$$

$$j = 0, 1, \ldots, m-1.$$

The synaptic weights w_{ij} between the input layer and the output layer are updated in accordance to the Hebbian learning rule, that is,

$$\Delta w_{ij}(n) = \eta\, x_i(n) y_j(n),$$

$$i = 0, 1, \ldots, p-1; \qquad j = 0, 1, \ldots, m-1$$

where η is the learning-rate parameter.

The lateral synaptic weights u_{lj} are adjusted according to the anti-Hebbian learning rule

$$\Delta u_{lj}(n) = -\mu\, y_l(n) y_j(n), \qquad l < j$$

where μ is another positive learning parameter.

Mao and Jain [6] found that introducing the momentum term and letting the learning rate and momentum decay with time can speed up the convergence. So, one can use

$$\Delta w_{ij}(n+1) = \eta(n) x_i y_j + \beta(n) \Delta w_{ij}(n)$$

and

$$\Delta u_{lj}(n+1) = -\mu(n) y_l y_j + \beta(n) \Delta u_{lj}(n)$$

where $\eta(n+1) = max(\alpha\eta(n), 0.0001)$, $\mu(n+1) = max(\alpha\mu(n), 0.0002)$, $\beta(n+1) = max(\alpha\beta(n), 0.0001)$ and α is the decay factor.

Because this neural network model includes such improvements in its structure and learning rules in comparison to the network presented previously, it has been chosen to be used in the experiment of this work.

5 Discussion of the Results

For evaluating the performance of the neural network chosen for performing PCA, a large medical image set has been used for compression task. The data are 480 x 640 pixel grayscale images taken from human liver by means of a laser microscope.

In our experiment, blocks with different dimensions, such as, blocks containing 8x8, 16x16 and 32x32 pixels were taken from original image to be compressed. However, after several tests it was observed that 32x32 blocks provided a high compression rate indeed a good performance. More especifically, each image was divided in 32x32 blocks in which each column represents a certain attribute of the image block. In this case, the original variables that represent a image block were considered as being $X_1, X_2, ..., X_{32}$, where X is the variable representing the pixel values in the *ith* column in the block. Then our objective was findind the first few principal components that could be enough to represent the most important information of each block, reducing the amount of bits necessary to store the image set.

To evaluate how many principal components would be used in the experiments, the classical PCA method was applied for compressing one block taken from an image in the

considered set. This block was successively transformed on basis of 1,2....,32 principal components, yielding 32x1, 32x2,..., 32x32 blocks, respectively. Next, the reconstruction step and the calculation of the mean square error (MSE) between the original block and its recovered block were performed in order to observe the lossy effect caused by the compression process.

Analysing the results, showed in Figure 2, it was found that transforming each block on basis on 1, 2 and 3 principal components, corresponding to compression ratios of 96.875%, 93.75% and 90.625%, it would be enough to construct a low-dimensional representation of the whole image while keeping its most relevant information.

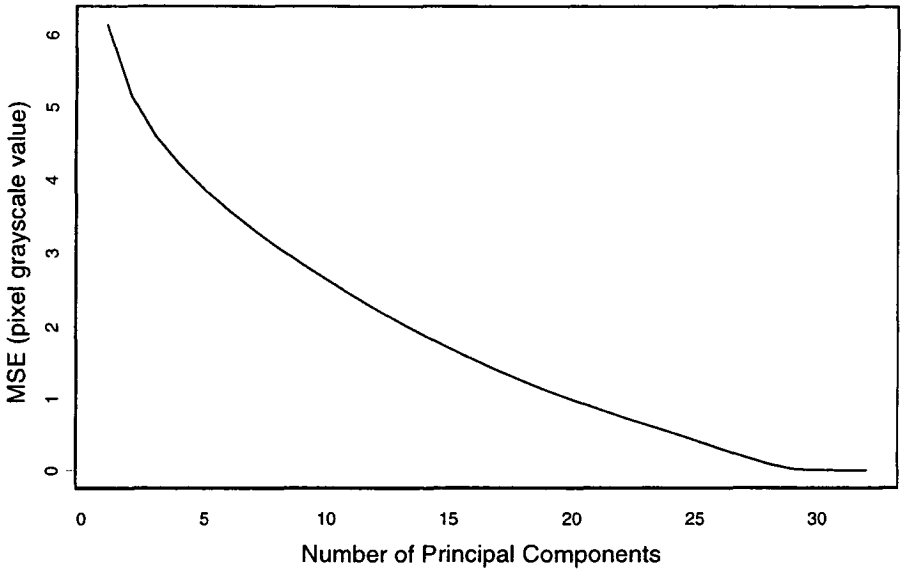

Figure 2: Results for compressing one image block

In the next step, both classical PCA and PCA neural network were applied to compress all the images in the set, according to the number of principal components that was chosen in the previous stage. Results obtained with these experiments were compared with other results obtained by applying the JPEG standard for compressing the same images. Table 1 shows a comparison among the three approaches considering 10 images taken from the whole set. The compression ratios for JPEG are not exactly equal to values adopted for PCA approach, because in the latest method the compression ratio is limited by the number of principal components that has been considered.

From Table 1, it can be viewed that PCA neural network really approximates the classical statistical analysis in terms of MSE. However, as the PCA neural network has the ability to learn, it was expected that the time necessary for the compression decrease as the number of images increase. So, this was tested for this particular set of images and the results can

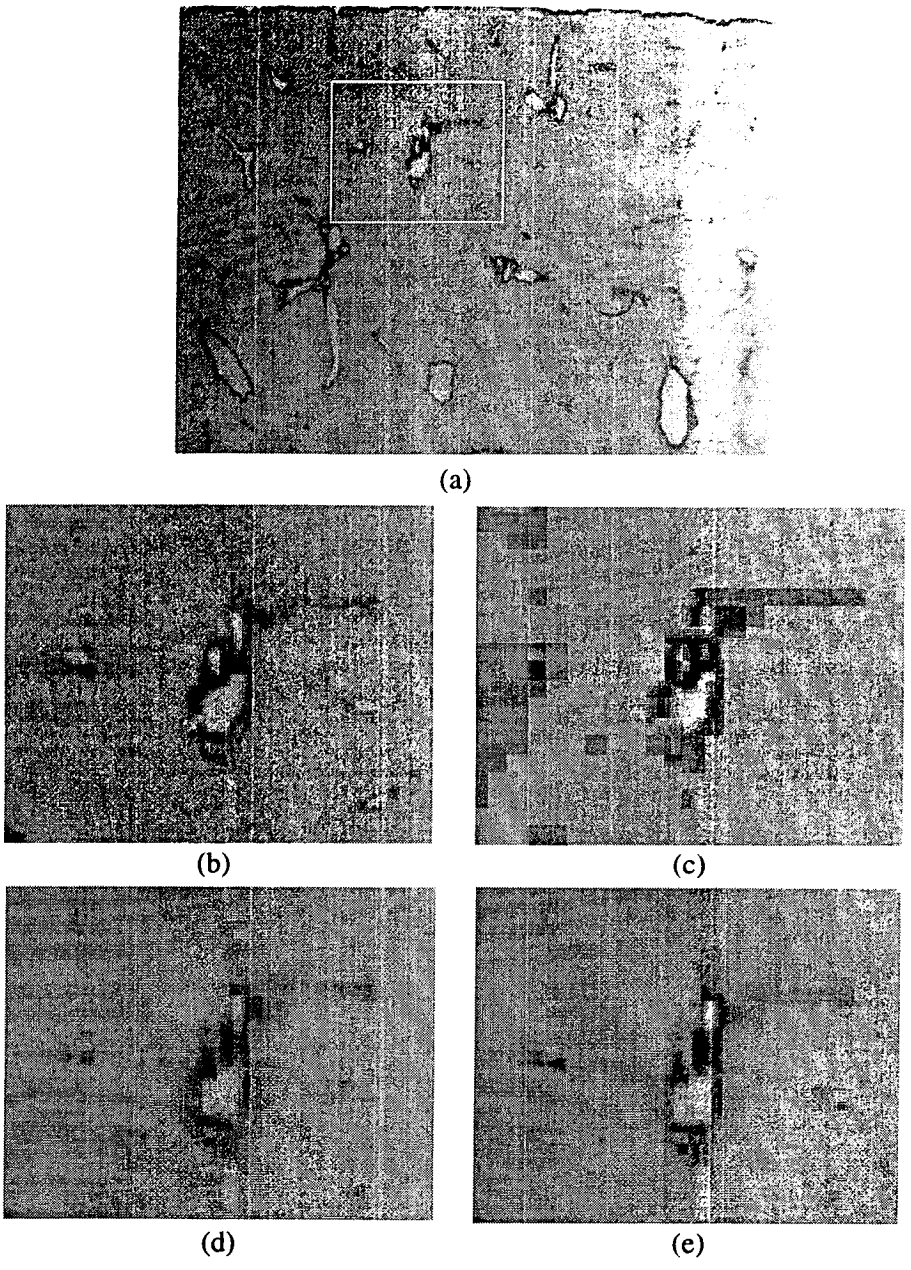

Figure 3: Results obtained by compression techniques. (a) original image, (b) a region taken from original image, (c) JPEG, (d) Classical PCA and (e) PCA network

be seen in Figure 4. In this Figure, the time spent by a PCA network with 3 principal components for compression are shown for 25 images, where it can be noted that consumed time decreases as the number of image increases.

In other words, this means that considering a large image sequence, it is expected as greater the number of images that has been compressed as better the neural network perfor-

	MSE (pixel grayscale value)					
	JPEG		PCA Network		Classical PCA	
	96 %	93 %	96.875%	93.75%	96.875%	93.75%
Image 1	9.72	7.01	5.82	4.91	5.43	4.69
Image 2	9.97	7.58	6.75	5.49	6.17	5.17
Image 3	8.91	6.44	5.20	4.08	4.64	3.77
Image 4	9.30	6.72	5.80	4.36	5.18	4.07
Image 5	9.47	6.76	6.08	4.50	5.30	4.16
Image 6	8.90	6.44	5.19	4.00	4.64	3.75
Image 7	9.04	6.51	5.41	3.96	4.81	3.74
Image 8	8.99	6.48	5.39	4.06	4.79	3.81
Image 9	8.93	6.40	5.32	3.92	4.66	3.68
Image 10	9.47	6.63	5.61	4.22	5.01	3.97

Table 1: Results for PCA and JPEG Approaches

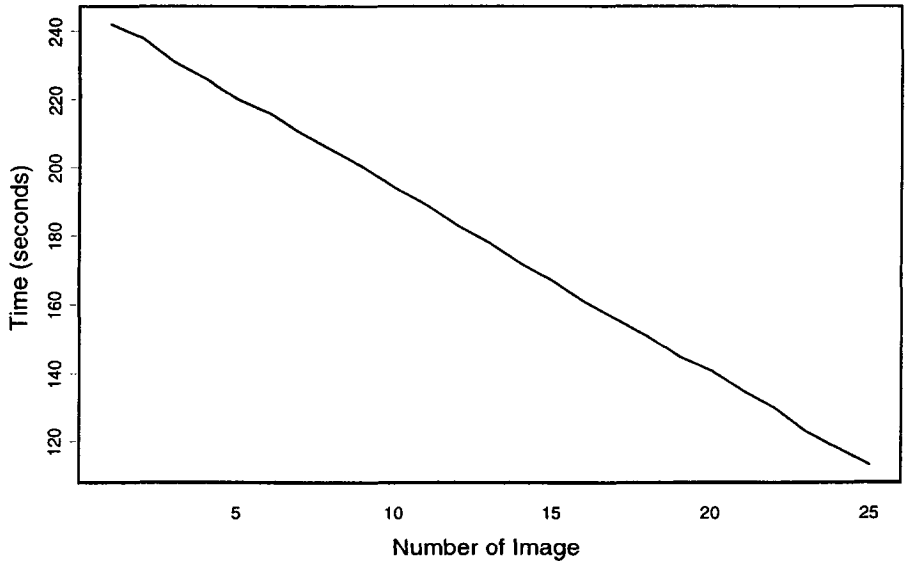

Figure 4: Results for compressing an image sequence

mance. Apparentely, this would be an advantage of the PCA network in relation to classical PCA but additional tests with different image sets should be analyzed in order to obtain

more precise results. In relation to the classical PCA, the time consumed will be the same independently of the number of images considered.

On the other hand, a comparison between the PCA network and JPEG performance shows that, for all compression ratios used in this experiment, the results obtained by PCA network were better than those obtained by JPEG, as it can be seen in Table 1.

Another point to be considered in comparing these two approaches is that the calculations used by the neural network are simpler than those performed by either JPEG standard or classical statistical approach. Figure 4 shows the original image and a region taken from it and the recovered images with compression ratio of 96.875% for both PCA approaches and 96% for JPEG standard.

It can be noted from Figure 3(c) that JPEG compression makes appear some vertical patterns. The presence of these patterns difficult the image analysis and the visualization of small objects in Figure 3(b). In Figures 3(d) and (e), it can be noticed that most of the small structures is kept. Keeping the small structures is important in analyzing liver images because some of these structures can indicate some kind of disease.

6 Conclusion and Future Work

Three tecniques for image compression, JPEG, classical PCA and PCA neural network were presented and used to compress a sequence of medical images.

The comparative study showed that the performance of PCA neural network was quite similar to that presented by the classical method and superior to JPEG for all the compression ratios assumed to compress the images in our experiments.

As a future work, the use of a nonlinear compression technique to reduce the dimensionality of image data will be investigated and compared to the results presented in this work. For instance, multilayer perceptrons operating in the self-organized manner promises to perform a more effective data compression, acting as feature detectors in many applications.

To verify the generality of the conclusions presented in this article, different sequences of images will be analyzed.

7 Acknowledgments

We are grateful to Dr. H. Lenzi of the Pathology Department of Fiocruz (Fundação Oswaldo Cruz - Rio de Janeiro) by the image set used in this work and to the partial financial support from Fapesp and Capes.

References

[1] P. Foldiak. Adaptive network for optimal linear feature extraction. Proc. International Joint Conference on Neural Networks, 1:401–405, 1989.

[2] S. Haykin. Neural Networks - A Comprehensive Foundation. Macmillan College Publishing Company, 1994.
[3] B. Lahme and R. Miranda. Karhunen-loeve decomposition in the presence of simmetry - part I. IEEE Transactions on Image Processing, 8(9):1183–1190, 1999.
[4] J. Lucca Filho. Implementação Modular da Técnica de Compressão e Decompressão JPEG para Imagens. Dissertação de Mestrado, IFSC-USP, São Carlos, SP, 1994.
[5] B. F. J. Manly. Multivariate Statistical Methods. Chapman and Hall, 1986.
[6] J. Mao and A. K. Jain. Artificial neural networks for feature extraction and data projection. IEEE Transactions on Neural Networks, 6(2):296–317, 1995.
[7] M. Nelson. The Data Compression Book. MT Books, New York, 1992.
[8] E. Oja. A simplified neuron model as a principal component analyzer. Journal of Mathematical Biology, 15:267–273, 1982.
[9] J. Rubner and P. Tavan. A self-organization network for principal component analysis. Europhysics Letters, 10:693–698, 1989.
[10] T. D. Sanger. Optimal unsupervised learning in a single-layer feedforward neural network. Neural Networks, 12:459–473, 1989.

NOTAS DO ICMC

SÉRIE COMPUTAÇÃO

048/2000 LINS, L.; LINS, S.; MORABITO, R. – Na n-tet graph approach for non-guillotine packings of n-dimensional boxes into na n-container.

047/2000 OLIVEIRA JR., O. N.; MARCHI, A.R.; MARTINS, M.S.; MARTINS, R.T.; NUNES, M.G.V. – A critical analysis of the performance of english-portuguese-english MT systems.

046/2000 SOUZA, P.S.L.; SANTANA, M.J.; SANTANA, R.H.C. – Escalonamento de processos: uma contribuição para a convergência da área.

045/2000 MARTINS, R.T.; RINO, L.H.M.; NUNES, M.G.V.; MONTILHA, G.; OLIVEIRA JR., O. N. – An interlingua diming at communication on the Web: how language-independent can it be?

044/99 FELTRIM, V.D.; FORTES,R.P.M. – Evaluation of a reverse engineering method through the application in a hypermedia system.

043/99 PANSANATO, L.T.E.; NUNES, M.G.V. – EHDM: Método para projeto de hiperdocumentos para ensino.

042/99 MORABITO, R.; ARENALES, M. – Optimizing the cutting of stock plates in a furniture company.

041/98 SANTOS-MEZA, E.; SANTOS, M.O.; ARENALES, M.N. – Lot sizing and scheduling in na automated foundry.

040/98 FELTRIM, V.D.; FORTES, R.P.M. – Uma modelagem do domínio de engenharia reversa de software utilizando o método OOHDM.

039/98 FERREIRA, V.G.; MELLO, O .D.; OLIVEIRA, J.N.; FORTUNA, A . O . - Tópicos teóricos e computacionais em escoamentos de fluidos.