Using C-NLPCA to extracting movements from video sequences

S.C. Botelho, R. de Bem, M. Figueiredo, W. Lautenschläger, T. Centeno
FURG - Av. Italia Km8, 96201-000 Rio Grande/RS - Brazil

Abstract

This article presents a methodology to extract principal components of large set of data, considering the existence of possible nonlinear relations among them, called C-NLPCA (Cascaded nonlinear principal component analysis), and evaluates its use in the extraction of main human movements in image series, aiming for the development of methodologies and techniques for skill transfer from humans to robotic/virtual agents. The experimental analysis uses a video images sequence to obtain principal movements of a human hand. The results are confronted with the linear traditional techniques.

1 Introduction

The principal components analysis (PCA) is a statistic method used for data multivariate analysis [4]. Such a method provides linear relations among the elements in a set of variables, giving as output the principal components (PC) which describe the variability patterns in such a set. However, when the variables have nonlinear relationships among then, classical PCA is not applied. In this case, we have a set of alternatives methods [6, 2, 8] like, for example NLPCA (Nonlinear Principal Component Analysis) [5], based in artificial neural networks (ANN). Nevertheless, due to internal saturation problems of the ANN [7, 3], the application of the NLPCA method is restricted to the data analysis with limited number of variables involved (limited dimensional size).

The authors have been working for the extension of the NLPCA technique to allow treatment of higher dimension data [1]. So, it is proposed a structure in layers, where NLPCAs are cascaded. Such a methodology, C-NLPCA (Cascaded Nonlinear Principal Components Analysis), was initially used for the determination of temporal series variability patterns of satellite images representing the ocean superficial temperature. The results obtained were encouraging, motivating the interest of its use in other applications like, for instance, the skill transfer tasks problem, that needs a formalization of intelligent human behavior and decision making processes into an algorithmic framework.

We argue that the high dimension problem associated with skill transfer can be a good application for our Cascaded Nonlinear Principal Components Analyses. Thus, in this paper, it is studied the use of C-NLPCA for the determination of movement patterns made by a human hand obtained from images temporal series, captured by a camera. These patterns could be used to transfer skills between humans and robots.

The first focus of this paper will be an explanation over NLPCA method. Later on, it will be done a specific discussion about peculiar details of the C-NLPCA, in order to finally show experiments and results, as well as an analysis of those.

2 NonLinear Principal Component Analysis Methods

2.1 The theory of NLPCA

Principal Component Analysis (PCA) only allows a linear mapping from a input vector $\vec{X}$ to a principal component u. On the other hand, NLPC is obtained using a multi layer autoassociative Neural Network, see figure 1. To perform NLPCA, the ANN contains 3 hidden layers of neurons between the input and output layers. Hidden layers have nonlinear activation functions between the input and bottleneck layers and between the bottleneck and output layers. Hence, the network models a composition of functions. The five-layer NLPCA network has p nodes in the input layer, r nodes in the third (bottleneck) layer, and p in the output layer. The bottleneck layer provides a dimensional reduction from p of original data to r.

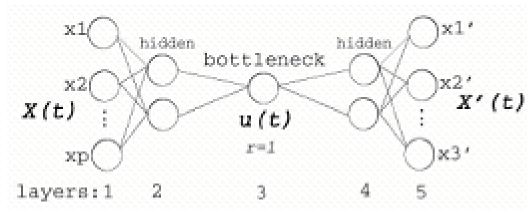

Figure 1: NLPCA: Neural Network to map Nonlinear Components.

3 C-NLPCA: Cascaded Nonlinear Principal Component Analysis

When it is necessary to run NLPCA with large dimension datasets, like images, there is a noticeable increase of parameters (weights) associated with the neurons of the ANN, thus leading to the necessity to have a bigger number of temporal samples, so that this value can be near to the parameters of the ANN [3].

Thus, when the original dataset have many dimensions, several authors opt to filter the data before the NLPCA analysis, like the use of PCA reduction techniques [1, 3], leading to erroneous outputs or, at least, producing coarser results.

Our C-NLPCA has the aim to allow the direct and totally nonlinear analysis of high dimension dataset, using a cascaded set of successive NLPCAs, see figure 2. The architecture is compose by two main stages: *reduction* and *expansion*, both detailed posteriorly.

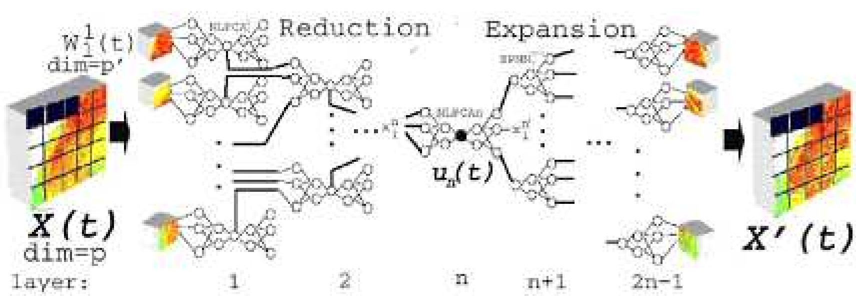

Figure 2: C-NLPCA: a set of layers grouped in Reduction and Expansion Stages. The pointed out neuron gives the pattern associated to the reduction of the original dataset.

Obtaining the C-NLPC - The Reduction Process C-NLPCA assumes that p' is the *ideal* dimension for the input data. Thus, we divide the original input image with dimension p into smaller windows with dimension p'. These windows are used directly as input of a first layer of NLPCAs. Each NLPCA, constrained by the saturation requirements, finds a local principal component (local reduction) of one window. The resulting patterns (reductions) are used as input to a new layer of C-NLPCA. The process is repeated up to layer n composed by only one NLPCA network. The output of the bottleneck neuron of this only NLPCA is the nonlinear principal component, $u_n(t)$ of the original large dataset.

Obtaining C-NLPCAs - The expansion Stage The second role of the Principal Component Analyses, called expansion, is to obtain the data associated with each principal component reduction (PC, C-NLPC) in the original dimension of the image (PCA, C-NLPCA). Hence, due to the cascading process, we have lost the original dimension of the input image, it is then necessary a method to obtain the expansion of the reduction. Expansion Stage is trained to expand the nonlinear principal component, $u_n(t)$, in a set of nonlinear expanded images. In fact, we propose a bottleneck layered structure to obtain the reduction/expansion of NLPCs. Expansion layers are symmetric with reduction layers, resulting in a total of $(2 * n - 1)$ layers. They are composed by simple backpropagation networks **BPNN** (without bottleneck neuron). The input of each **BPNN** is an output of the last layer. We use the original propagated image to train the desired outputs of **BPNN** networks. To obtain the next component, is necessary to calculate the residues associated with the current, which will be the input to a new C-NLPC analysis.

4 Experiments and Results

In this paper, we present some experiments involving skill transference. More precisely, we are interested to identify principal movements made in a sequence of human gestures, captured through a static camera. These movements, once expanded, will be performed by a robotic or virtual hand.

In the experiments, in a 25 frames/second rate, a temporal series of 128 gray scale images were analyzed, each one with 60 pixels in width and height, resulting in an image vector with 3600 elements. The video was obtained in a real and usual light conditions, without any kind of specific treatment. Besides the proper C-NLPCA, the linear method, PCA, was also applied to these dataset, analysing some results/features, according to:

- Temporal patterns (principal components - PC).
- Reconstructed Image: Obtained image as return of the methods, through the composition of spatial and temporal principal components, that aims to rebuild the samples.

The experiment made consisted of the oscillation of a human hand (left-right), as global movement, associated to the movement of opening and closing it, being that a local movement. The figure 3 illustrates some samples of the series.

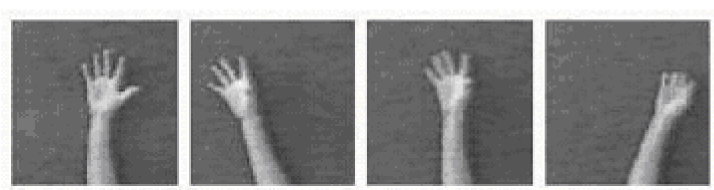

Figure 3: Some frames of the 128 analyzed images series.

For a better analysis we evaluate the nonlinear and linear spatial patterns, PCs (solid lines) and NLPCs (dashed lines), in the same figure 4. One can notice that the first linear principal component (PC1, represented in the figure 4 (a) in solid line), has a high frequency behavior which does not characterize the movement of the hand, but the local lighting conditions. It does not occur in the C-NLPC1.

In the C-NLPC1, as well as in the C-NLPC2 (and even in the PC2), it is possible to notice clearly the oscillating patterns of the movement: both represent the left-right hand movement, but with distints means to peaks and valleys. High frequency signals in C-NLPC3 represents closing and opening hand movements. We can notice that C-NLPC3 shows a different pattern from the others, showing more sensitivity to isolated cases, like the fact of the closed hand to exceed the image limits, showed by a deep valley.

The non-linearity feature of C-NLPCA method will appear whit the reconstructed data. Once this peculiarity represents more characteristics of original data than the linear method, peaks and valleys in input are represented in C-NLPCA expansion, despite the PCA reconstruction.

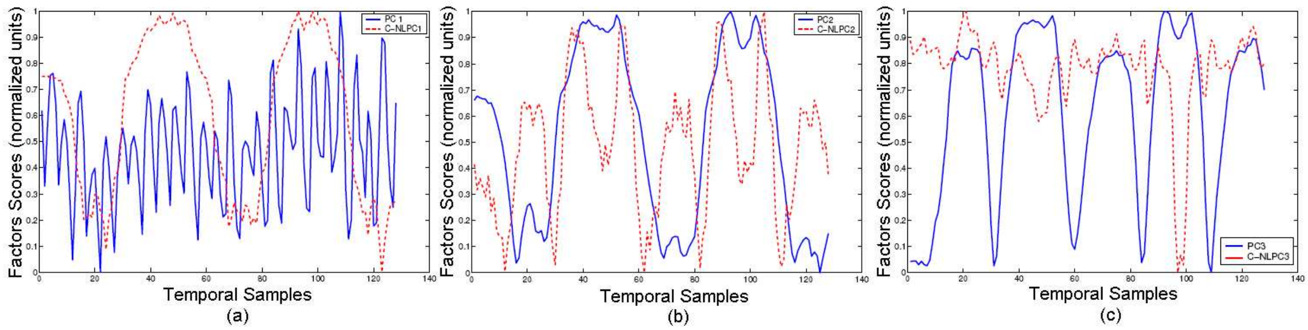

Figure 4: Comparison of nonlinear and linear principal components: (a) first component (b) second component (c) third component

The principal components of movements can be mapped to robot/vitual coordinated space (joint actuators, motor of wheels, etc...). We are using a neural network to do this.

5 Conclusion

In the present study, we propose an original method, called C-NLPCA, to extract principal components of human movement captured by a camera, aiming the skill transfer between humans and virtual/robotic agents. C-NLPCA reduces the dimension of large dataset, obtaining principal components of them. We use a cascaded Neural Network in a bottleneck structure to obtain dimension reduction, giving the principal components of the data variability. The same structure is also used to expand the data from obtained principal component.

From a video images sequence, the method was applied to obtain principal movements of a human hand. As results, the PCA finds as first mode the lighting conditions variability (not principal, but whole linear). On the other hand, the C-NLPCA network has demonstrated the capability of isolating the principal variability (hand translation) in its first PC. Moreover, closing/opening movements, and besides the fact of the closed hand to exceed the image limits were situations extracted of the three principal C-NLPCs obtained. The experiments conducted to satisfactory results, signaling the C-NLPCA use in other applications, as for instance, the navigation in autonomous vehicles based on images, data compression, etc.

References

[1] S. Botelho, R. de Bem, M. Mata, and I. Almeida. Applying neural networks to study the mesoscale variability of oceanic boundary currents. *LNAI/LNCC*, 2871:684–688, 2003.

[2] K. I. Diamantaras and S. Y. Kung. *Principal Component Neural Networks: Theory and Applications*. Wiley, 1996.

[3] W. Hsieh. Nonlinear principal component analysis by neural network. *Tellus*, 53A:599–615, 2001.

[4] Jr. J. F. Hair, R. E. Anderson, R. L. Tatham, and W. C. Black. *Multivariate data analysis: with readings*. Prentice Hall, 4th edition, 1995.

[5] M. A. Kramer. Nonlinear principal component analysis using autoassociative neural networks. *AlChe Journal*, 37:233–43, February 1991.

[6] C. Lee. *Learning Reduced-Dimension Models of Human Actions*. PhD thesis, The Robotics Institute - CMU, May 2000.

[7] Edward C. Malthouse. Limitations of nonlinear PCA as performed with generic neural networks. *IEEE Trans. NN*, 9(1), January 1998.

[8] M. Scholz and R Vigrio. Nonlinear pca: a new hierarchical approach. In *ESANN'2002 proceedings - Europena Symposium on Artificial Neural Networks*, pages 439–444, Bruges(Belgium), April 2002.