

Rodrigo Andrade de Bem

*CNLPCAMove - Extração de Componentes
Principais de Movimento com o método
C-NLPCA*

Rio Grande

Janeiro 2005

Rodrigo Andrade de Bem

*CNLPCAMove - Extração de Componentes
Principais de Movimento com o método
C-NLPCA*

Projeto de conclusão de curso de Engenharia
de Computação da Fundação Universidade
Federal de Rio Grande, apresentado pelo
acadêmico Rodrigo Andrade de Bem.

Orientadora:

Prof.^a Dr.^a Silvia Silva da Costa Botelho

FUNDAÇÃO UNIVERSIDADE FEDERAL DO RIO GRANDE
ENGENHARIA DE COMPUTAÇÃO

Rio Grande

Janeiro 2005

Prof.^a Dr.^a Silvia Silva da Costa Botelho
Departamento de Física - Furg
Orientadora

Prof. Ms. Alessandro de Lima Bicho
Departamento de Matemática - Furg

Prof. Ms. Denis Franco
Departamento de Física - Furg

A Carol.

Agradecimentos

Agradeço a Deus por ter me acompanhado em todos instantes, principalmente nas noites em claro e nos momentos mais difíceis. Agradeço a Prof^a. Silvia, que com sua maneira obstinada de trabalhar, cobrar resultados e de cobrar resultados :-), foi de fundamental importância na realização do trabalho, e como minha orientadora durante praticamente todo curso de graduação. Agradeço ao Prof. Vinícius de Oliveira, o popular “Bauru”, e ao amigo Rafael Cunha, pelo empréstimo dos livros, *Numerical Recipes in C* e *Multivariate Data Analysis*, respectivamente, sem os quais o trabalho ficaria muito mais difícil. Agradeço a minha namorada, Carol Cunha, que me apoio e incentivou sempre, e principalmente entendeu as minhas ausências. Agradeço a meus pais, irmãos e família pelo apoio e torcida. E a todos os amigos e colegas que de alguma forma contribuíram na elaboração do trabalho.

“Em um grande problema, nenhum detalhe é pequeno.”

Provérbio Francês

Resumo

Este trabalho tem como principal objetivo apresentar a metodologia empregada para execução do C-NLPCA (*Cascaded Nonlinear Principal Component Analysis*), e avaliar seu uso na extração de componentes principais de movimento de séries de imagens digitais 2D, visando o desenvolvimento de metodologias e técnicas que permitam a execução de tal tarefa com eficácia e robustez. O método C-NLPCA é um método de análise multivariada de dados, baseado no NLPCA (*Nonlinear Principal Component Analysis*), que tem como principais características a capacidade de extrair componentes principais de variação de grande conjuntos de variáveis e de considerar possíveis relações não lineares existentes entre as mesmas. Sobre o problema da extração de padrões de movimento, para que se tenha um melhor entendimento, pode-se considerar uma imagem 2D como um grande conjunto de variáveis (pixels), e uma série de imagens, como observações dessas variáveis ao longo do tempo. Assim, os padrões de movimento são os componentes principais de variação desta série, ou seja, os fatores que melhor caracterizam a variação temporal e espacial dos pixels nas mesmas. A determinação eficaz dos padrões de movimentos de séries temporais de imagens pode ser aplicada para solução de problemas como reconhecimento de gestos, transferência de habilidades (*skill transfer*) para agentes robóticos e virtuais, teleoperação, animação gráfica e interação homem-computador (*HCI*). Serão apresentados ao longo do trabalho uma revisão bibliográfica, a descrição detalhada dos métodos PCA, NLPCA e C-NLPCA, os resultados de dois experimentos realizados e suas respectivas análises, e por fim, as conclusões e considerações finais.

Abstract

The main goal of this work is to present the methodology adopted in the execution of C-NLPCA (Cascaded Nonlinear Principal Component Analysis), and to evaluate its performance in the extraction of movements principal components from 2D digital image series. This evaluation will allow the development of new methodologies and techniques to perform this task with the robustness and efficiency required. The C-NLPCA is a multivariate data analysis method, based on NLPCA (Nonlinear Principal Component Analysis). Its main features are the capacity to extract principal components from large datasets and the power to consider nonlinearities in the data. To have a better comprehension about the extraction of movements principal components, we can consider one 2D image as a large set of variables (pixels), and one image series, as the observations of this variables in time. Thus, the movements patterns are the series' variability principal components, and it means that these are the factors that better describes the temporal and spacial variability in the dataset. The efficient extraction of image series' movements patterns can be applied to treat several problems, as gestures recognition, skill transfer to robotic and virtual agents, teleoperation, graphics animation and human-computer interaction. We will present an overview of related works, a detailed description of PCA, NLPCA and C-NLPCA methods, the performed experiments and their results, and finally, the conclusions and final considerations.

Sumário

1	Revisão bibliográfica e trabalhos relacionados	p. 16
1.1	PCA, NLPCA e C-NLPCA - Métodos de análise de componentes principais	p. 16
1.1.1	Problema da determinação dos componentes principais	p. 18
1.1.1.1	Vetor imagem	p. 18
1.1.1.2	Espaço das imagens	p. 18
1.1.2	Análise dos Componentes Principais	p. 18
1.1.2.1	Teoria do NLPCA	p. 20
	Considerações sobre o NLPCA	p. 22
1.1.3	Trabalhos relacionados	p. 23
1.2	Aplicações	p. 26
2	C-NLPCA: Análise Cascadeada dos Componentes Principais	p. 28
2.1	Obtendo o C-NLPC - Estágio de Redução	p. 29
2.2	Obtendo os C-NLPCAs - Estágio de Expansão	p. 30
3	Experimentos realizados	p. 32
3.1	Primeiro experimento	p. 34
3.1.1	Padrões temporais de movimento - PCs	p. 35
3.1.2	Padrões espaciais de movimento	p. 35
3.1.3	Reconstrução das observações	p. 36
3.1.4	Projeção das reconstruções das amostras no espaço 3D	p. 37
3.1.5	Percentual de erro inexplicado por cada padrão	p. 37

3.1.6	Análise dos resultados	p. 37
3.2	Segundo experimento	p. 49
3.2.1	Análise dos resultados	p. 49

Lista de Figuras

1	“Scatter plot” do conjunto de dados exemplo.	p. 17
2	Primeiro componente principal do conjunto de dados.	p. 17
3	Construção do vetor imagem.	p. 18
4	Um conjunto de <i>vetores imagem</i> , formando uma série temporal de imagens $\vec{X}(t) = \{x_1, x_2, x_3\}$, no <i>espaço das imagens</i> , com os componentes principais (a1 e a2).	p. 19
5	NLPCA: rede neural para mapear componentes principais não lineares.	p. 21
6	C-NLPCA: um conjunto de camadas agrupadas nos estágios de Redução e Expansão. O neurônio destacado fornece o padrão associado à redução do conjunto de dados original.	p. 29
7	Alguns <i>frames</i> da primeira série de 128 imagens analisada.	p. 33
8	Alguns <i>frames</i> da segunda série de 128 imagens analisada.	p. 33
9	Comparação entre PC1 e C-NLPC1.	p. 40
10	Comparação entre PC2 e C-NLPC2.	p. 40
11	Comparação entre PC3 e C-NLPC3.	p. 41
12	Padrões espaciais lineares em tons de cinza. (a) Primeiro padrão. (b) Segundo padrão. (c) Terceiro padrão. (d) Soma dos 3 padrões lineares. A escala utilizada é mostrada ao lado da imagem (d).	p. 41
13	Padrões espaciais lineares em <i>pseudocolor</i> . (a) Primeiro padrão. (b) Segundo padrão. (c) Terceiro padrão. (d) Soma dos 3 padrões lineares. A escala utilizada é mostrada ao lado da imagem (d).	p. 42
14	Expansão das observações referentes ao valor mínimo do C-NLPC de cada padrão, em tons de cinza. (a) Primeiro padrão. (b) Segundo padrão. (c) Terceiro padrão. A escala utilizada é a mesma mostrada na Figura 12.	p. 42

15	Expansão das observações referentes ao valor mínimo do C-NLPC de cada padrão, em <i>pseudocolor</i> . (a) Primeiro padrão. (b) Segundo padrão. (c) Terceiro padrão. A escala utilizada é a mesma mostrada na Figura 13.	p. 43
16	Expansão das observações referentes ao valor máximo do C-NLPC de cada padrão, em tons de cinza. (a) Primeiro padrão. (b) Segundo padrão. (c) Terceiro padrão. A escala utilizada é a mesma mostrada na Figura 12.	p. 43
17	Expansão das observações referentes ao valor máximo do C-NLPC de cada padrão, em <i>pseudocolor</i> . (a) Primeiro padrão. (b) Segundo padrão. (c) Terceiro padrão. A escala utilizada é a mesma mostrada na Figura 13.	p. 44
18	Reconstrução, a partir dos 3 primeiros padrões de movimento, de uma das 128 observações analisadas, em tons de cinza. (a) Imagem original. (b) Reconstrução com o método linear (PCA). (c) Reconstrução (expansão) com o método não linear (C-NLPCA). A escala utilizada é a mesma mostrada na Figura 12.	p. 44
19	Reconstrução, a partir dos 3 primeiros padrões de movimento, de uma das 128 observações analisadas, em <i>pseudocolor</i> . (a) Imagem original. (b) Reconstrução com o método linear (PCA). (c) Reconstrução (expansão) com o método não linear (C-NLPCA). A escala utilizada é a mesma mostrada na Figura 13.	p. 45
20	Evolução temporal do valor de cada variável (pixel) da série de imagens originais.	p. 45
21	Evolução temporal do valor de cada variável (pixel) da reconstrução da série de imagens originais com o método PCA.	p. 46
22	Evolução temporal do valor de cada variável (pixel) da reconstrução da série de imagens originais com o método C-NLPCA.	p. 46
23	Percentual de erro inexplicado por cada método, de acordo com o número de padrões utilizados para reconstruir os dados.	p. 47
24	Mão ultrapassando os limites da imagem.	p. 47
25	Comparação entre o C-NLPC3 e os PCs de mais alta ordem.	p. 48
26	Comparação entre PC1 e C-NLPC1.	p. 51
27	Comparação entre PC2 e C-NLPC2.	p. 51

28	Comparação entre PC3 e C-NLPC3.	p. 52
29	Padrões espaciais lineares em tons de cinza. (a) Primeiro padrão. (b) Segundo padrão. (c) Terceiro padrão. (d) Soma dos 3 padrões lineares. A escala utilizada é mostrada ao lado da imagem (d).	p. 52
30	Padrões espaciais lineares em <i>pseudocolor</i> . (a) Primeiro padrão. (b) Segundo padrão. (c) Terceiro padrão. (d) Soma dos 3 padrões lineares. A escala utilizada é mostrada ao lado da imagem (d).	p. 53
31	Expansão das observações referentes ao valor mínimo do C-NLPC de cada padrão, em tons de cinza. (a) Primeiro padrão. (b) Segundo padrão. (c) Terceiro padrão. A escala utilizada é a mesma mostrada na Figura 29.	p. 53
32	Expansão das observações referentes ao valor mínimo do C-NLPC de cada padrão, em <i>pseudocolor</i> . (a) Primeiro padrão. (b) Segundo padrão. (c) Terceiro padrão. A escala utilizada é a mesma mostrada na Figura 30.	p. 54
33	Expansão das observações referentes ao valor máximo do C-NLPC de cada padrão, em tons de cinza. (a) Primeiro padrão. (b) Segundo padrão. (c) Terceiro padrão. A escala utilizada é a mesma mostrada na Figura 29.	p. 54
34	Expansão das observações referentes ao valor máximo do C-NLPC de cada padrão, em <i>pseudocolor</i> . (a) Primeiro padrão. (b) Segundo padrão. (c) Terceiro padrão. A escala utilizada é a mesma mostrada na Figura 30.	p. 55
35	Reconstrução, a partir dos 3 primeiros padrões de movimento, de uma das 128 observações analisadas, em tons de cinza. (a) Imagem original. (b) Reconstrução com o método linear (PCA). (c) Reconstrução (expansão) com o método não linear (C-NLPCA). A escala utilizada é a mesma mostrada na Figura 29.	p. 55
36	Reconstrução, a partir dos 3 primeiros padrões de movimento, de uma das 128 observações analisadas, em <i>pseudocolor</i> . (a) Imagem original. (b) Reconstrução com o método linear (PCA). (c) Reconstrução (expansão) com o método não linear (C-NLPCA). A escala utilizada é a mesma mostrada na Figura 30.	p. 56
37	Evolução temporal do valor de cada variável (pixel) da série de imagens originais.	p. 56

38	Evolução temporal do valor de cada variável (pixel) da série de imagens originais.	p. 57
39	Evolução temporal do valor de cada variável (pixel) da série de imagens originais.	p. 57
40	Percentual de erro inexplicado por cada método, de acordo com o número de padrões utilizados para reconstruir os dados.	p. 58
41	<i>Frames</i> mostrando o ápice do movimento das mãos esquerda (a) e direita (b), na série de imagens.	p. 58

Introdução

Este trabalho tem como principal objetivo apresentar a metodologia empregada para execução do C-NLPCA (*Cascaded Nonlinear Principal Component Analysis*), e avaliar seu uso na extração de componentes principais de movimento¹ de séries de imagens digitais 2D, visando o desenvolvimento de metodologias e técnicas que permitam a execução de tal tarefa com eficácia e robustez.

O método C-NLPCA (BEM; BOTELHO; MATA, 2003) é um método de análise multivariada de dados, baseado no NLPCA (*Nonlinear Principal Component Analysis*), que tem como principais características a capacidade de extrair componentes principais de variação de grandes conjuntos de variáveis e de considerar possíveis relações não lineares existentes entre as mesmas. Na extração de padrões de movimento de séries temporais de imagens, essas duas características são encontradas. A primeira delas devido à dimensão das imagens, que em qualquer aplicação ultrapassa facilmente os 1000 pixels (cada pixel é uma variável); e a segunda, devido a modificações não homogêneas nos valores dos pixels das imagens, como o deslocamento de objetos com velocidades e direções distintas, o que muitas se caracteriza como uma variação não linear.

No problema da extração de padrões de movimento, para que se tenha um melhor entendimento, pode-se considerar uma imagem 2D como um grande conjunto de variáveis (pixels), e uma série de imagens, como observações dessas variáveis ao longo do tempo. Assim, a extração dos padrões de movimento existentes nessas imagens, consiste na busca por padrões de modificação de tais variáveis (pixels) ao longo do tempo. Essa busca pode ser feita em séries de imagens que mostram cenas em diferentes contextos, desde os movimentos de um atleta surfando, até a vista aérea de uma cidade. No entanto, o padrão de modificação das variáveis será diferente em cada caso.

Os padrões de movimento são considerados os componentes principais de variação de uma série de imagens, ou seja, os fatores que melhor caracterizam a variação temporal e espacial dos pixels nessas imagens. Em uma série de imagens que mostra, por exemplo,

¹Com o objetivo facilitar as definições dos métodos de análise, e diferenciar componentes principais, de componentes principais de movimento, este último termo será substituído por “padrões de movimento” ao longo do texto. Na Conclusão o termo voltará a ser empregado sem prejuízo para compreensão.

os movimentos de pássaro voando, um dos os componentes principais seria a translação do pássaro na imagem, considerando que esse movimento predomina a variação existente. Outro componente certamente caracterizaria o bater de suas asas, supondo que esse movimento seja registrado nitidamente nas imagens.

A determinação eficaz dos padrões de movimentos de séries temporais de imagens pode ser aplicada na solução de problemas como reconhecimento de gestos, transferência de habilidades (*skill transfer*) para agentes robóticos e virtuais, teleoperação, animação gráfica e interação homem-computador (*HCI*). Essa abordagem, se adotada para o tratamento de tais problemas, baseada na análise de imagens no espaço 2D, poderá trazer alguns problemas que terão que ser tratados, como a oclusão de pontos, e estabelecimento de correspondência entre pontos dos objetos nas imagens e possíveis modelos utilizados. No entanto, uma vez superadas tais limitações, o método poderá representar uma simplificação frente as abordagens encontradas na bibliografia.

No Capítulo 1 será apresentada uma visão geral dos tópicos, métodos e aplicações estudados, relacionados aos temas principais deste trabalho. O problema da determinação dos componentes principais será caracterizando, além de ser feita uma descrição detalhada dos métodos PCA e NLPCA. A metodologia empregada para execução do C-NLPCA será apresentada no Capítulo 2. O Capítulo 3 define os experimentos realizados, mostra os resultados obtidos e apresenta ainda uma discussão sobre esses resultados. Por fim, no Capítulo 3.2.1 é feita a avaliação final do trabalho elaborado, além de descritas as contribuições, códigos implementados e trabalhos futuros.

1 *Revisão bibliográfica e trabalhos relacionados*

Neste capítulo, pretende-se mostrar de modo geral os temas e trabalhos relacionados ao problema da extração de padrões de movimento, bem como abordar os assuntos que tiveram maior relevância na execução do trabalho. Pode-se dividir claramente os tópicos abordados e revisados no desenvolvimento do presente trabalho da seguinte maneira: métodos **PCA**, **NLPCA** e **C-NLPCA** e **aplicações** para os padrões de movimento extraídos através da análise de componentes principais. A seguir esses dois temas serão detalhados.

1.1 **PCA, NLPCA e C-NLPCA - Métodos de análise de componentes principais**

A análise de fatores (*factor analysis*) é um nome genérico dado a uma classe de métodos de análise multivariada que têm como principal objetivo a extração dos fatores ou dimensões que melhor representam a estrutura de interdependência (correlação) entre um grande número de variáveis. Nesse tipo de análise, chamada de interdependente, todas as variáveis são analisadas simultaneamente, cada uma em relação a todas as outras, com o objetivo de determinar os fatores que maximizam a “explicação” da variação existente em todo conjunto de dados. Assim como outros métodos de análise multivariada, os métodos de análise de fatores são indicados para análise de variáveis que possuem relações lineares entre si (HAIR et al., 1995). O método de análise dos componentes principais (*principal component analysis* - PCA), também conhecido como transformada de Karhunen-Loève (LEE, 2000), é um desses métodos, e é largamente utilizado em várias áreas do conhecimento. Sua principal característica é a determinação de componentes principais¹ ortogonais entre si, que constituirão uma nova base que maximize a variação dos dados analisados. A busca de tais componentes baseia-se no fato que regularidades no

¹Na análise dos componentes principais o termo fator é substituído por componentes principais.

conjunto de dados farão com que as observações das variáveis organizem-se em aglomerados (*clusters*), motivo pelo qual a variação no conjunto dados poderá ser representada em uma dimensão reduzida, formada pelos componentes principais (BJÖRNSSON; VENEGAS, 1997). Como um exemplo simples, na Figura 1 são mostradas n observações de duas variáveis, x e y , ao longo do tempo. Pode-se notar facilmente a existência de uma relação suavemente não linear entre as duas. Destaca-se a dificuldade de visualização dos dados em grandes conjunto de variáveis, sendo apenas viável traçar o gráfico das variáveis duas a duas, ou três a três. Na Figura 2 é mostrado o primeiro componente principal do conjunto de dados, determinado de modo que a variação dos dados ao longo do mesmo seja máxima, além de minimizar o erro quadrático médio em relação às observações, que é caracterizado pela distância de cada ponto ao componente principal. Nas seções seguintes, o problema da análise dos componentes principais será abordado mais detalhadamente.

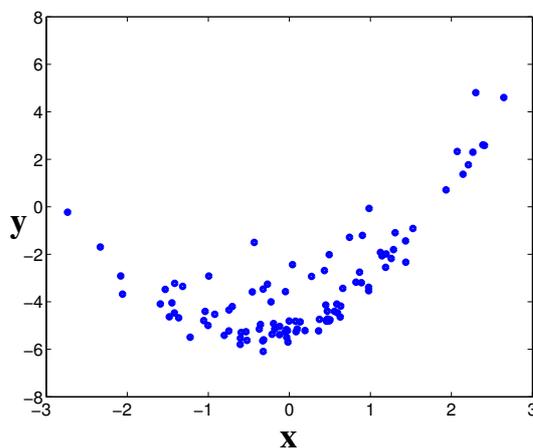


Figura 1: “Scatter plot” do conjunto de dados exemplo.

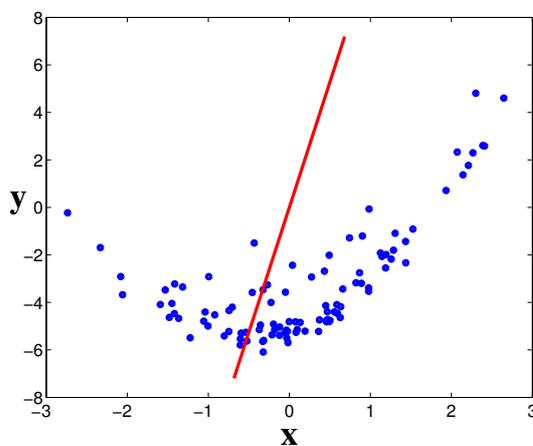


Figura 2: Primeiro componente principal do conjunto de dados.

1.1.1 Problema da determinação dos componentes principais

1.1.1.1 Vetor imagem

Independentemente de sua natureza, séries de dados temporais podem ser vistas como um vetor. Pode-se considerar, por exemplo, um conjunto de imagens como uma série de dados, tendo w pixels de largura, e h pixels de altura. Assim, o número de componentes desse vetor será $w * h$, onde cada pixel é uma coordenada do mesmo. A construção de tal vetor, chamado de *vetor imagem* \vec{X} , a partir de uma imagem bidimensional é executada com um simples concatenação - as linhas da imagen são inseridas uma após a outra (ROMDHANI; PSARROU; GONG, 1999), como mostrado de forma ilustrativa, onde cada quadrado representa um pixel, na Figura 3.

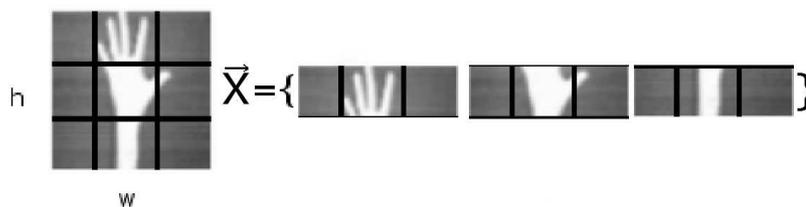


Figura 3: Construção do vetor imagem.

1.1.1.2 Espaço das imagens

O *vetor imagem* pertence a um espaço, chamado *espaço das imagens*, que é o espaço de todas as imagens cuja dimensão é $w * h$ pixels. Nesse espaço os *vetores imagem* passam a ser representados por pontos. A Figura 4 mostra *vetores imagem* de dimensão 3 (correspondentes a imagens hipotéticas com 3 pixels) dispostos no respectivo *espaço das imagens*. Cada ponto neste espaço equivale a uma imagem (vetor imagem \vec{X}). Percebe-se que os *vetores imagem*, devido a regularidades e semelhanças nas imagens, tendem a agruparem-se formando aglomerados (*clusters*) de pontos no *espaço das imagens*.

1.1.2 Análise dos Componentes Principais

Nota-se na Figura 4 que o *espaço das imagens* não é um espaço ótimo para análise dos dados. Isso acontece porque algumas coordenadas, dos diferentes vetores, podem trazer informações redundantes, o que se deve a semelhanças que podem existir entre eles, como descrito na seção anterior. Assim, em diversos casos pode ser útil a construção de um

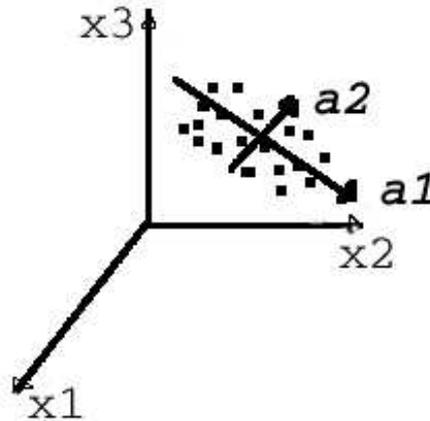


Figura 4: Um conjunto de *vetores imagem*, formando uma série temporal de imagens $\vec{X}(t) = \{x_1, x_2, x_3\}$, no *espaço das imagens*, com os componentes principais (**a1** e **a2**).

novo espaço, menor do que o original, para uma melhor descrição e análise dos dados. Os vetores base deste novo espaço construído são chamados de Componentes Principais.

Dessa forma, o objetivo da análise dos componentes principais é reduzir a dimensão do conjunto original, ou distribuí-lo de modo que a nova base descreva melhor os padrões de variação existentes nos dados. O PCA tenta representar e explicar a variação total no conjunto dos dados com poucos vetores (ou padrões) na nova base. A Figura 4 mostra a construção, por uma combinação linear, dos novos vetores da base (eixos). Os componentes nesta nova base não serão correlacionados. Pode-se ver, ainda na Figura 4, que a variação dos dados é máxima na direção do componente **a1**, que é definido como primeiro componente principal (primeiro padrão) de variação. Em segundo lugar, a maior variação dos dados dá-se na direção do componente **a2**, o qual é denominado segundo componente principal (segundo padrão) de variação, sendo necessariamente ortogonal ao primeiro.

Teoria do PCA Seja $\vec{X}(t) = \{x_1, \dots, x_p\}$ um conjunto de dados, com dimensão p , onde cada variável x_i , ($i = 1, \dots, p$) é uma série temporal contendo n observações. O PCA é obtido pela combinação linear de x_i , a série temporal u , e um vetor a associado:

$$u(t) = \mathbf{a} * \vec{X}(t), \quad (1.1)$$

então

$$\ll \| \vec{X}(t) - \mathbf{a}u(t) \|^2 \gg, \quad (1.2)$$

é minimizado ($\ll \dots \gg$ denota uma amostra ou série temporal). A variável u , chamada

de primeiro componente principal (PC), é uma série temporal, enquanto \mathbf{a} , o primeiro autovetor da matriz de covariância dos dados, descreve um padrão espacial de variação. A partir do resíduo, $\vec{X} - \mathbf{a}u$, pode-se obter repetir o processo para obtenção do segundo PCA, e assim por diante para os padrões de mais alta ordem (ver (HSIEH, 2001b) para maiores detalhes).

1.1.2.1 Teoria do NLPCA

O PCA permite somente o mapeamento linear de \vec{X} para u . Já o NLPCA possibilita o mapeamento não linear para a obtenção do componente principal (NLPC). Isso pode ser conseguido com a utilização de uma rede neural artificial (RNA) autoassociativa, multicamadas, com a presença de uma camada de gargalo (KRAMER, 1991)(HSIEH, 2001b)(KIRBY; SIROVICH, 1990), como exemplificado na Figura 5.

As RNAs são constituídas por componentes básicos chamados de neurônios. Existem na literatura implementações e propostas de redes neurais artificiais com diferentes topologias, modos de propagação, métodos de atualização de pesos e treinamento. As RNAs usadas no método NLPCA, que têm sua topologia mostrada na Figura 5, são redes MLP (*multi layer perceptron*) autoassociativas² (detalhes sobre RNAs em (FREEMAN; SKAPURA, 1992)). Nessas redes, cada neurônio possui um peso associado a cada entrada recebida dos neurônios da camada anterior. O somatório dessas entradas, multiplicadas pelos pesos correspondentes é por sua vez, a variável de entrada de uma função que fornece a saída do neurônio, que será propagada para próxima camada da rede. Esta função é chamada de função de ativação do neurônio. Quando funções não lineares, como a função sigmoideal ou a tangente hiperbólica, são adotadas como função de ativação de um determinado neurônio, o mesmo adquire um comportamento não linear em relação às entradas recebidas.

Os parâmetros da rede (pesos + polarizadores³) são atualizados durante o treinamento, objetivando que os valores desejados na camada de saída sejam obtidos. Para que isso seja possível, isto é, para que não haja **saturação**⁴ da RNA, é necessário que o número de amostras usadas no treinamento, no pior caso, seja semelhante ao número de parâmetros da rede, sendo que quanto maior o conjunto de amostras, mais o treinamento é facilitado. Em redes como a da Figura 5, usada no NLPCA, o número de parâmetros é dado por:

²Os valores desejados na camada de saída da rede, são os mesmos recebidos na camada de entrada.

³O polarizador é um peso associado a um valor constante, geralmente igual a 1.

⁴Não convergência para resultados satisfatórios.

$$2pm + 4m + p, \quad (1.3)$$

onde p é o número de neurônio das camadas de entrada e saída, e m é o número de neurônios das camadas escondidas 2 e 3. Quando o número de entradas recebidas pela rede, e conseqüentemente de neurônios nas camadas de entrada e saída, são muito elevados, há um aumento excessivo do número de parâmetros, obrigando que o conjunto de amostra usadas no treinamento seja, no mínimo, igualmente numeroso, sob pena de saturação da RNA durante o treinamento.

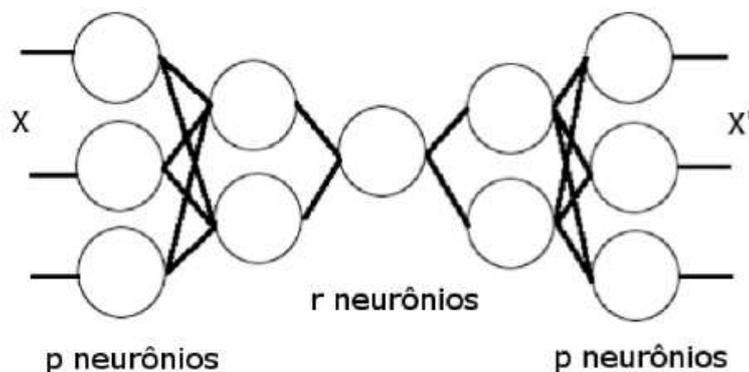


Figura 5: NLPCA: rede neural para mapear componentes principais não lineares.

Para executar o NLPCA, a RNA usualmente é formada por 3 camadas de neurônios escondidas entre uma camada de entrada e uma camada de saída. Os neurônios da camada escondida entre a entrada e o gargalo, também chamada de camada de codificação; e da camada escondida entre o gargalo e a saída, também chamada de camada de decodificação; têm funções de ativação não lineares. Enquanto os neurônios das camadas de entrada, gargalo e saída, possuem a função identidade, ou linear, como função de ativação. Devido a isso, a RNA modela uma composição de funções. A rede NLPCA de 5 camadas tem p nós na camada de entrada, r nós na terceira camada (gargalo), e p nós na saída. A camada de saída deve reproduzir os sinais de entrada apresentados a rede. A rede NLPCA permite a compressão/redução de dados porque, uma vez apresentada uma entrada à rede, esta última deve fornecer como resposta nos seus neurônios de saída, a mesma entrada (autoassociação). As entradas em um espaço p -dimensional devem passar pela camada do gargalo, isto é, mapeadas para o espaço r -dimensional, antes de reproduzir as saídas. Uma vez que a rede foi treinada, os valores de ativação do neurônio da camada de gargalo u fornecem o componente principal não linear dos dados (NLPC) (KRAMER, 1991).

Seja $\mathbf{f} : \mathcal{R}^p \rightarrow \mathcal{R}^r$ a função modelada pelas camadas 1, 2 e 3; e $\mathbf{s} : \mathcal{R}^r \rightarrow \mathcal{R}^p$ a função

modelada pelas camadas 3, 4 e 5. Usando essa notação, ao apresentar-se n imagens (série temporal de imagens) à rede, os pesos na rede NLPCA são determinados sob o objetivo de minimizar a seguinte função:

$$\min \sum_{l=1}^n \|\vec{X}_l - \vec{X}'_l\| \quad (1.4)$$

, onde \vec{X}'_l é a saída da rede para cada imagem \vec{X}_l apresentada. A relação 1.1 é agora generalizada para $u = \mathbf{f}(\vec{X})$, onde \mathbf{f} pode ser uma função não linear denotada pelo mapeamento feito pela RNA, da camada de entrada para camada de gargalo, e ao invés da Equação 1.2, $\|\vec{X}_l - \vec{X}'_l\|$ é minimizado pelas funções de mapeamento. $\vec{X}'_l = \mathbf{s}(u)$, onde \mathbf{s} mapeia os dados da camada de gargalo para camada de saída. Sobre o resíduo $\|\vec{X}_l - \vec{X}'_l\|$, pode ser repetido o mesmo processo para extração do segundo componente principal não linear, e assim por diante para os padrões de mais alta ordem (MONAHAN, 2000).

Considerações sobre o NLPCA Em geral, o mais sério problema relacionado ao NLPCA é a presença de mínimos locais na Equação 1.4. Como resultado disso, otimizações dessa Equação, realizadas através da atualização dos pesos na RNA, que partem de diferentes pontos iniciais, frequentemente convergem para mínimos diferentes, tornando o método instável. Uma das técnicas usadas para superar esse problema é a realização da análise do mesmo conjunto de dados a partir de vários pontos iniciais, o que é feito através da geração de vários conjuntos de pesos aleatórios para a rede neural artificial. Ao final das análises, apenas o conjunto que apresentou o menor erro (Equação 1.4) é considerado.

Outro recurso empregado é o uso de uma função penalidade associada aos pesos (HSIEH, 2001b) da rede neural. O propósito dessa técnica é limitar a capacidade não linear do NLPCA, que vem das funções de ativação da RNA. Algumas vezes uma “sensibilidade” excessiva as não linearidades encontradas nos dados analisados pode também tornar o método instável. Essa limitação é imposta adicionando-se um outro termo a equação de deve ser minimizada na análise, assim:

$$\min \left(\sum_{l=1}^n \|\vec{X}_l - \vec{X}'_l\| + p \sum_{i=1}^m (W_i^{(2)})^2 \right) \quad (1.5)$$

, onde p é o parâmetro de penalidade, $W^{(2)}$ é o conjunto de pesos da camada 2, e m é o número de neurônios dessa camada. Um valor alto de p aumenta concavidade da função, e força $W^{(2)}$ a assumir valores pequenos em magnitude. De acordo com a função de ativação utilizada, a magnitude dos pesos pode produzir um ou outro resultado. Na função

tangente hiperbólica, um peso w é muito pequeno, faz com que $\tanh(wx) \approx wx$, tornando a função praticamente linear. Com a função sigmoïdal, da mesma forma, valores de pesos grandes em magnitude, fazem com que a função torne-se constante, caracterizando um comportamento não linear. Desse modo, valores baixo para p fazem com que o NLPCA torne-se mais estável. Valores de p em torno de 1 são indicados (HSIEH, 2001b), mas esse ajuste deve variar de acordo com as funções de ativação usada e com os dados analisados. Vale ressaltar que apenas os pesos da camada de codificação (camada 2) são considerados, o que basta para limitar naturalmente a camada de decodificação (camada 4)

1.1.3 Trabalhos relacionados

Vários métodos já foram propostos para execução do PCA com a utilização de redes neurais artificiais, visando a superação de suas limitações, como em Oja (OJA, 1982) e Diamantaras e Kung (DIAMANTARAS; KUNG, 1996). Segundo Hsieh (HSIEH, 2001b), a análise de componentes principais não lineares (NLPCA) com o uso de RNAs foi proposta a primeira vez por Kramer (KRAMER, 1991) na literatura de engenharia química. Depois disso, o NLPCA tornou-se bastante usado em diversas áreas do conhecimento, e outros métodos foram propostos para sua execução através da utilização de RNAs autoassociativas, como descrito inicialmente por Kramer; análise das curvas principais (*principal curves analysis*), proposto por Hastie e Stuetzle (HASTIE; STUETZLE, 1982), e por Dong e McAvoy (DONG; MCAVOY, 1995); superfícies principais adaptativas (*adaptive principal surfaces*), descrito por LeBlanc e Tibshirani (LEBLANC; TIBSHIRANI, 1994); e otimização das entradas de redes neurais (*optimizing neural networks inputs*), apresentado por Tan e Mavrouniotis (TAN; MAVROVOUNIOTIS, 1995); foram alguns desses métodos, de acordo com Lee (LEE, 2000).

Existem ainda, variações sobre o método original proposto por Kramer (KRAMER, 1991), que em sua maioria mantêm como núcleo o algoritmo NLPCA alterando algumas características do processo. O S-NLPCA (*Sequential NLPCA*) é talvez a mais popular dessas variações; apresentada por Kramer no mesmo artigo em que apresenta o NLPCA (KRAMER, 1991), é muitas vezes citada indiscriminadamente como “NLPCA” na literatura, apesar de possuir algumas características próprias. Essas variações, assim como o H-NLPCA (*Hierarchical NLPCA*) (SCHOLZ; VIGÁRIO, 2002) e o C-NLPCA (BEM; BOTELHO; MATA, 2003), são apresentadas na Tabela 1.1.3 com suas respectivas características particulares. Ainda sobre o NLPCA, foram estudados mais alguns trabalhos, como o apresentado por Malthouse, Mahn e Tamhane (MALTHOUSE; MAH; TAMHANE,

1995), onde o método é relacionado com a análise das curvas principais; e o artigo de Malthouse (MALTHOUSE, 1998), no qual são apresentadas algumas de suas limitações.

Já na etapa de implementação, foram averiguados trabalhos e algoritmos para otimização do aprendizado em RNAs (algoritmos de minimização de funções), como aqueles apresentados por Bortoletti e outros (BORTOLETTI et al., 2003); e por Byrd e outros (BYRD et al., 1994). Esses algoritmos são, na verdade, melhorias propostas sobre o BFGS (Broyden-Fletcher-Goldfarb-Shanno), um algoritmo *quasi-Newton* muito popular e amplamente utilizado para otimização (minimização) multidimensional de funções. Foi utilizado nas implementações dos métodos NLPCA e C-NLPCA o BFGS, conforme apresentado por Press e outros (PRESS et al., 1990). É também importante ressaltar os recursos empregados na implementação, com o objetivo de aumentar a estabilidade da análise, como a utilização de uma **função custo** associada ao erro quadrático médio (*MSE*) das RNAs e o uso de uma **função penalidade** associada aos pesos das redes, como apresentado por Hsieh (HSIEH, 2001b), e utilizado em sua implementação do método NLPCA (HSIEH, 2001a).

Por fim, salientar-se que o método NLPCA de Kramer, e métodos relacionados têm sido aplicados com sucesso em muitos problemas diferentes (MALTHOUSE, 1998). Esses métodos são utilizados em várias áreas do conhecimento, desde a administração e *marketing*, como abordado por Hair e outros (HAIR et al., 1995), até a física, para estimativa da trajetória de partículas em um acelerador, como mencionado por Malthouse (MALTHOUSE, 1998), passando pelos problemas que têm relação com a extração de padrões de movimento de séries de imagens.

	NLPCA	S-NLPCA	H-NLPCA	C-NLPCA
Camada de gargalo	n neurônios	1 neurônio	n neurônios	n neurônios
Distinção do padrões	NÃO há hierarquização dos padrões de acordo com sua influência na variação total dos dados	a camada de gargalo, com apenas um neurônio, possibilita a identificação de cada padrão por ordem decrescente de influência na variação total	é possível identificar a importância de cada padrão de variação	permite a distinção dos padrões quando é executado com 1 neurônio na camada de gargalo
Reconstrução dos dados	dados originais são reconstruídos diretamente na saída da RNA	reconstrução direta	reconstrução direta	a reconstrução dos dados apresentados como entrada requer um estágio de expansão, como mostrado na Seção 2.2
Comentários	aplicado a problemas com número limitado de variáveis de entrada, devido a limitações intrínsecas às RNAs; nos quais a identificação dos padrões de variação não é importante (ex.: reconhecimento de faces, reconhecimento de gestos, etc.)	indicado geralmente para problemas onde se faz uma análise de natureza exploratória, na qual a busca e distinção de cada padrão de variação é fundamental (ex.: análise de comportamento de consumidores, identificação de fatores que influenciam mudanças climáticas, etc.)	indicado para o tratamento de problemas em que há necessidade de mapeamento para camada de gargalo com r ($r > 1$) neurônios, com hierarquização dos padrões (ex.: reconhecimento de faces, com determinação dos pontos determinantes no processo)	essa abordagem é indicada quando existe a necessidade de análise de um número elevado de variáveis, já que o método não oferece restrições quanto a isso (ex.: análise de imagens)

1.2 Aplicações

A extração de padrões de movimento pode ser empregada para solução de uma série de problemas nas áreas de visão computacional, reconhecimento de padrões e robótica. Entre as principais aplicações estão: reconhecimento de gestos, transferência de habilidades (*skill transfer*) para agentes robóticos e virtuais, teleoperação, animação gráfica e interação homem-computador (*HCI*). Os trabalhos encontrados na bibliografia podem ser divididos em dois grupos; no primeiro estão aqueles que empregam os métodos PCA e NLPCA na solução dos problemas; e no segundo grupo estão os trabalhos sem uma relação direta com os métodos de análise, mas relacionados à extração de padrões de movimento.

Neste ponto, deve-se deixar claro que em alguns trabalhos a análise de dados através dos métodos PCA e NLPCA é referenciada como uma redução dimensional, por mapear dados de um espaço *p-dimensional*, para um espaço *r-dimensional*, sendo $p > r$. Na verdade, na maior parte dos casos, não há diferenças quanto à análise, entre redução dimensional, extração de padrões de variação dos dados (ex.: padrões de movimento) e determinação de componentes principais, mas sim quanto à formulação e caracterização dos problemas a serem resolvidos. Essas diferentes abordagens, provavelmente frutos da grande popularização desses métodos em várias áreas do conhecimento, complementam-se e confundem-se algumas vezes.

No primeiro grupo descrito acima, encontram-se trabalhos como o apresentado por Lee (LEE; XU, 1998)(LEE, 2000), no qual o autor emprega os métodos PCA e NLPCA como uma alternativa para modelagem de ações humanas, como realização de gestos, em um espaço de dimensões reduzidas. Nos experimentos, a extração dos componentes principais é realizada sobre os dados enviados por uma *cyberglove* usada na execução dos movimentos. A proposta de Lee pode ser utilizada, entre outras coisas para transferência de habilidades. Nessa área, Tatani e Nakamura (TATANI; NAKAMURA, 2003) apresentam uma topologia semelhante a adotada no C-NLPCA para transferência de habilidade para um robô humanóide de 20 DOF. Os dados analisados são enviados por sensores localizados nas juntas do robô.

Já na área de reconhecimento de gestos e padrões em imagens, pode-se citar o trabalho de Black e Jepson (BLACK; JEPSON, 1998), que propõem o uso do PCA para realização de *tracking* e *matching* de objetos articulados; o reconhecimento de faces com uso do PCA, proposto de Romdhani (ROMDHANI,), e por Sim e outros (SIM et al., 2000); o modelo de redução dimensional não linear baseado no NLPCA de Kramer, para compressão de

imagens, proposto por DeMers e Cottrell (DEMERS; COTTRELL, 1993); e o método para execução do NLPCA sobre padrões visuais, como objeto, faces e mãos, proposto por Chalmond e Girard (CHALMOND; GIRARD, 1999).

No segundo grupo de trabalhos estudados destacam-se aqueles relacionados ao reconhecimento de gestos e ao *tracking* e modelagem de objetos e mãos em imagens. Dentre eles pode-se citar a proposta de Campos (CAMPOS, 2003) para modelagem de mãos e objetos no espaço 3D, com o objetivo de reconhecer intenções de movimento; o trabalho apresentado por Dorner (DORNER, 1994) mostra um interessante método para reconstrução de gestos no espaço 3D virtual, a partir do *tracking* de mãos em séries de imagens no espaço 2D. A autora propõe o uso de uma luva com marcas coloridas para estabelecer a correspondência entre pontos da mão real e do modelo virtual 3D. Outros trabalhos apresentando diferentes abordagens para modelagem de movimentos (COHEN et al., 1997)(LIN; WU; HUANG, 2000) e para reconhecimento de gestos foram analisados, mas as técnicas empregadas, como por exemplo, *hidden Markov models* (HMM) (STROBEL et al., 2002)(MACLEAN et al., 2001) e filtro de Kalman(KOHLER, 1997), estão fora do escopo deste trabalho.

Por fim, vale salientar que na extração de padrões de movimento de imagens com o método C-NLPCA, não é necessário que haja modelagem dos objetos ou movimentos analisado, o que torna o método passível de ser aplicado sem alterações em séries de imagens contendo qualquer tipo de movimento. Além disso, um fato importante é que não foram encontrados trabalhos na bibliografia estudada, que possibilitassem a análise não linear, através do NLPCA, de grandes conjuntos de variáveis (mais de 50), sem necessidade de uma pré-filtragem dos dados com o método PCA, como apresentado por Hsieh (HSIEH, 2001b) e por DeMers e Cottrell (DEMERS; COTTRELL, 1993). Tal tarefa torna-se possível com o método C-NLPCA.

2 C-NLPCA: Análise Cascadeada dos Componentes Principais

Quando se tem necessidade de executar o NLPCA sobre dados em uma dimensão de ordem elevada, há um aumento considerável no número de parâmetros (pesos+polinizadores) associados aos neurônios da RNA. Em consequência disso, há também a necessidade de utilização de uma maior quantidade de amostras temporais na análise, a fim de que esse número fique próximo (ou menor) do número de parâmetros da RNA (HSIEH, 2001b). Sabe-se que, dependendo da aplicação, algumas vezes não é possível atender a esse requisito de maneira razoável, mesmo sob o risco de saturação da RNA e não obtenção da redução dimensional. Além disso, a introdução de mais amostras aumenta de maneira expressiva as demandas por capacidade computacional e por tempo de processamento para a execução da análise.

Assim, quando o conjunto de dados original tem muitas dimensões (variáveis), muitos autores optam por filtrar o conjunto antes de realizar a análise pelo método NLPCA, usando técnicas para redução, como o PCA (BOTELHO et al., 2003; HSIEH, 2001b). Mas, devido ao fato do método PCA ser linear, como visto anteriormente, uma simplificação indesejada da variabilidade existente nos dados pode estar sendo introduzida na análise.

Dessa forma, com o C-NLPCA tem-se a pretensão de executar a análise direta e totalmente não linear de dados em dimensões de ordem elevada, com o uso de um conjunto de NLPCAs cascadeados sucessivamente, ver Figura 6. A arquitetura é composta por dois estágios: redução e expansão. Os dados (imagens) são decompostos em uma série de pequenas janelas, as quais serão reduzidas e agrupadas sucessivamente pelas redes NLPCA no **estágio de redução**. Ao final desse estágio, o neurônio de gargalo da última rede NLPCA fornecerá o primeiro componente principal (C-NLPC). No segundo estágio (**estágio de expansão**), este valor é expandido, resultando em uma saída na mesma dimensão dos dados originais de entrada.

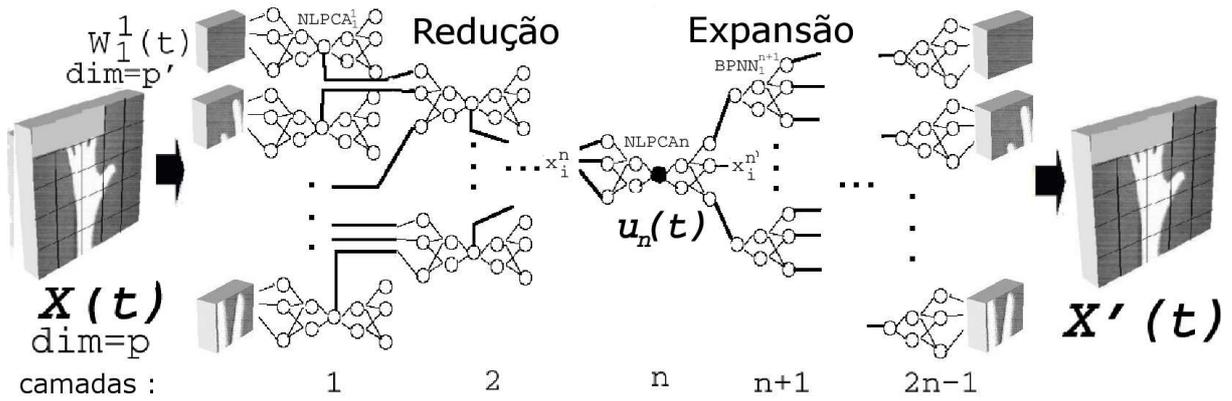


Figura 6: C-NLPCA: um conjunto de camadas agrupadas nos estágios de Redução e Expansão. O neurônio destacado fornece o padrão associado à redução do conjunto de dados original.

2.1 Obtendo o C-NLPC - Estágio de Redução

O C-NLPCA assume que p' é a dimensão *ideal* para os dados de entrada. O conceito de *ideal* é associado ao relacionamento entre o número de parâmetros (pesos + polarizadores) (Equação 1.3) e o número de amostras (observações) temporais (HSIEH, 2001b). Assim, divide-se a imagem original de entrada com dimensão p em janelas menores com dimensão p' . Essas janelas são usadas diretamente como entrada da primeira camada de NLPCAs. Cada NLPCA, limitado pelo critério de saturação, mencionado na seção 1.1.2.1, encontra o componente principal local (redução local) de uma janela. Os padrões resultantes (reduções) são usados como entrada para uma nova camada do C-NLPCA. Este processo é repetido n vezes até que reste apenas um padrão, configurando assim a redução final do conjunto inteiro de dados. Apesar de, no primeiro passo as janelas serem analisadas independentemente umas das outras, no segundo passo as relações de vizinhança são consideradas, uma vez que os resultados são agrupados sucessivamente.

Considera-se a janela $\vec{W}(t)$ como um subconjunto de pixels consecutivos do *vetor imagem*, $\vec{X}(t)$:

$$\vec{X}(t) = [\vec{W}_i(t) \mid 1 < i < m_1 = p/p'], \quad (2.1)$$

com:

$$\vec{W}_i(t) \cap \vec{W}_j(t) = 0, \forall i \neq j. \quad (2.2)$$

Para cada componente $\vec{W}_i(t)$ de $\vec{X}(t)$, a análise não linear é executada (seção 1.1.2.1)

com uma rede NLPCA “padrão”. Cada rede associada com $\vec{W}_i(t)$ é chamada $NLPCA_i$. Elas compõem a primeira camada com $m_1 = p/p'$ redes. Os resultados $u_i(t)$ (fornecidos pelos gargalos) de cada janela $NLPCA_i$ são agrupados em uma segunda camada de NLPCAs. Cabe salientar que, o número de NLPCAs da segunda camada é $m_2 = m_1/p'$ ¹.

O processo é repetido até a camada n , composta por apenas uma rede NLPCA. A saída do neurônio de gargalo dessa única rede NLPCA é um componente principal, $u_n(t)$ do conjunto de dados original.

2.2 Obtendo os C-NLPCAs - Estágio de Expansão

O segundo papel da análise de componentes principais, chamado de expansão, é obter os dados associados a cada componente principal (PC, C-NLPC) na dimensão original da imagem de entrada (PCA, C-NLPCA). Além disso, quando um componente principal k é calculado, e o componente $k + 1$ é desejado, o processo de expansão é também necessário para o cálculo dos resíduos associados a k , os quais serão a entrada para a obtenção do C-NLPC($k + 1$).

Devido ao processo de cascadeamento, a dimensão dos dados de entrada é perdida, sendo então necessário um método para obter a expansão, a partir da redução. O Estágio de Expansão tem como objetivo expandir o componente principal não linear, $u_n(t)$, em um conjunto de imagens que é a reconstrução da série de dados originais, levando-se em conta apenas o componente principal $u_n(t)$.

As camadas de expansão são simétricas posicionadas em relação às camadas de redução, resultando em um total de $(2 * n - 1)$ camadas. Elas são compostas por redes “backpropagation” simples *BPNN* (sem neurônio de gargalo). A entrada de cada *BPNN* é a saída da camada anterior do C-NLPCA. A metodologia adotada no presente trabalho foi de não re-treinar as *BPNN*, de modo que as entradas são apenas propagadas através das mesmas, que utilizam os mesmos pesos usados no Estágio de Redução.

Propagação nas *BPNN*: Cada saída resultante x_i^m da camada de gargalo NLPCA n é usada como entrada para propagação na respectiva $BPNN_i^{(n+1)}$ na próxima camada $(n + 1)$. Esse processo é repetido até a camada $(2n - 1)$. Os pesos são os mesmos usados nas camadas de decodificação e saída das redes NLPCAs utilizadas para redução dos

¹Os sub-índices são usados para descrever, respectivamente, a primeira e a segunda camada do cascadeamento).

dados (Estágio de Redução).

Assim, a imagem expandida é a reconstrução da entrada original levando-se em conta apenas o componente principal atual. Todos os componentes do conjunto original de dados são considerados, assim como suas relações de vizinhança e as variações temporais. O método pode ser aplicado independentemente da dimensão do conjunto de dados. A não linearidade associada as RNAs também é mantida, evitando a restrição de saturação inerente as mesmas.

3 *Experimentos realizados*

Com o objetivo de avaliar a extração de padrões de movimento de séries de imagens digitais 2D por meio do método C-NLPCA, foram realizados alguns experimentos que serão descritos neste capítulo. O método C-NLPCA (assim como o PCA e o NLPCA) pode ser aplicado na análise de qualquer tipo de movimento, sem que haja necessidade de modelagem dos objetos movimentados ou dos próprios movimentos realizados. Assim, uma grande quantidade de movimentos poderia ser utilizada nos experimentos, como por exemplo, movimentos faciais, movimentos da própria câmera (móvel) em um ambiente estático ou dinâmico, movimentos corporais, movimentos realizados por um agente robótico equipado com sensores, etc. Optou-se pela apresentação da análise de gestos e movimentos executados com mãos e capturados por uma câmera digital estática.

Foram capturados e analisados um conjunto de 11 movimentos, sendo 9 gestos manuais, 1 gesto manual associado ao movimento de 1 objeto, e uma série de movimentos faciais. Dentre os testes realizados, foram escolhidos para apresentação dos resultados neste trabalho, os experimentos com 1 dos gestos manuais, e o gesto manual associado ao movimento de 1 objeto. Tal escolha deu-se pelo fato desses movimentos possuírem um componente não linear exercendo maior influência na variação dos dados (pixels), em relação aos demais.

O primeiro movimento apresentado consiste no movimento oscilatório (esquerda-direita) da mão do agente humano (movimento global), associado ao movimento de abrir e fechar a mão (movimento local) (Figura 7); já o segundo é a movimentação de um objeto com o uso de uma e posteriormente das duas mãos de um agente humano (Figura 8).

A análise e modelagem de movimentos manuais não é uma tarefa simples. Isso é evidenciado pelos modelos de mãos encontrados na bibliografia, contendo em torno de 30 graus de liberdade (*30 DOF*) (LIN; WU; HUANG, 2000) (CAMPOS, 2003). Embora exista um grande número de dimensões nas representações resultantes dos gestos capturados, é possível mapear essas representações para um espaço de dimensão muito mais reduzida,

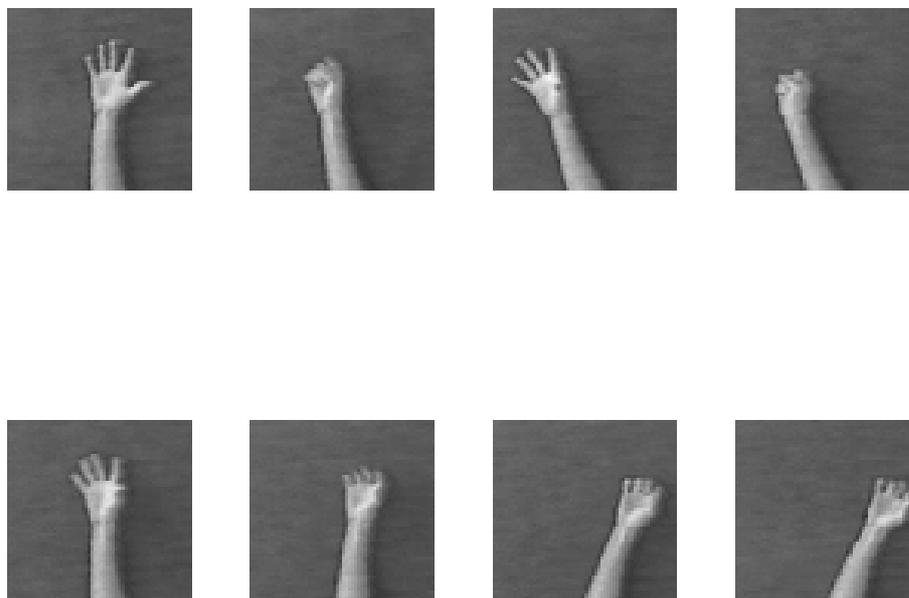


Figura 7: Alguns *frames* da primeira série de 128 imagens analisada.

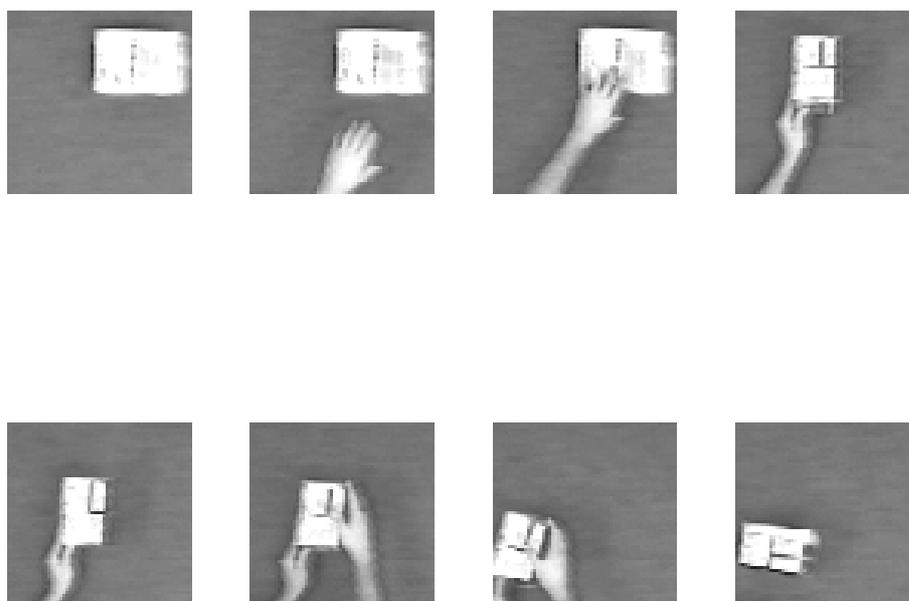


Figura 8: Alguns *frames* da segunda série de 128 imagens analisada.

como já descrito no capítulo anterior, e através desse mapeamento, extrair os padrões de movimento existentes nas séries de imagens.

As imagens contendo os gestos executados foram capturadas a uma frequência de 25 *frames* por segundo. O conjunto de imagens correspondete a cada gesto é composto por uma série de 128 imagens, cada uma com 60 pixels de largura (w), e 60 pixels de altura (h), totalizando 3600 ($w * h$) pontos (variáveis). Conforme descrito no Capítulo 2, cada uma dessas imagens pode ser vista como um *vetor imagem*, pertencente ao *espaço das imagens* $w * h$ -dimensional. O sistema de cores usados foi o HSV (*Hue/Saturarion/Value*), sendo que apenas a banda V, com 8 bits por pixel, foi analisada. Essa decisão foi tomada com o objetivo de desvincular o processo de análise da cor dos objetos, mantendo-o apenas sensível à iluminação do ambiente, que foi padronizada nos experimentos. Os dados foram normalizados antes de serem submetidos a análise.

Para avaliação do método C-NLPCA, foi utilizado como parâmetro de comparação a análise com o PCA sobre as séries de imagens. Nas próximas duas seções serão comparados os resultados de cada um dos métodos de análise, sobre as duas séries de imagens (movimentos) selecionadas, bem como uma discussão sobre os resultados obtidos.

3.1 Primeiro experimento

Nesta seção serão apresentados os resultados da análise sobre a primeira série de imagens selecionadas, contendo o movimento mostrado na Figura 7. Primeiramente será feita um comparação entre os padrões temporais de movimento (PCs ou *factor scores*) lineares, obtidos pela execução do método PCA, e não lineares, extraídos no Estágio de Redução do método C-NLPCA, como descrito na Seção 2. Logo após, serão apresentados os padrões espaciais de movimento, calculados com o método PCA, e as amostras “expandidas”, calculadas no Estágio de Expansão (Seção 6) do C-NLPCA, referentes aos valores máximos e mínimos de cada padrão temporal de movimento (C-NLPC). Por fim, será apresentada a reconstrução de uma das 128 amostras analisadas, com o uso do método linear e do método não linear. Será mostrado ainda, o erro (variação) inexplicado com o uso de cada método, relativo ao número de padrões utilizados na reconstrução da série original de imagens.

3.1.1 Padrões temporais de movimento - PCs

Os padrões temporais de movimento (variação), ou componentes principais (*principal components - PCs*), também conhecidos como *factor scores*, são compostos por uma série de valores (*scores*), cada um correspondente a uma das amostras (imagens) analisadas. Para o cálculo do *score* de cada amostra, são levados em consideração os valores de cada variável (pixels) na amostra, e os valores das projeções dessas variáveis sobre o novo eixo estabelecido. Assim, uma observação que tem valores altos em um grupo de variáveis que têm projeções com pesos consideráveis em um dado eixo, terá um *score* alto em relação a esse eixo. O PC mostra então, que uma determinada observação possui uma característica particular muito bem representada por esse eixo (padrão) (HAIR et al., 1995). O cálculo do PC foi apresentado anteriormente na Equação 1.1.

Nas Figuras 9, 10 e 11, respectivamente, é apresentado o primeiro, segundo e terceiro componente principal, calculado com cada um dos métodos. De acordo com a análise feita com o PCA, cada um deles é responsável por explicar 94,2%, 1,5% e 1,15%, respectivamente, da variação total encontrada nos dados. Pode-se observar a evolução temporal das amostras da série de imagens, em relação a cada um dos novos padrões espaciais de variação estabelecidos (novos eixos). De maneira mais específica em relação a série analisada, quando se tem um *score* alto para uma determinada amostra, significa que o posicionamento da mão na imagem aproxima-se da posição em que a mão aparece nos padrões espaciais, mostrados na seção seguinte. Caso contrário, tem-se um *score* baixo para a observação.

3.1.2 Padrões espaciais de movimento

Os padrões espaciais de movimento são, no método PCA, os autovetores da matriz de covariância da série de dados analisados, como mencionado na Seção 1.1.2. Esses autovetores são formados pelas projeções de cada uma das variáveis analisadas sobre os eixos da nova base estabelecida para representar os dados. Quando algumas variáveis têm projeções com valores elevados em um determinado padrão espacial, significa que essas variáveis têm grande influência na variação explicada por esse padrão. Com relação aos padrões mostrados nas Figuras 12 e 13, as variáveis mais significativas em relação a cada padrão são aquelas mostradas em tons mais claros e em vermelho escuro, respectivamente.

Quando a análise dos dados é realizada com o método C-NLPCA, os padrões espaciais de movimento não são explicitamente calculados. Apesar disso, uma aproximação desses

padrões pode ser obtida através da observação das amostras “expandidas”, no estágio de expansão do C-NLPCA, referentes aos valores máximos e mínimos dos PCs correspondentes. Isso se deve ao fato, já mencionado, de amostras com *scores* elevados possuírem uma alta correlação com os padrões espaciais, e vice-versa. As Figuras 14 e 15, mostram as observações correspondentes ao valor mínimo do PC de cada padrão de movimento; e as Figuras 16 e 17, as correspondentes ao valor máximo.

3.1.3 Reconstrução das observações

Utilizando o método PCA, depois de calculados os padrões temporais e espaciais de movimento (variação), é possível reconstruir o conjunto original de dados a partir dos mesmos. A equação usada para reconstrução dos dados, de acordo com Venegas (BJÖRNSSON; VENEGAS, 1997), é:

$$\vec{X}(t)' = \sum_{i=1}^{w * h} u_i(t) * a_i, \quad (3.1)$$

onde $\vec{X}(t)'$ é a reconstrução da matriz de dados original, $\sum_{i=1}^{w * h}$ é o somatório de todos os padrões de movimento existentes, $u_i(t)$ é o padrão temporal i (PC_i), e a_i é o padrão espacial (autovetor) i . Já no método C-NLPCA, a reconstrução $\vec{X}(t)'$ dos dados originais é obtida por meio da soma algébrica das amostras “expandidas” no estágio de expansão, na extração de cada padrão.

Nos experimentos realizados, os 3 primeiros padrões de movimento extraídos com o método PCA, são responsáveis por 97.24%, no primeiro experimento, e 96% no segundo, da variação total encontrada em toda série de imagens. Dessa forma, foram usados somente os 3 primeiros padrões para reconstrução das observações apresentadas nesta seção, assim a Equação 3.2 foi utilizada para reconstrução no método PCA, enquanto a Equação 3.3 foi utilizada no método C-NLPCA. As reconstruções calculadas são mostradas nas Figuras 18 e 19.

$$\vec{X}(t)' = u_1(t) * a_1 + u_2(t) * a_2 + u_3(t) * a_3 \quad (3.2)$$

$$\vec{X}(t)' = \vec{X}_1(t)' + \vec{X}_2(t)' + \vec{X}_3(t)' \quad (3.3)$$

3.1.4 Projeção das reconstruções das amostras no espaço 3D

Devido a dificuldade de visualização e comparação das 128 imagens originais, além de suas respectivas reconstruções pelo método PCA e C-NLPCA, uma outra forma de viabilizar essa tarefa com maior facilidade é através da projeção dos dados no espaço 3D. Nas Figuras 20, 21 e 22 são mostrados, respectivamente, a série de dados originais, a série de dados reconstruída com o PCA e a série reconstruída com o método C-NLPCA.

3.1.5 Percentual de erro inexplicado por cada padrão

Com o objetivo de avaliar a parcela de variação não explicada por cada um dos padrões de movimento extraídos pelos métodos PCA e C-NLPCA, foi utilizada a magnitude relativa dos resíduos, conforme apresentado por Lee (LEE, 2000):

$$\%Erro = \frac{\|\vec{X}(t) - \vec{X}(t)'\|}{\|\vec{X}(t)\|}, \quad (3.4)$$

onde $\vec{X}(t)$ é a matriz de dados originais, $\vec{X}(t)'$ é somatório dos n padrões de movimento utilizados para reconstrução dos dados, e $\|\cdot\|$ é a norma (L2) de Frobenius¹. A Figura 23 mostra a relação entre o percentual de variação inexplicada (ou erro) e o número de padrões usado na reconstrução dos dados.

3.1.6 Análise dos resultados

Analisando os padrões temporais de variação, constata-se que o primeiro padrão linear (PC1) tem um forte componente de alta frequência, mostrado na Figura 9, que não é encontrado no C-NLPC1, mostrado na mesma figura. Concluiu-se que essa variação de alta frequência é consequência da iluminação usada no local de realização dos experimentos, devido ao fato dessa frequência ser encontrada mesmo em trechos da série de imagens onde não havia movimento significativo da mão, como acontece nos primeiros *frames* capturados. Assim, o C-NLPCA desconsiderou o ruído existente na série de imagens, o que é uma característica muito interessante, encontrando como primeiro padrão temporal de variação (C-NLPC1) um “sinal” muito parecido com o PC2, mostrado na Figura 10. Já o C-NLPC2, também mostrado na Figura 10, agrega características do PC2 e PC3, este último visto na Figura 11.

¹ $\sqrt{(\sum \text{diag}(A^t * A))}$.

Pode-se afirmar, ao analisar os padrões C-NLPC1 e C-NLPC2, em conjunto com os padrões espaciais de variação (Figuras 12 e 13), com as amostras referentes aos seus pontos máximos e mínimos (Figuras 14, 15, 16 e 17) e com os dados originais, que os mesmos representam essencialmente a oscilação da mão nas imagens, da esquerda para direita², como era esperado. No C-NLPC1, os 3 vales encontrados correspondem aos trechos da série de imagens em que a mão estava na esquerda, e 2 picos correspondentes aos trechos nos quais esta encontrava-se na direita. Quando a mão encontra-se na região central das imagens, os *scores* assumem valores de transição entre os pontos máximos e mínimos, o que significa que as variáveis (pixels) do centro das imagens possuem uma projeção de valor intermediário sobre o novo eixo encontrado. Já no padrão C-NLPC2, os 3 picos mais baixos correspondem ao posicionamento da mão à esquerda, enquanto os 2 picos de maiores amplitude e período, correspondem à mão posicionada à direita. O movimento de abrir e fechar a mão é representado por um sinal de mais alta frequência, que pode ser visto como um “ruído”, localizado geralmente nos vales e picos de ambos padrões.

É necessário ressaltar que, diferente do primeiro padrão, que representa a variação principal encontrada na série de imagens, o C-NLPC2 foi capaz de captar as peculiaridades do movimento, visto por exemplo, por meio das diferenças constatadas entre os picos representando “esquerda” e “direita”. Essas diferenças foram encontradas porque nos movimentos para direita, a mão ultrapassou os limites da imagem, como mostrado na Figura 24.

O C-NLPC3, observado na Figura 11, mostrou-se um padrão bem diferente dos demais. Observando os padrões de mais alta ordem extraídos com o método PCA, vistos na Figura 25, verifica-se que o vale encontrado no C-NLPC3, também é encontrado no PC7, em destaque. Isso deve-se aos parâmetros usados para execução do C-NLPCA nesse experimento, que deram mais “sensibilidade” à análise. Esse vale corresponde justamente a imagem mostrada na Figura 24, que constitui uma singularidade da série de dados, e não deveria dominar o padrão C-NLPC3, como constatado. No entanto, esse fato não se caracteriza como um erro na análise, e sim uma questão de ajuste de parâmetros.

Apesar do problema abordado no parágrafo anterior, o gráfico da Figura 23 mostra 19.35% de redução no erro (variação) inexplicado pelo C-NLPC1, em comparação ao PC1, e 3.7%, quando o terceiro padrão foi considerado. Mesmo assim há necessidade de um ajuste mais refinado nos parâmetros, como pode ser constatado pela observação das expansões das amostras referentes aos mínimos e máximos *scores* dos padrões 2 e 3,

²Ponto de vista do observador.

(Figuras 14, 15, 16 e 17); da reconstrução de uma das amostras, vista na Figuras 18 e 19; e das projeções no espaço 3D (Figuras 20, 21 e 22), onde a reconstrução com o C-NLPCA captou alguns picos existentes na série de dados originais, o que demonstra sua capacidade de lidar com não linearidades, ao contrário do PCA.

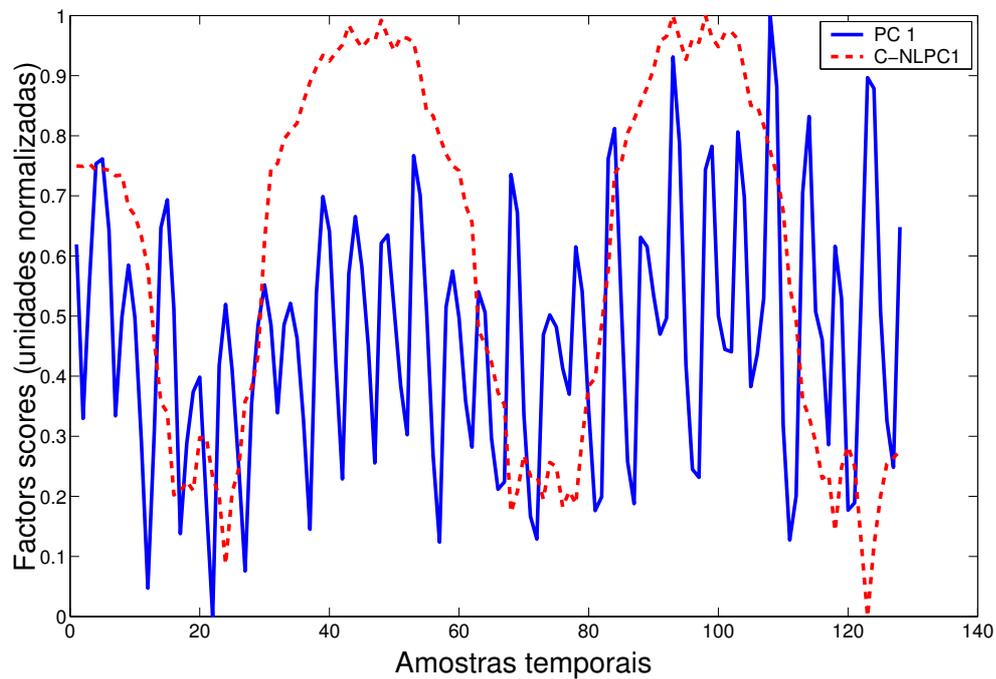


Figura 9: Comparação entre PC1 e C-NLPC1.

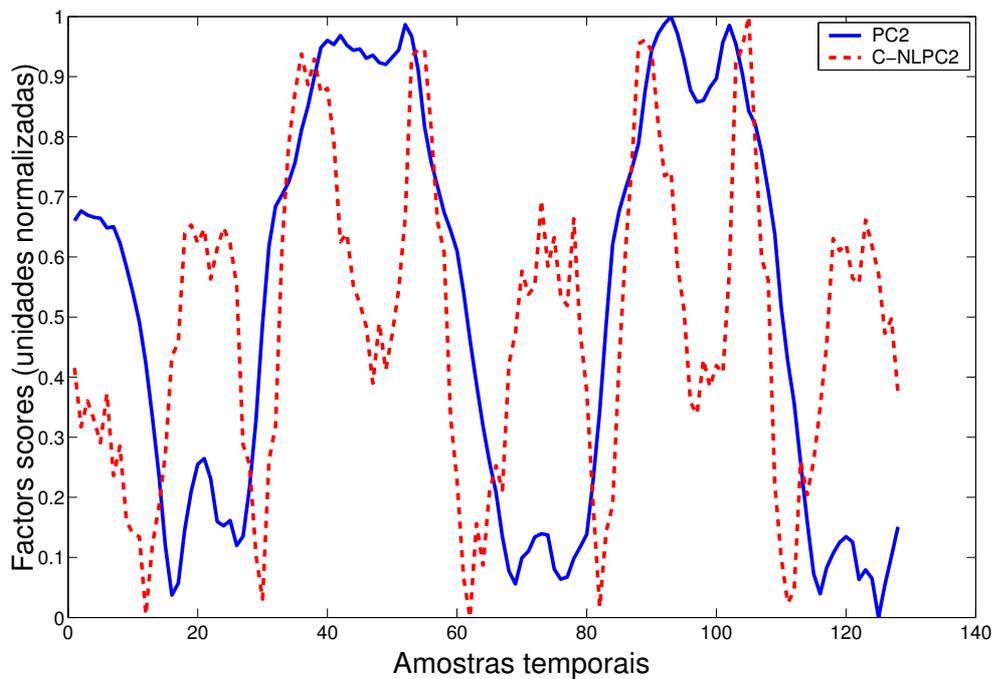


Figura 10: Comparação entre PC2 e C-NLPC2.

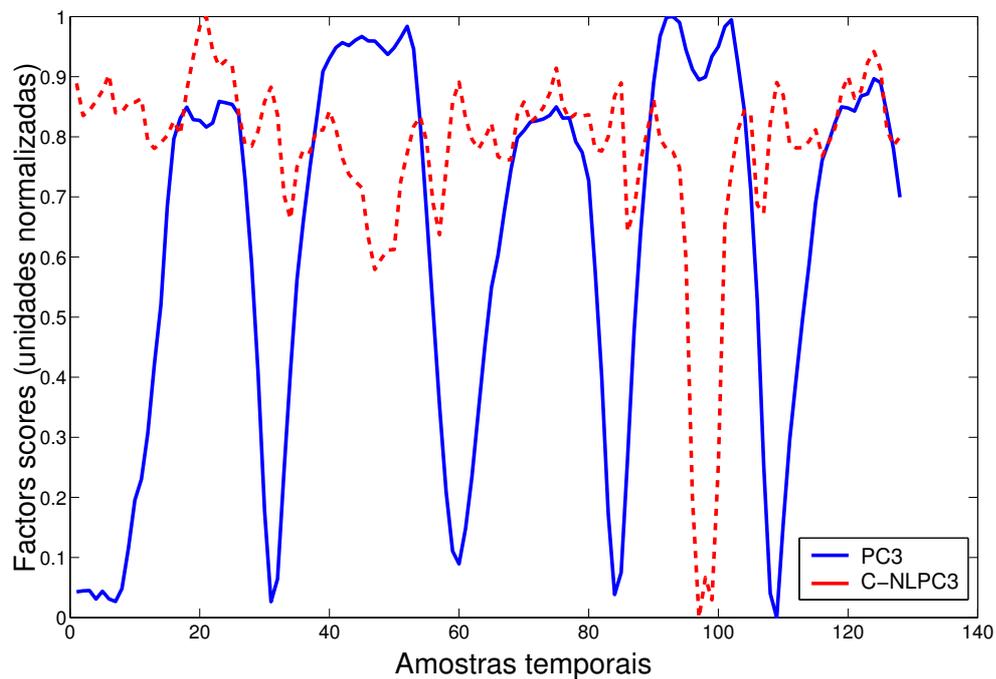


Figura 11: Comparação entre PC3 e C-NLPC3.

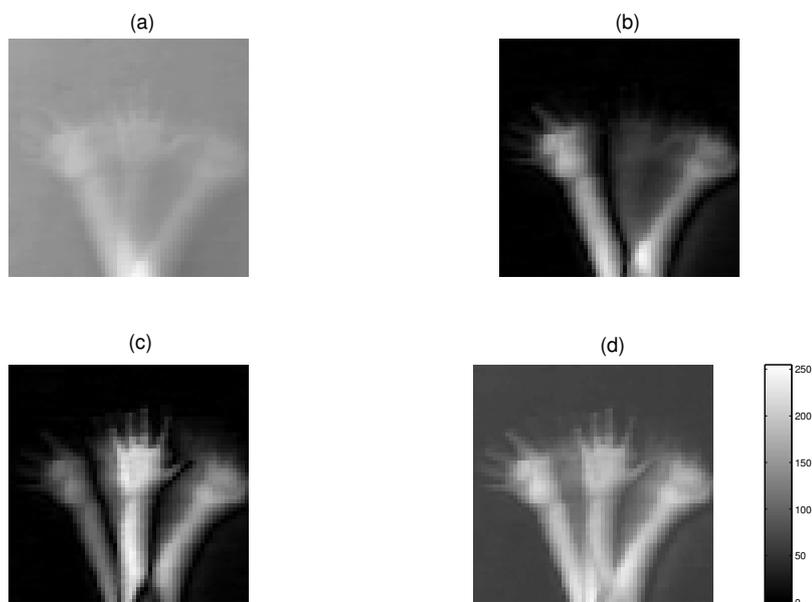


Figura 12: Padrões espaciais lineares em tons de cinza. (a) Primeiro padrão. (b) Segundo padrão. (c) Terceiro padrão. (d) Soma dos 3 padrões lineares. A escala utilizada é mostrada ao lado da imagem (d).

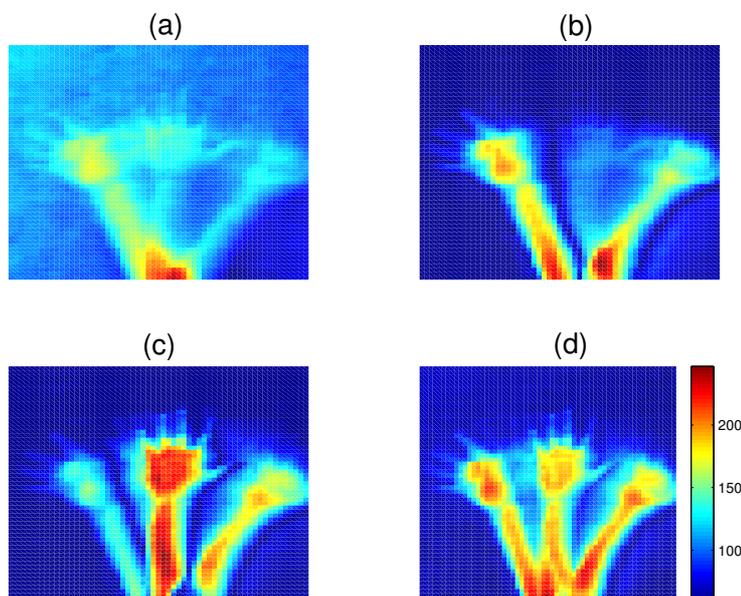


Figura 13: Padrões espaciais lineares em *pseudocolor*. (a) Primeiro padrão. (b) Segundo padrão. (c) Terceiro padrão. (d) Soma dos 3 padrões lineares. A escala utilizada é mostrada ao lado da imagem (d).

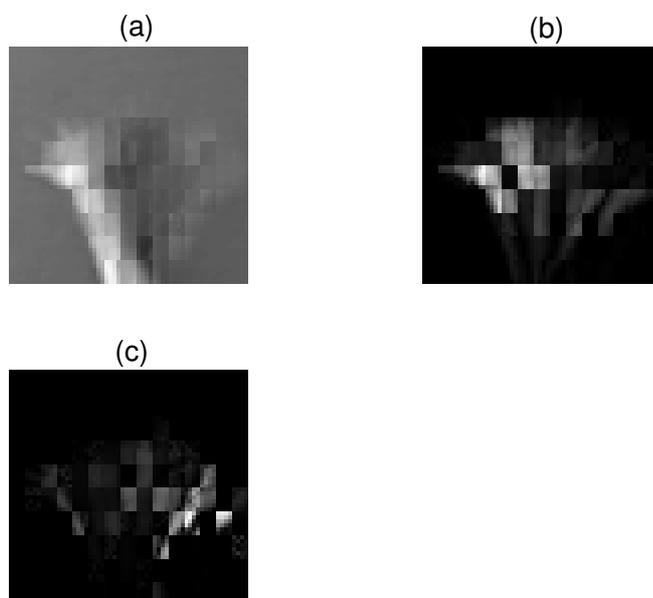


Figura 14: Expansão das observações referentes ao valor mínimo do C-NLPC de cada padrão, em tons de cinza. (a) Primeiro padrão. (b) Segundo padrão. (c) Terceiro padrão. A escala utilizada é a mesma mostrada na Figura 12.

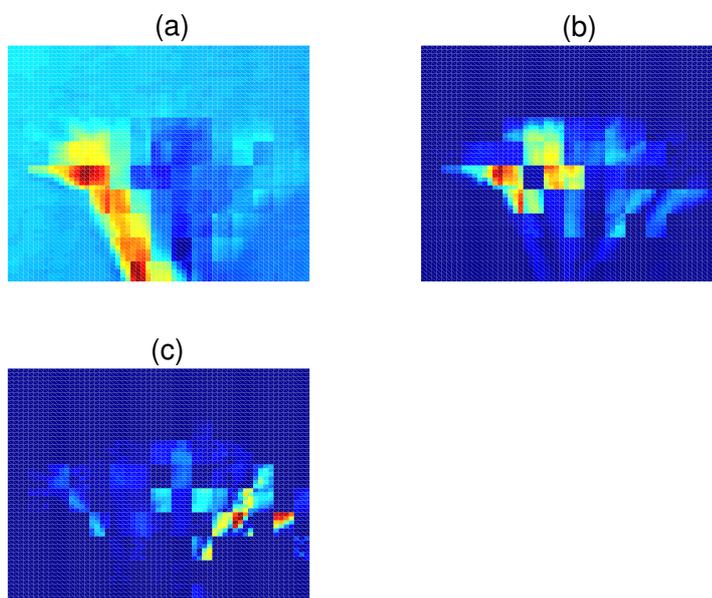


Figura 15: Expansão das observações referentes ao valor mínimo do C-NLPC de cada padrão, em *pseudocolor*. (a) Primeiro padrão. (b) Segundo padrão. (c) Terceiro padrão. A escala utilizada é a mesma mostrada na Figura 13.

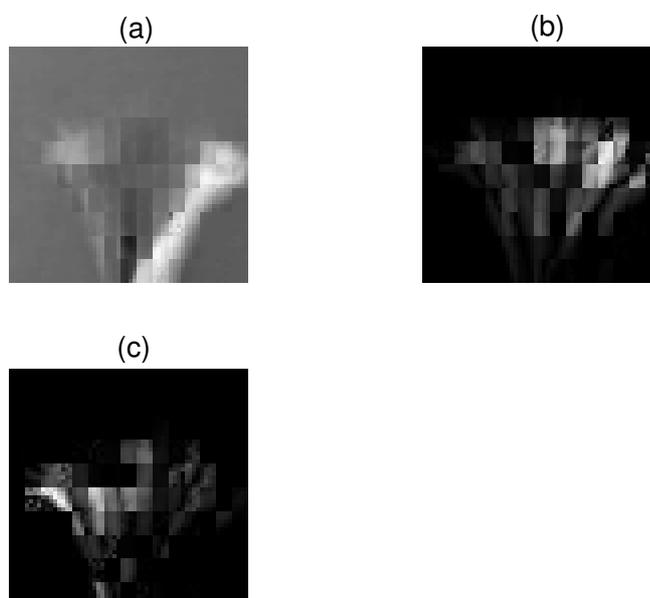


Figura 16: Expansão das observações referentes ao valor máximo do C-NLPC de cada padrão, em tons de cinza. (a) Primeiro padrão. (b) Segundo padrão. (c) Terceiro padrão. A escala utilizada é a mesma mostrada na Figura 12.

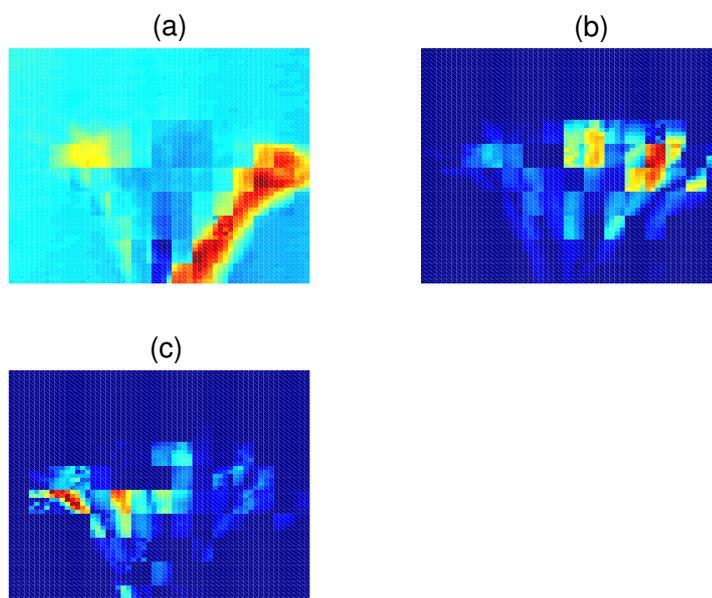


Figura 17: Expansão das observações referentes ao valor máximo do C-NLPC de cada padrão, em *pseudocolor*. (a) Primeiro padrão. (b) Segundo padrão. (c) Terceiro padrão. A escala utilizada é a mesma mostrada na Figura 13.

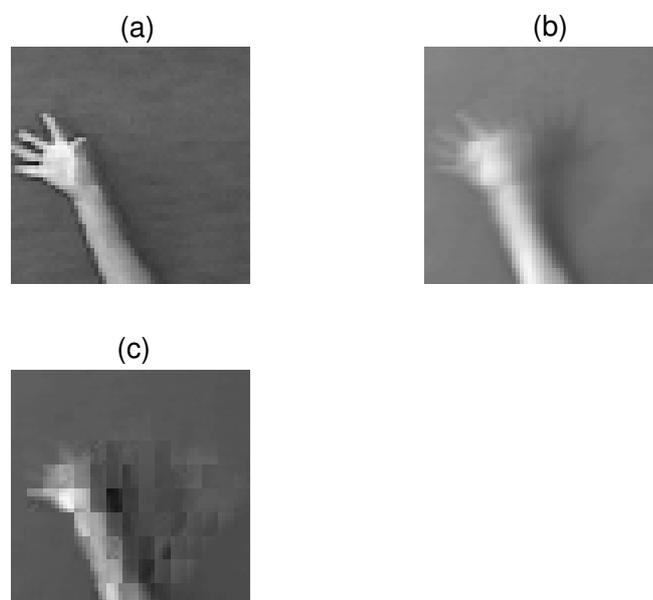


Figura 18: Reconstrução, a partir dos 3 primeiros padrões de movimento, de uma das 128 observações analisadas, em tons de cinza. (a) Imagem original. (b) Reconstrução com o método linear (PCA). (c) Reconstrução (expansão) com o método não linear (C-NLPCA). A escala utilizada é a mesma mostrada na Figura 12.

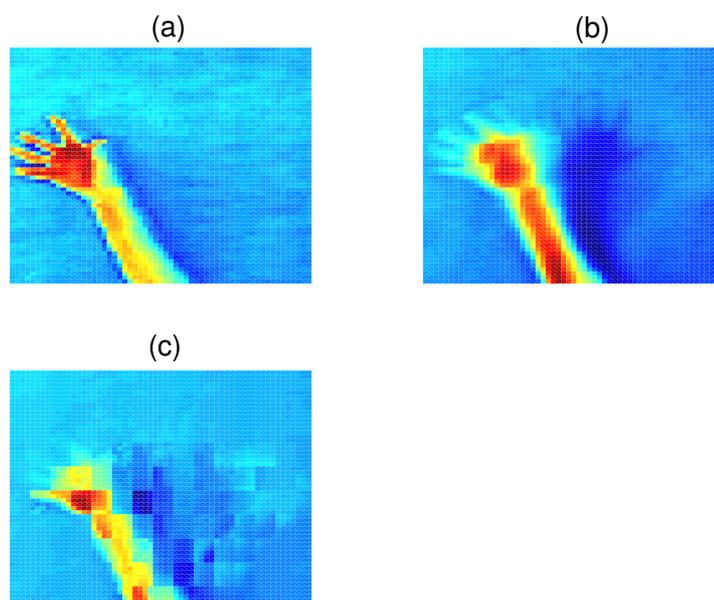


Figura 19: Reconstrução, a partir dos 3 primeiros padrões de movimento, de uma das 128 observações analisadas, em *pseudocolor*. (a) Imagem original. (b) Reconstrução com o método linear (PCA). (c) Reconstrução (expansão) com o método não linear (C-NLPCA). A escala utilizada é a mesma mostrada na Figura 13.

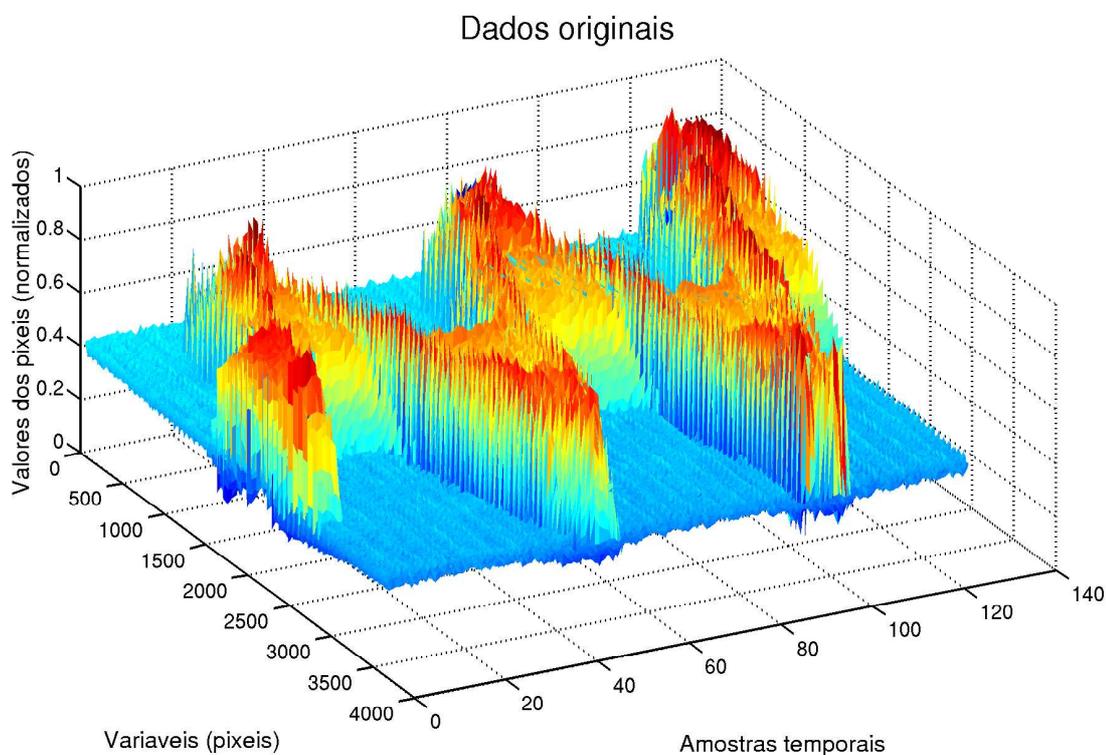


Figura 20: Evolução temporal do valor de cada variável (pixel) da série de imagens originais.

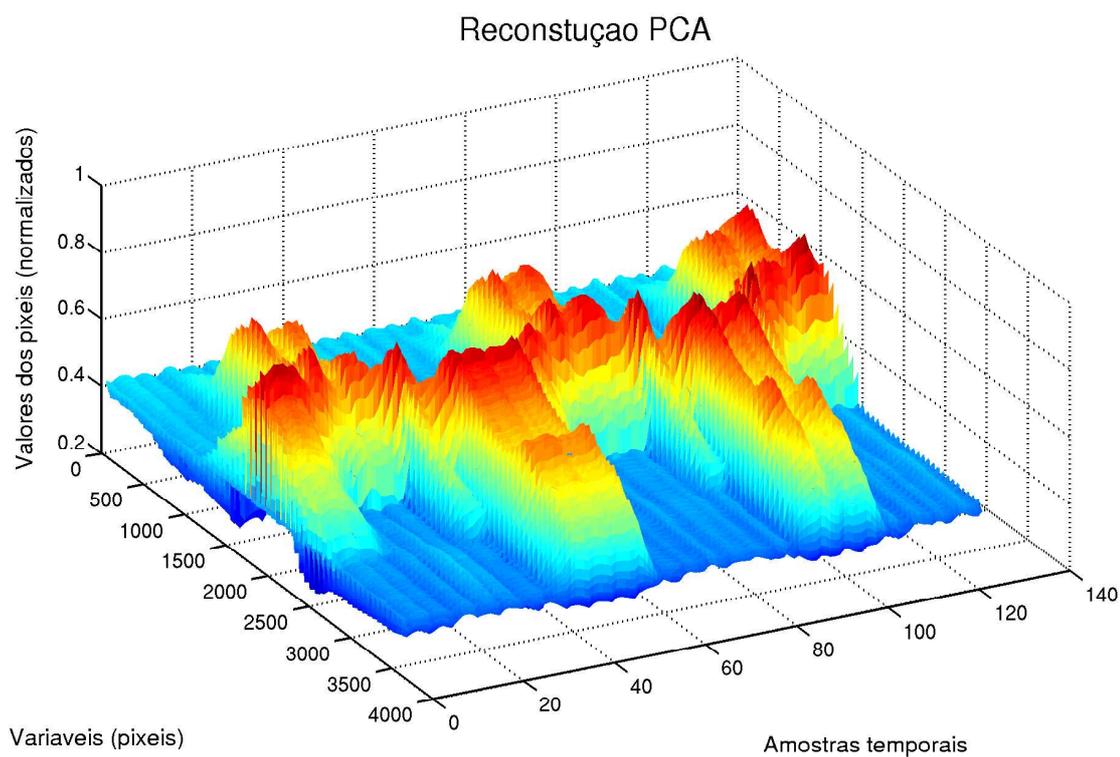


Figura 21: Evolução temporal do valor de cada variável (pixel) da reconstrução da série de imagens originais com o método PCA.

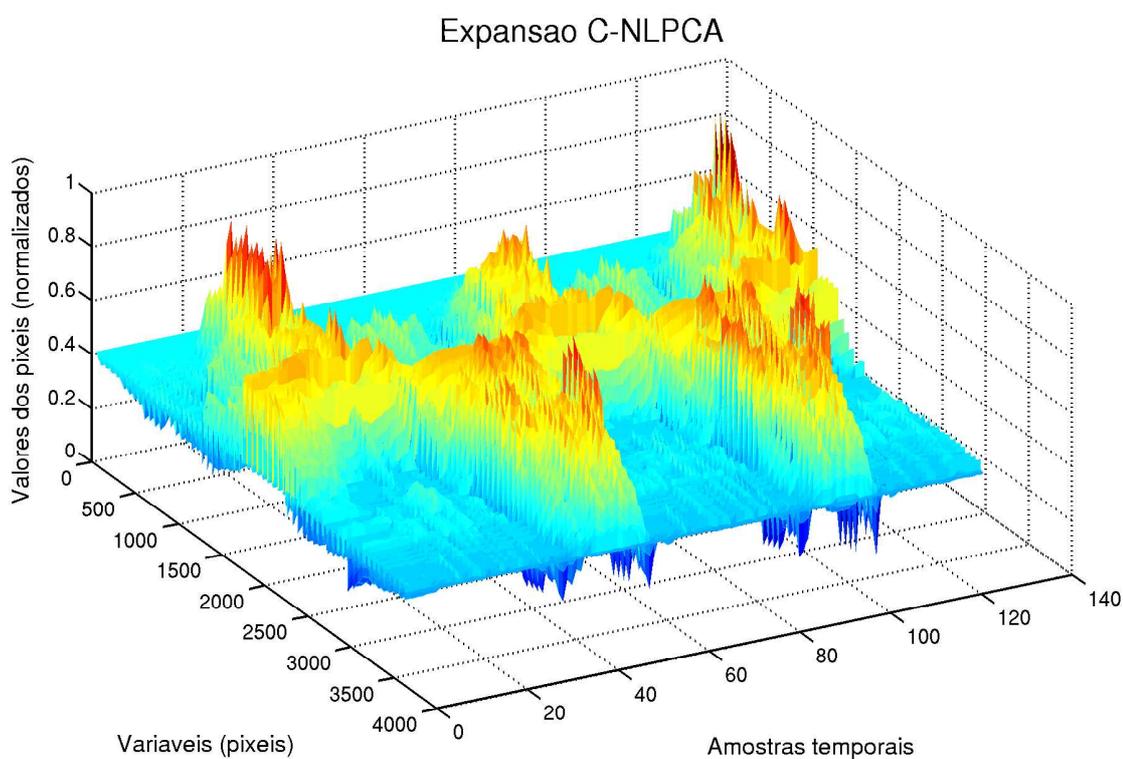


Figura 22: Evolução temporal do valor de cada variável (pixel) da reconstrução da série de imagens originais com o método C-NLPCA.

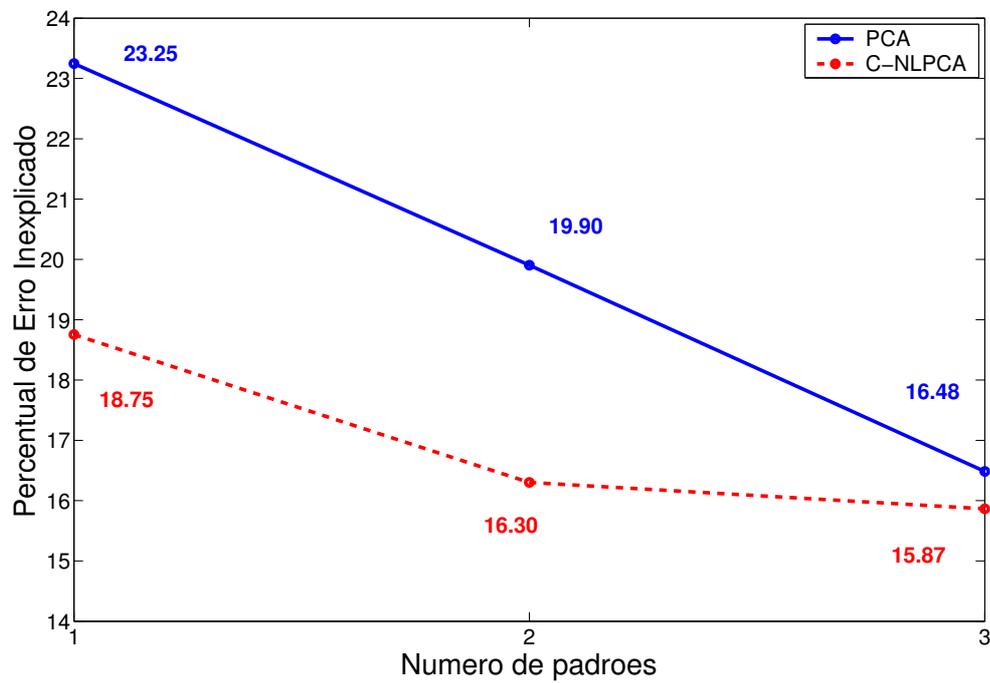


Figura 23: Percentual de erro inexplicado por cada método, de acordo com o número de padrões utilizados para reconstruir os dados.

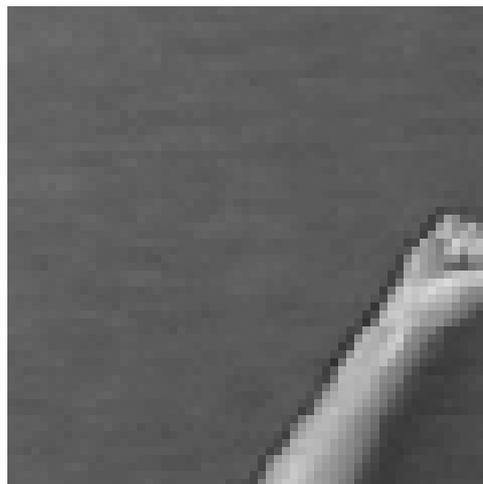


Figura 24: Mão ultrapassando os limites da imagem.

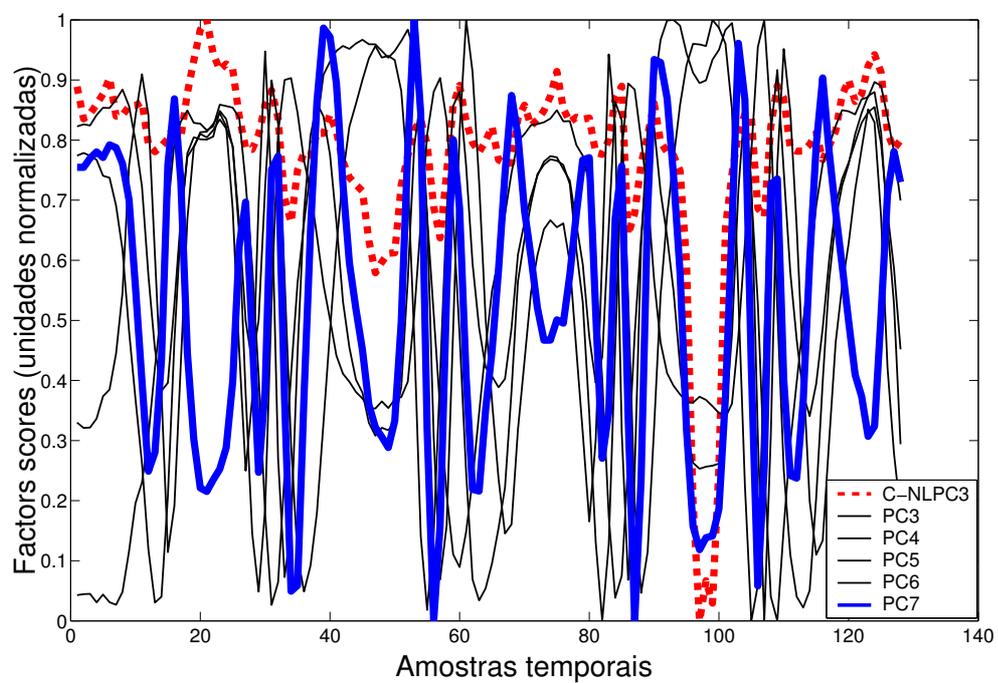


Figura 25: Comparação entre o C-NLPC3 e os PCs de mais alta ordem.

3.2 Segundo experimento

Nesta seção serão apresentados e discutidos os resultados da análise sobre a segunda série de imagens selecionadas, contendo o movimento mostrado na Figura 8. Todos os detalhes a respeito da metodologia empregada na obtenção dos resultados, já foram descritos na seção anterior.

3.2.1 Análise dos resultados

Neste experimento, com o método PCA foram encontrados os percentuais de 93%, 3% e 1.3%, respectivamente, de influência de cada componente principal na variação total existente nos dados. Semelhante ao constatado no experimento anterior, o primeiro padrão temporal linear de movimento (PC1) também apresenta neste caso, um forte componente de alta frequência, presente em toda sua extensão. O que realmente comprova tratar-se de um “ruído” gerado pela iluminação do ambiente. Nota-se também, nos padrões PC1 e C-NLPC1, a existência de um “degrau”, sendo que este último conseguiu minimizar consideravelmente a alta frequência encontrada no primeiro. O “degrau” em questão corresponde ao deslocamento do objeto encontrado nas imagens, de sua posição inicial, para posição final.

Destaca-se que, enquanto na análise linear os padrões PC1 e PC2 mostram o “degrau” já mencionado, na análise não linear apenas o C-NLPC1 apresenta essa característica. Isso significa que o primeiro padrão não linear foi capaz de “explicar” praticamente toda variação referente ao deslocamento do objeto nas imagens. Assim, além do padrão C-NLPC2 não apresentar o “degrau”, ele conseguiu caracterizar o surgimento de cada uma das mãos na série de imagens. Isso é mostrado com clareza na Figura 27, onde o pico do C-NLPC2 representa o aparecimento da primeira mão, Figura 41 (a); e o vale o aparecimento da segunda, Figura 41 (b). Esse movimento das mãos foi praticamente ignorado nos padrões lineares, a não ser por um leve pico encontrado no PC1, coincidente a um pico observado no C-NLPC1, Figura 26.

Também similar a “sensibilidade” excessiva na análise constatada no primeiro experimento, o padrão C-NLPC3 apresenta um pico exagerado, provavelmente fruto de especificidades encontradas nas amostras. A exemplo do ocorrido na análise da primeira série de imagens, isso pode ser percebido nas expansões referentes aos *scores* máximos e mínimos, especialmente no segundo e terceiro padrões, Figuras 31, 32, 33 e 34.

Nesse ponto, é relevante chamar a atenção para o aspecto fragmentado apresentado

pelas expansões citadas no parágrafo anterior, bem como para a mesma característica apresentada pela reconstrução de uma das amostras por meio do C-NLPCA, Figuras 35 e 36; e também para o ruído percebido na projeção da série de dados reconstruída com o C-NLPCA, Figura 39. Esses problemas surtem porque algumas RNAs encontram mínimos locais, ou especializam-se demais, no processamento das janelas de dados, o que afeta tanto o Estágio de Redução, quanto o de Expansão. No entanto, no primeiro deles, essas falhas acabam por “desaparecer” em meio aos padrões extraídos corretamente das demais janelas, visto que estes são agrupados sucessivamente no processo de análise. Por outro lado, na Expansão, os resultados de RNAs com “problemas” de mínimo local ou muito “especializadas”, mesmo que sejam em número reduzido, fatalmente aparecerão na reconstrução dos dados e prejudicarão o cálculo do resíduo para extração dos padrões de ordem mais elevada.

Neste experimento o ganho apresentado com o primeiro padrão foi perdido ao serem acrescentados o segundo e terceiro padrões na reconstrução dos dados, Figura 40. Isso ocorreu devido aos problemas já mencionados anteriormente.

Por fim, vê-se nas projeções das reconstruções no espaço 3D, que o C-NLPCA conseguiu representar bem algumas regiões não bem definidas pelo método PCA, o que denota a capacidade de lidar com não linearidades do primeiro método. Isso pode ser notado, por exemplo, na região em torno da amostra 80, nas Figuras 37, 38 e 39. Apesar disso, percebe-se uma grande quantidade de ruído presente na reconstrução C-NLPCA, o que fez com que aumentasse a magnitude do erro inexplicado pelos 3 padrões em conjunto.

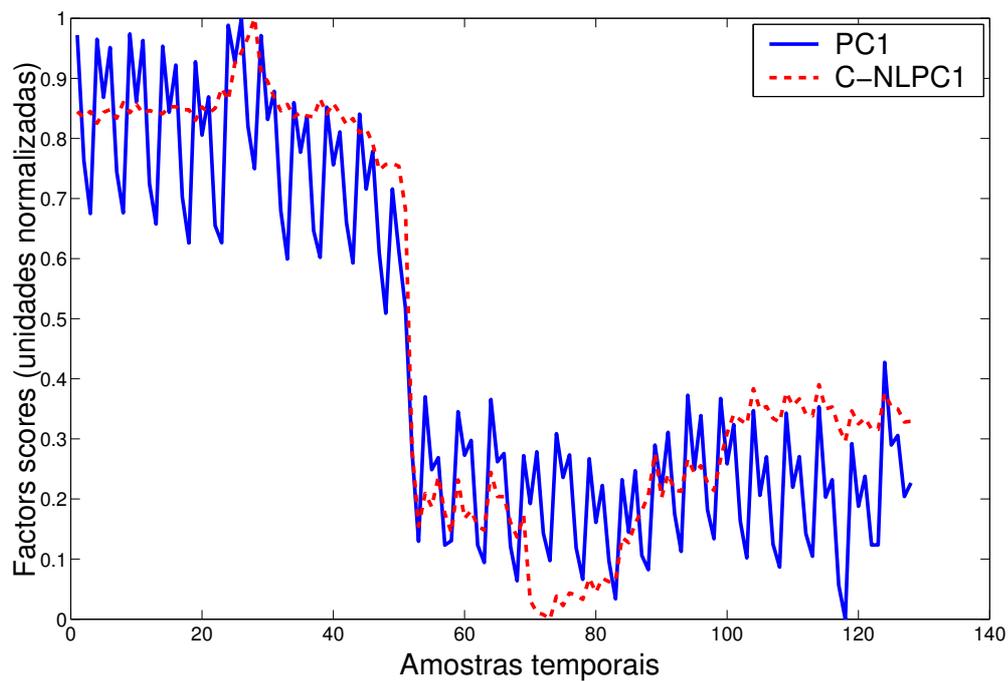


Figura 26: Comparação entre PC1 e C-NLPC1.

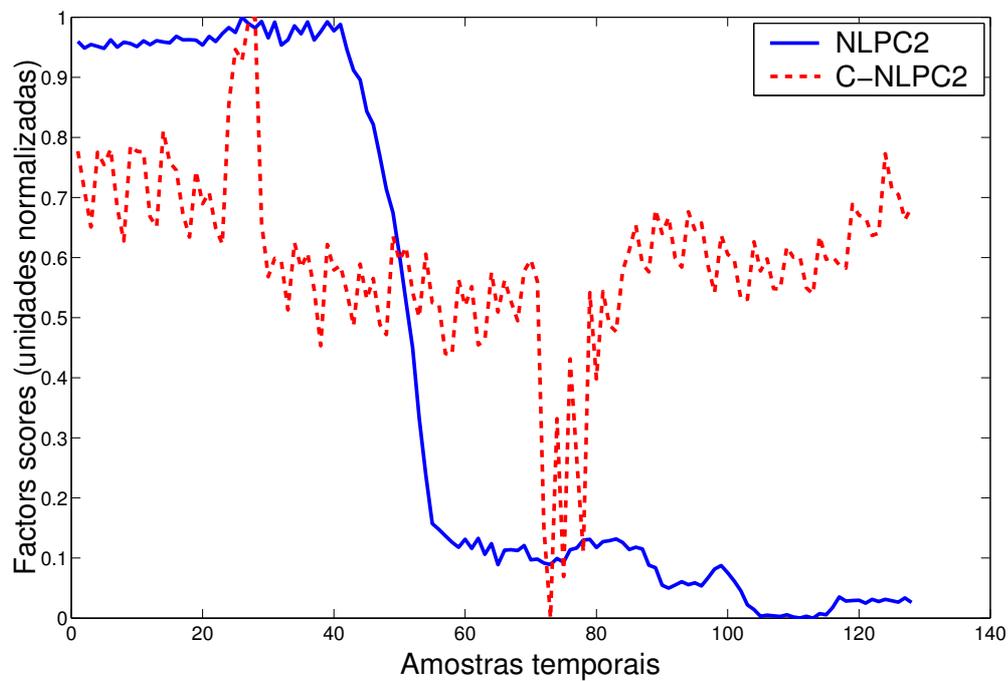


Figura 27: Comparação entre PC2 e C-NLPC2.

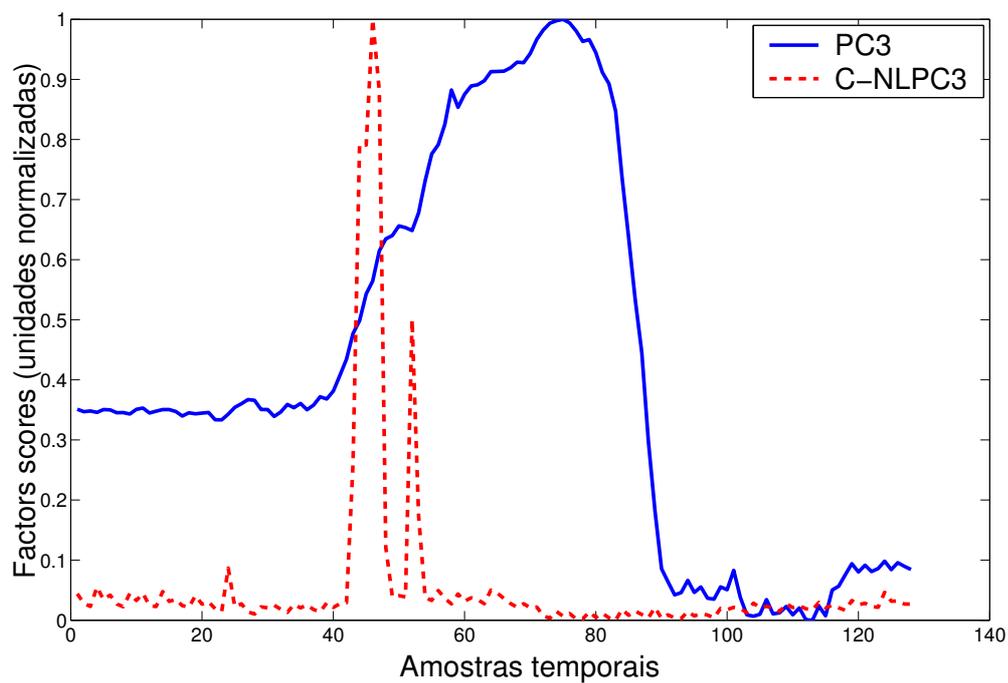


Figura 28: Comparação entre PC3 e C-NLPC3.

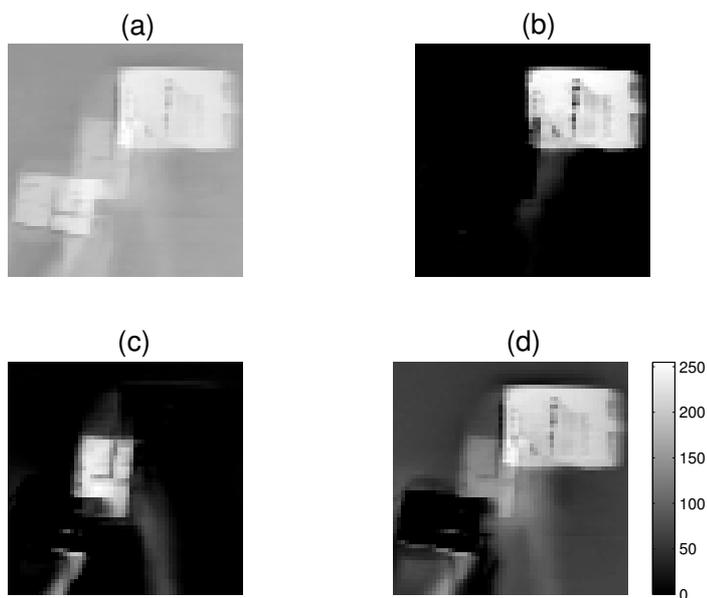


Figura 29: Padrões espaciais lineares em tons de cinza. (a) Primeiro padrão. (b) Segundo padrão. (c) Terceiro padrão. (d) Soma dos 3 padrões lineares. A escala utilizada é mostrada ao lado da imagem (d).

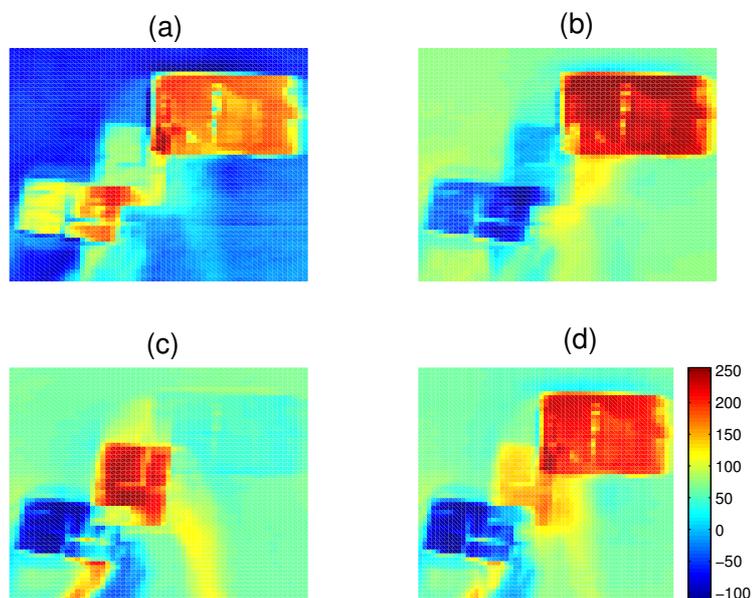


Figura 30: Padrões espaciais lineares em *pseudocolor*. (a) Primeiro padrão. (b) Segundo padrão. (c) Terceiro padrão. (d) Soma dos 3 padrões lineares. A escala utilizada é mostrada ao lado da imagem (d).

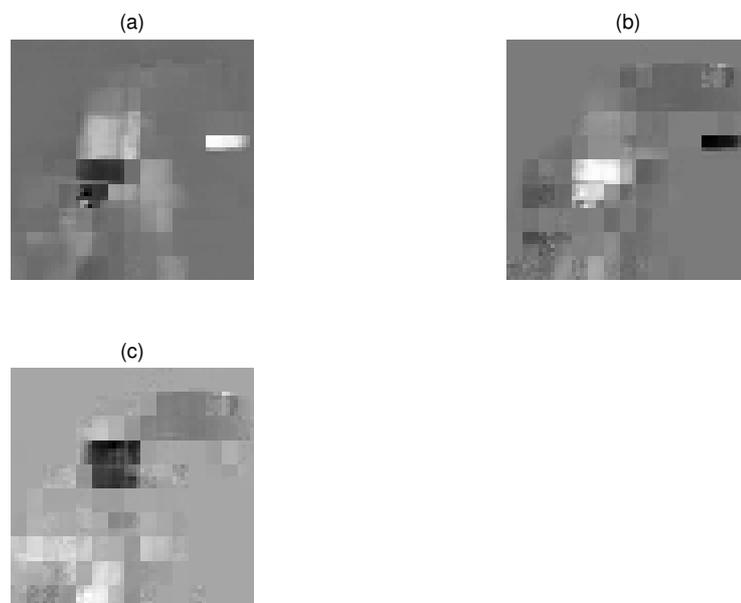


Figura 31: Expansão das observações referentes ao valor mínimo do C-NLPC de cada padrão, em tons de cinza. (a) Primeiro padrão. (b) Segundo padrão. (c) Terceiro padrão. A escala utilizada é a mesma mostrada na Figura 29.

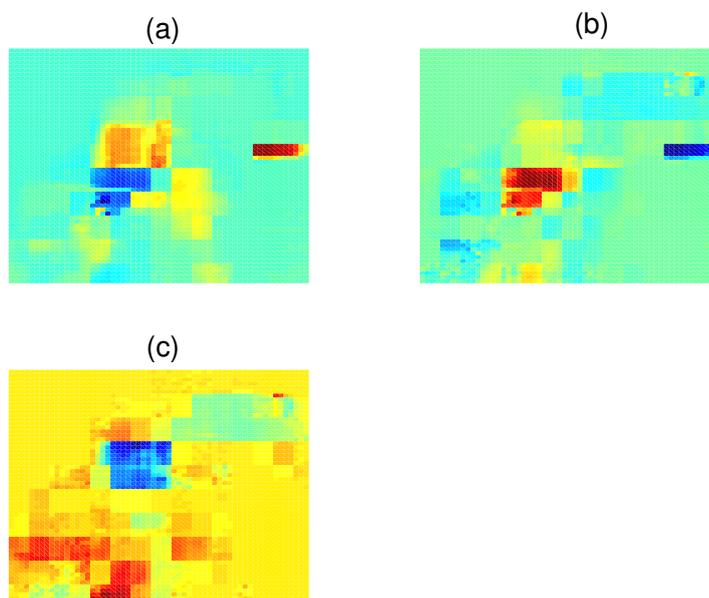


Figura 32: Expansão das observações referentes ao valor mínimo do C-NLPC de cada padrão, em *pseudocolor*. (a) Primeiro padrão. (b) Segundo padrão. (c) Terceiro padrão. A escala utilizada é a mesma mostrada na Figura 30.

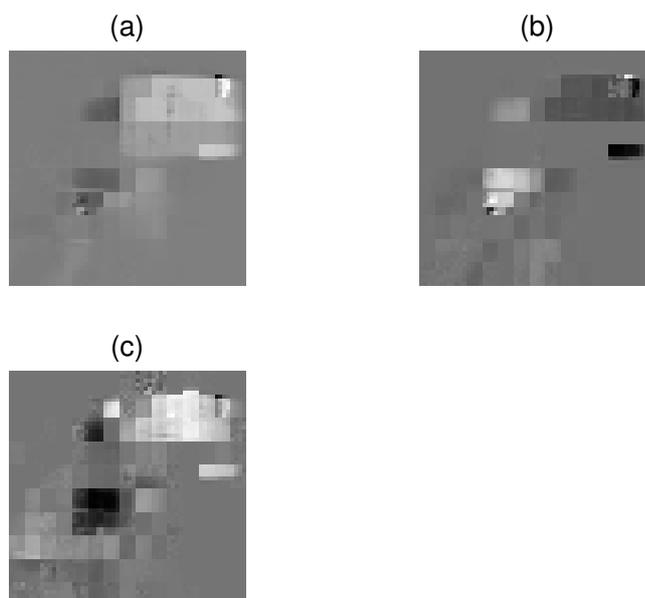


Figura 33: Expansão das observações referentes ao valor máximo do C-NLPC de cada padrão, em tons de cinza. (a) Primeiro padrão. (b) Segundo padrão. (c) Terceiro padrão. A escala utilizada é a mesma mostrada na Figura 29.

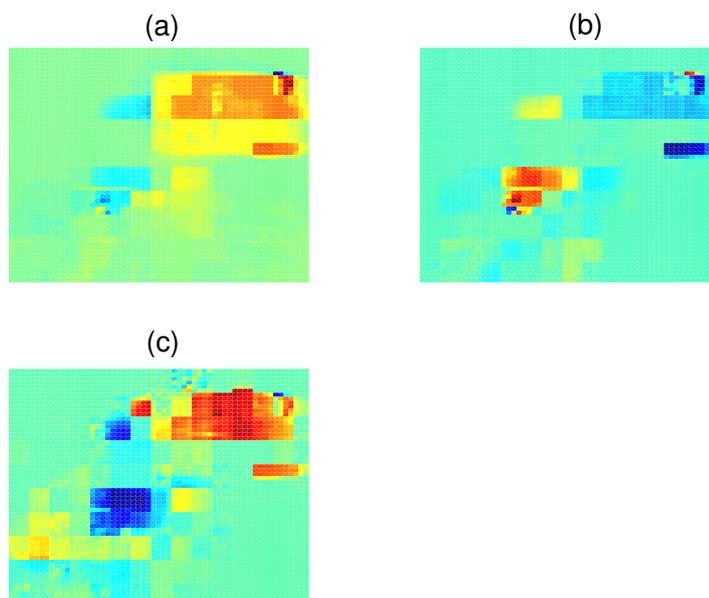


Figura 34: Expansão das observações referentes ao valor máximo do C-NLPC de cada padrão, em *pseudocolor*. (a) Primeiro padrão. (b) Segundo padrão. (c) Terceiro padrão. A escala utilizada é a mesma mostrada na Figura 30.

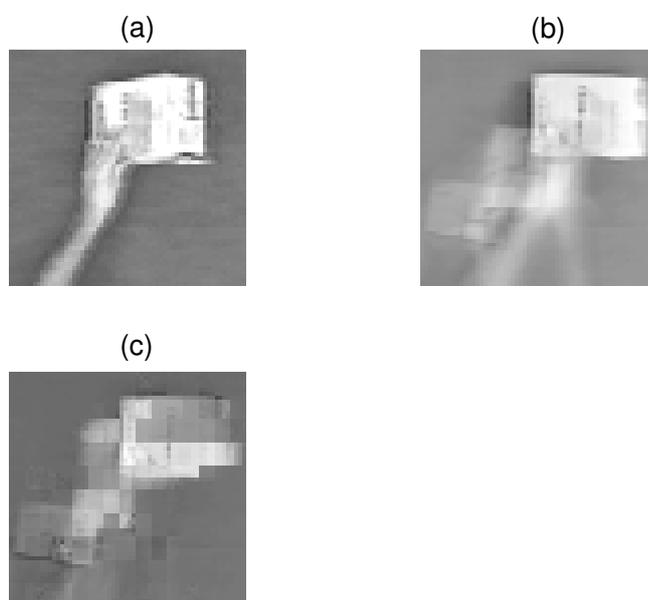


Figura 35: Reconstrução, a partir dos 3 primeiros padrões de movimento, de uma das 128 observações analisadas, em tons de cinza. (a) Imagem original. (b) Reconstrução com o método linear (PCA). (c) Reconstrução (expansão) com o método não linear (C-NLPCA). A escala utilizada é a mesma mostrada na Figura 29.

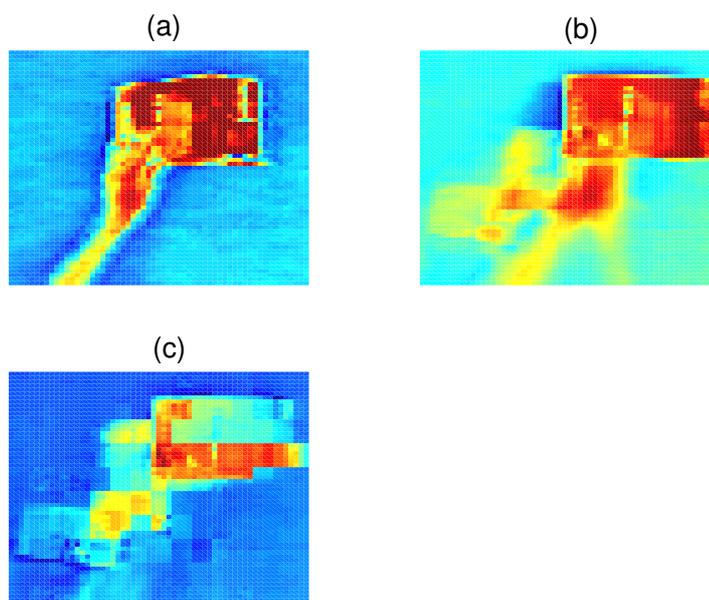


Figura 36: Reconstrução, a partir dos 3 primeiros padrões de movimento, de uma das 128 observações analisadas, em *pseudocolor*. (a) Imagem original. (b) Reconstrução com o método linear (PCA). (c) Reconstrução (expansão) com o método não linear (C-NLPCA). A escala utilizada é a mesma mostrada na Figura 30.

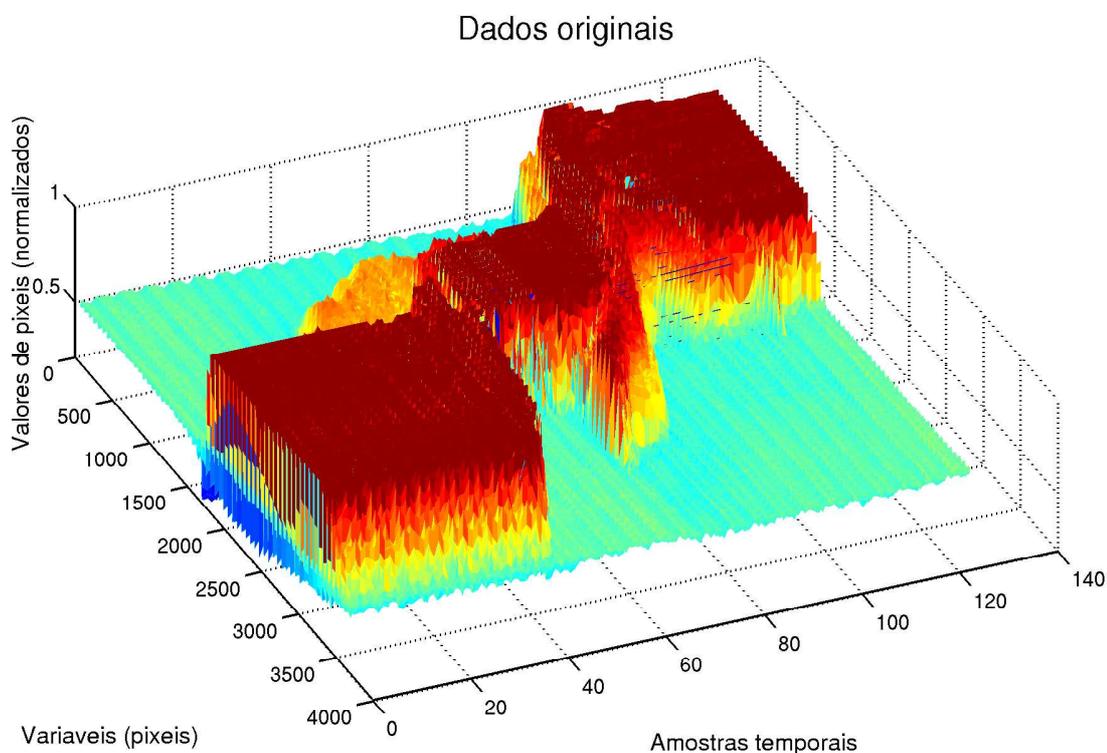


Figura 37: Evolução temporal do valor de cada variável (pixel) da série de imagens originais.

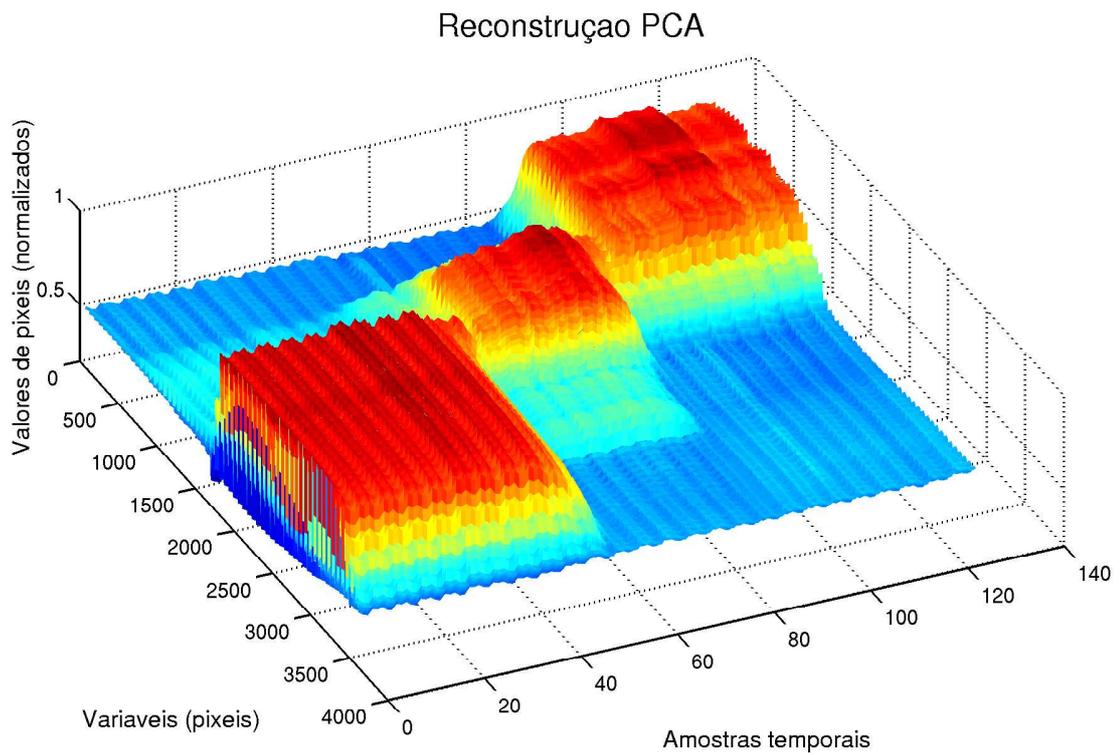


Figura 38: Evolução temporal do valor de cada variável (pixel) da série de imagens originais.

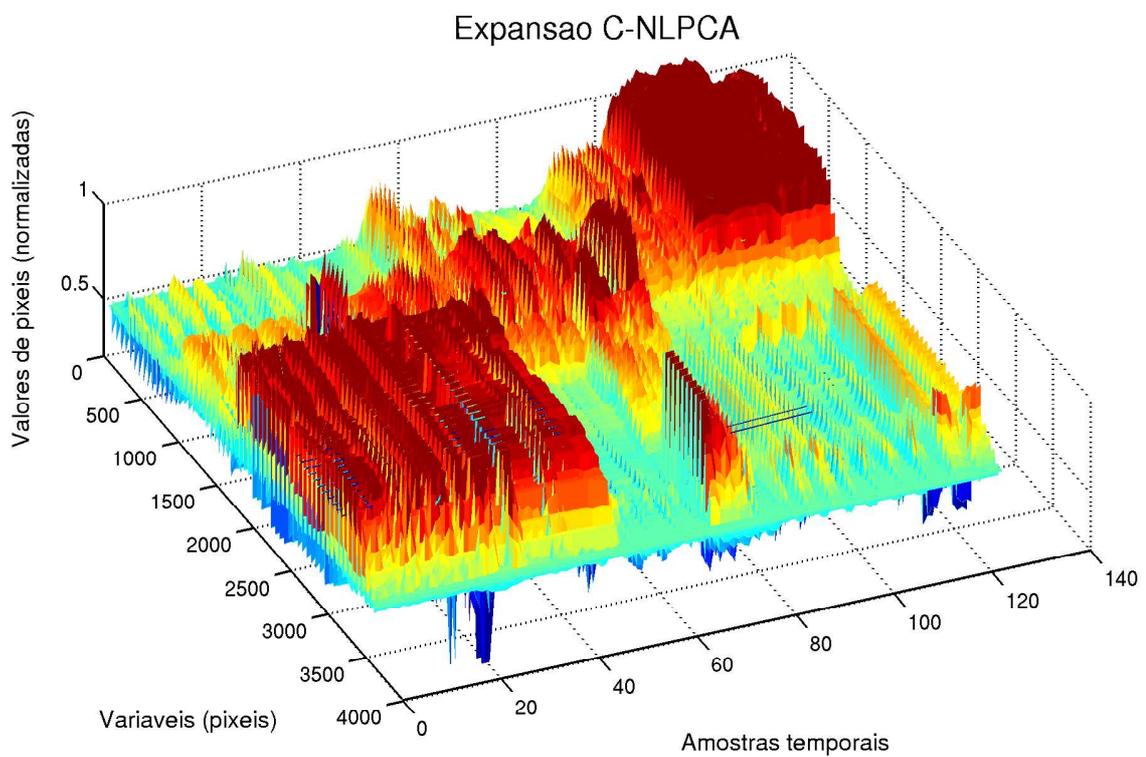


Figura 39: Evolução temporal do valor de cada variável (pixel) da série de imagens originais.

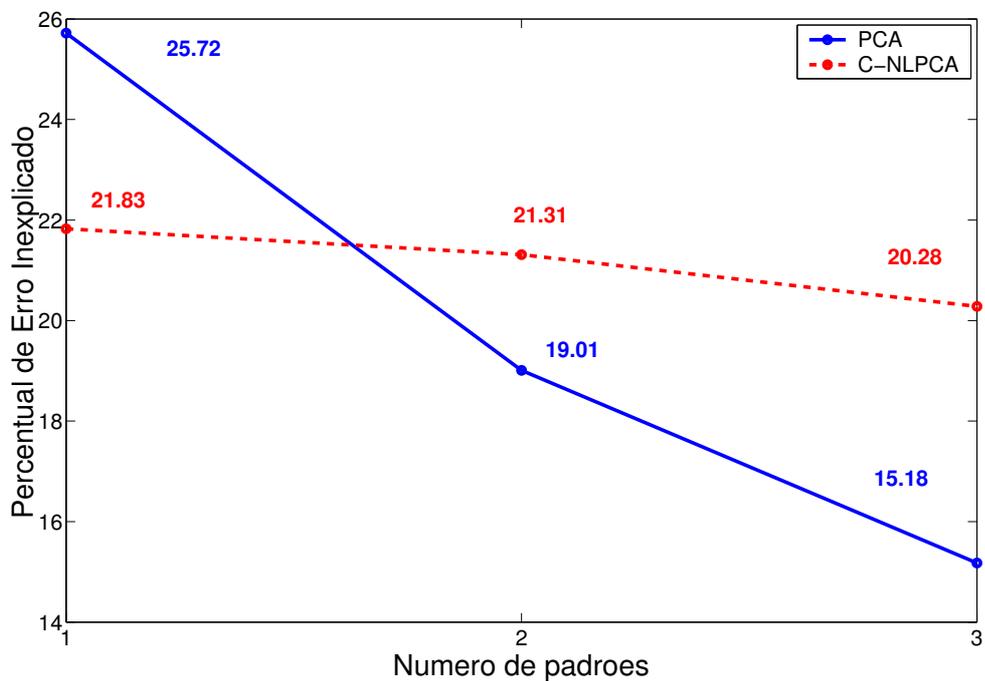


Figura 40: Percentual de erro inexplicado por cada método, de acordo com o número de padrões utilizados para reconstruir os dados.

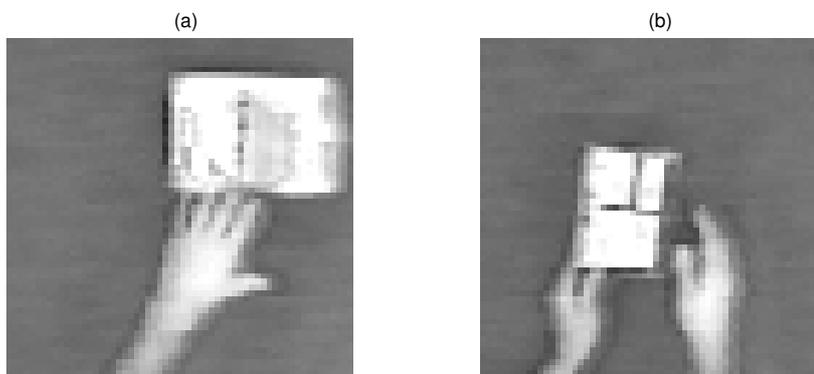


Figura 41: *Frames* mostrando o ápice do movimento das mãos esquerda (a) e direita (b), na série de imagens.

Conclusão

O presente trabalho apresentou a metodologia empregada para execução do C-NLPCA, propôs e avaliou o uso desse método para extração de componentes principais de movimento de séries de imagens no espaço 2D. Sendo o C-NLPCA e a extração de componentes principais de movimento os temas centrais no desenvolvimento do trabalho, nas seções seguintes serão apresentadas as contribuições, os programas implementados e os trabalhos futuros relacionados a esses dois assuntos.

Contribuições

Em relação ao C-NLPCA, as principais contribuições do trabalho foram o detalhamento da metodologia utilizada e o desenvolvimento de programas para execução do método (ver seção seguinte). Vale ressaltar que o C-NLPCA apresenta vantagens, sobre outros métodos de análise multivariada de dados, mencionados no Capítulo 1, como a capacidade de lidar com não linearidades existentes nos dados e a possibilidade de analisar um número ilimitado de variáveis. Algumas desvantagens também foram constatadas, como a instabilidade na análise intrínseca ao método NLPCA, como foi citado por Hsieh (HSIEH, 2001b); a considerável “sensibilidade” ao ajuste dos parâmetros da análise, com visto na discussão dos resultados dos experimentos; e a complexidade na implementação e execução da análise, em comparação aos métodos PCA e NLPCA.

A respeito da extração de componentes principais de movimento, pode-se perceber, através dos resultados dos experimentos realizados, que o C-NLPCA foi capaz de extrair esses componentes das séries de imagens analisadas, através do Estágio de Redução, e posteriormente reconstruir a série de dados originais, no Estágio de Expansão. Além disso o método conseguiu proporcionar um ganho, em relação ao PCA, na variação explicada pelo primeiro componente principal encontrado. No entanto, devido a problemas no ajuste dos parâmetros da análise, os componentes de mais alta ordem apresentaram ruídos, e captaram peculiaridades referentes a algumas amostras, que não deveriam estar presentes nos mesmos.

Outro ponto importante, é a possibilidade de uso da extração de componentes principais de movimento para solução de vários problemas encontrados na literatura, como por exemplo, reconhecimento de gestos, transferência de habilidades e animação gráfica. As abordagens, para o tratamento desses e de outros problemas relacionados, adotadas nos trabalhos estudados, concentram-se na modelagem de movimentos e objetos, na análise de poucos dados obtidos por meio de sensores, ou ainda, na pré-filtragem de grandes conjunto de dados por meio de métodos como o PCA. A extração de componentes principais de movimento de imagens 2D, através do C-NLPCA, apresenta-se como uma abordagem alternativa para o tratamento de grande conjuntos de variáveis, principalmente quando as mesmas possuem relações não lineares entre si.

Programas implementados

Foi desenvolvida em C++, e aprimorada durante a realização do trabalho, a biblioteca de funções RNALib, contendo as classes e respectivos métodos necessários para implementação, treinamento e aplicação de RNAs, e para implementação dos métodos NLPCA e C-NLPCA. Foram também desenvolvidos, utilizando a biblioteca RNALib, um programa para execução do NLPCA e outro para execução do C-NLPCA. Foram elaborados também uma série de *scripts* para manipulação dos dados no Matlab©. Com essas ferramentas, novos trabalhos e experimentos podem ser realizados com facilidade e rapidez, sem demanda de tempo de desenvolvimento de *software*, que foi significativa durante o projeto.

Trabalhos Futuros

Observando os resultados dos experimentos, foi possível determinar que o C-NLPCA extraiu os componentes principais de movimento das imagens analisadas, apesar de existência de alguns problemas já mencionados. Dessa forma, faz-se necessário a realização de mais testes, com o objetivo de calibrar os parâmetros do método, de modo a proporcionar mais ganhos na análise.

Outra questão importante, é a determinação criteriosa dos ganhos obtidos com o uso do C-NLPCA, em relação ao emprego do NLPCA com pré-filtragem através do método PCA, na análise de grandes conjuntos de variáveis com relações essencialmente não lineares entre si. Para isso é indicada a elaboração de um conjunto de dados sintético para ser

submetido a análise por cada um dos métodos.

Por fim, durante a realização do trabalho foram identificados várias aplicações onde a extração de componentes principais através dos métodos abordados poderia ser empregada. Essas aplicações vão desde a determinação de mapas topológico de redes de computadores, até a navegação de veículos autônomos baseada em imagens. Assim, considerando os *softwares* desenvolvidos neste trabalho, uma vasta gama de experimentos podem ser realizados com facilidade futuramente.

Referências

- BEM, R. A. de; BOTELHO, S. S.; MATA, M. M. Aplicação de redes neurais artificiais na análise de padrões de variabilidade de mesoescala. In: *Anais XI Simpósio Brasileiro de Sensoriamento Remoto - SBSR*. Brasil: [s.n.], 2003. p. 2007–2014.
- BJÖRNSSON, H.; VENEGAS, S. A. A manual for eof and svd analyses of climatic data. February 1997.
- BLACK, M. J.; JEPSON, A. D. Eigentracking: Robust matching and tracking of articulated objects using a view-based representation. *International Journal of Computer Vision*, Kluwer Academic Publishers, v. 26, n. 1, p. 63–84, 1998.
- BORTOLETTI, A. et al. A new class of quasi-newtonian methods for optimal learning in MLP-networks. *IEEE Transactions on Neural Networks*, v. 14, n. 2, mar. 2003.
- BOTELHO, S. et al. Applying neural networks to study the mesoscale variability of oceanic boundary currents. *Lectures Notes in Artificial Intelligence*, v. 2871, p. 684–688, 2003.
- BYRD, R. H. et al. *A Limited Memory Algorithm for Bound Constrained Optimization*. [S.l.], maio 1994. Technical Report NAM-08.
- CAMPOS, T. E. de. *3D Hand and Object Tracking for Intention Recognition*. [S.l.], 2003.
- CHALMOND, B.; GIRARD, S. C. Nonlinear modeling of scattered multivariate data and its application to shape change. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 5, n. 21, May 1999.
- COHEN, C. *Dynamical system representation, generation, and recognition of basic oscillatory motion gestures and applications for the control of actuated mechanisms*. Tese (Doutorado) — University of Mich, 1996.
- COHEN, C. J. et al. Dynamic system representation of basic and non-linear in parameters oscillatory motion gestures. In: *IEEE International Conference on Systems, Man, and Cybernetics*. Orlando, USA: [s.n.], 1997. v. 5, p. 4513–4518.
- DEMERS, D.; COTTRELL, G. Nonlinear dimensionality reduction. *Neural Inform. Processing Syst.*, v. 5, p. 580–587, 1993.
- DIAMANTARAS, K. I.; KUNG, S. Y. *Principal Component Neural Networks: Theory and Applications*. [S.l.]: Wiley, 1996.
- DONG, D.; MCAVOY, T. J. Nonlinear principal component analysis-based on principal curves and neural networks. *Computers & Chemical Engineering*, v. 20, n. 1, p. 65–78, July 1995.

- DORNER, B. *Chasing the Colour Glove: visual hand tracking*. Dissertação (Mestrado) — Computing Science - Simon Frase University, June 1994.
- FREEMAN, J.; SKAPURA, D. *Neural Networks - Algorithms, Applications, and Programming Techniques*. [S.l.]: Addison-Wesley, 1992.
- HAIR, J. J. F. et al. *Multivariate data analysis: with readings*. 4th. ed. [S.l.]: Prentice Hall, 1995.
- HASTIE, T.; STUETZLE, W. Principal curves. *Journal of the American Statistical Society*, v. 84, n. 406, p. 502–16, June 1982.
- HSIEH, W. W. *Manual for Nonlinear Principal Component Analysis and Nonlinear Canonical Correlation Analysis*. Vancouver, Canada, 2001. Version 2.2.
- HSIEH, W. W. Nonlinear principal component analysis by neural networks. *Tellus Press*, v. 53A, p. 599–615, 2001.
- KIRBY, M.; SIROVICH, L. Application of karhunen-loeve procedure for the characterization of human faces. *IEEE transaction on pattern analysis and machine intelligence*, 1990.
- KOHLER, M. R. J. System architecture and techniques for gesture recognition in unconstrained environments. In: *Proc. of the 1997 Int. Conf. on Virtual Systems and MultiMedia (VSMM'97)*. Geneva: [s.n.], 1997. p. 137–146.
- KRAMER, M. A. Nonlinear principal component analysis using autoassociative neural networks. *AlChe Journal*, v. 37, p. 233–43, February 1991.
- LEBLANC, M.; TIBSHIRANI, R. Adaptive principal surfaces. *Journal of the American Statistical Society*, v. 89, n. 425, p. 53–64, March 1994.
- LEE, C. *Learning Reduced-Dimension Models of Human Actions*. Tese (Doutorado) — The Robotics Institute - Carnegie Mellon University, May 2000.
- LEE, C.; XU, Y. Reduced-dimension representation of human performance data for human-to-robot skill transfer. In: *Proceedings of the 1998 IEEE/RSJ Intl. Conference on Intelligent Robots and Systems*. Canada: [s.n.], 1998.
- LIN, J.; WU, Y.; HUANG, T. S. Modeling the constraints of human hand motion. In: *Proc of IEEE Human Motion Workshop*. [S.l.: s.n.], 2000. p. 121–126.
- MACLEAN, W. J. et al. Fast hand gesture recognition for real-time teleconferencing applications. In: *II Int. Workshop on Recognition, Analysis and Tracking of Faces and Gestures in Real-time Systems, IEEE Computer Society*. Vancouver, Canada: [s.n.], 2001.
- MALTHOUSE, E.; MAH, R.; TAMHANE, A. Some theoretical results on nonlinear principal components analysis. In: *Proceedings of the American Control Conference*. Seattle, Washington: [s.n.], 1995.
- MALTHOUSE, E. C. Limitations of nonlinear PCA as performed with generic neural networks. *IEEE Transactions on Neural Networks*, v. 9, n. 1, jan. 1998.

- MONAHAN, A. *Nonlinear principal component analysis of climate data*. Tese (Doutorado) — University of British Columbia, 2000.
- OJA, E. A simplified neuron model as a principal component analyzer. *J. Math. Biology*, v. 15, p. 267–273, 1982.
- PRESS, W. H. et al. *Numerical Recipes in C - The Art of Scientific Computing*. [S.l.]: Cambridge University Press, 1990.
- ROMDHANI, S. *Face Recognition using Principal Components Analysis*. [S.l.].
- ROMDHANI, S.; PSARROU, A.; GONG, S. Multi-view nonlinear active shape model using kernel pca. In: *Tenth British Machine Vision Conference*. [S.l.: s.n.], 1999.
- SCHOLZ, M.; VIGÁRIO, R. Nonlinear pca: a new hierarchical approach. In: *ESANN'2002 proceedings - European Symposium on Artificial Neural Networks*. Bruges(Belgium): [s.n.], 2002. p. 439–444.
- SIM, T. et al. Memory-based face recognition for visitor identification. In: *Proceedings of 4th Intl. Conf. on Face and Gesture Recognition*. Grenoble, France: [s.n.], 2000. p. 214–220.
- STROBEL, M. et al. Using spatial context knowledge in gesture recognition for commanding a domestic service robot. In: *Proceedings of the 2002 IEEE Int. Workshop on Robot and Human Interactive Communication*. Berlim, Germany: [s.n.], 2002.
- TAN, S.; MAVROVOUNIOTIS, M. L. Reducing data dimensionality through optimizing neural network inputs. *AIChe Journal*, v. 41, n. 6, p. 1471–1480, June 1995.
- TATANI, K.; NAKAMURA, Y. Dimensionality reduction and reproduction with hierarchical nlpc neural networks - extracting common space of multiple humanoid motion patterns. In: *Proceedings of the 2003 IEEE International Conference on Robotics and Automation*. Taipei: [s.n.], 2003. p. 1927–32.