Robust hierarchical image representation using non-negative matrix factorisation with sparse code shrinkage preprocessing

Botond Szatmáry, Gábor Szirtes, András L?rincz, Julian Eggert, Edgar Körner

2003

Preprint:

Pattern Anal Applic (2003) 6: 0–0
DOI 10.1007/s●●●

ORIGINAL ARTICLE

B. Szatmáry, G. Szirtes, A. Lörincz, J. Eggert and E. Körner

Robust hierarchical image representation using non-negative matrix factorization with sparse code shrinkage preprocessing

Received: ●●● / Accepted: ●●●

Abstract When analysing patterns, our goals are (i) to find structure in the presence of noise, (ii) to decompose the observed structure into sub-components, and (iii) to use the components for pattern completion. Here, a novel loop architecture is introduced to perform these tasks in an unsupervised manner. The architecture combines sparse code shrinkage with non-negative matrix factorisation, and blends their favourable properties: sparse code shrinkage aims to remove Gaussian noise in a robust fashion; non-negative matrix factorisation extracts substructures from the noise filtered inputs. The loop architecture performs robust pattern completion when organised into a two-layered hierarchy. We demonstrate the power of the proposed architecture on the so-called 'bar-problem' and on the Feret facial database.

Keywords Hierarchy · Non-negative matrix factorisation · Sparse code shrinkage

Introduction

In this paper we propose a novel architecture, which is mainly based on two algorithms. The aim of our architecture is to distinguish pattern from noise. The first algorithm is Independent Component Analysis (ICA), also called blind source separation, or demixing, which has almost a ten year history [1–5]. ICA solves the following model: there are n unknown independent sources, the components of the source vector $\mathbf{s} \in \mathbb{R}^n$, and an unknown mixing matrix, $\mathbf{A} \in \mathbb{R}^{n \times m}$ ($m = n$ is possible. ICA tries to estimate matrix $\mathbf{W} \in \mathbb{R}^{m \times n}$ the (pseudo)inverse of $\mathbf{A}$, and to recover the original sources $\mathbf{s}$ by means of a set of mixed signals $\mathbf{x}_i$, $i = 1, 2, \ldots N$, ($\mathbf{x}_i = \mathbf{A}\mathbf{s}_i$).[1] The estimation is based on entropy minimisation (e.g. see Hyvärinen [6] and references therein). ICA transformation efficiently removes higher order correlations from input components.

A recent extension of this ICA technique is Sparse Code Shrinkage (SCS) [7]. SCS is concerned with the removal of noise. In most pattern recognition problems, noise filtering is a central issue. The separation of noise from data is, however, problem-dependent. In information theory, noise is considered structure-free (i.e. of maximal entropy).[2] The SCS algorithm aims to separate Gaussian noise from structured components locally, i.e. component by component. In turn, SCS makes use of ICA transformation and works on components with minimised mutual information. SCS applies a shrinkage nonlinearity (function g) to each ICA component, which eliminates small amplitude ICA components. This novel approach can be considered as a generalisation of wavelet denoising [7]. The generalisation concerns the process of learning the underlying basis set given an ensemble of inputs, and then performing *denoising* via thresholding, as in wavelet bases.

Decomposition of multivariate data into 'positive components' (positively coded substructures) can be efficient in pattern completion problems. The objective of learning is to seek subparts in individual inputs of an ensemble so as to enable inferencing. A powerful recent technique is non-negative matrix factorifation (NMF) [8,9]. In NMF, each input is built from non-negative components. NMF minimises the $\Sigma_i (\mathbf{x}_i - \mathbf{Q}\mathbf{h}_i)^2$ cost function, where rows of matrix $\mathbf{Q} \in \mathbb{R}^{n \times m}$ contain the NMF basis set, $\mathbf{h} \in \mathbb{R}^m$ is the hidden representation vector, and $\mathbf{x}_i$ $i = 1, 2, \ldots n$ denote the set of input vectors. It is assumed that, for all i, j $\mathbf{x}_i$, $\mathbf{Q}_{ij}$, $\mathbf{h}_i \geq 0$. In Lee and Seung

B. Szatmáry ✉, G. Szirtes, A. Lörincz
Eötvös University Department of Information Systems,
Budapest, Hungary

J. Eggert and E. Körner
HONDA R&D Europe,
Future Technology Research, Offent

[1] The row vectors of matrix $\mathbf{W}$ are called the ICA basis set.
[2] For continuous variables, Gaussian distribution of unit variance has the maximal entropy.

[9], the following batch mode learning rule for the hidden variables and the basis set is suggested:

$$\mathbf{h}_i \leftarrow \mathbf{h}_i \frac{(\mathbf{Q}^T\mathbf{x})_i}{(\mathbf{Q}^T\mathbf{Q}\mathbf{h})_i},\ \mathbf{Q}_{ij} \leftarrow \mathbf{Q}_{ij} \frac{(\mathbf{x}\mathbf{h}^T)_{ij}}{(\mathbf{Q}\mathbf{h}\mathbf{h}^T)_{ij}}$$

The interested reader can find details of the algorithms elsewhere [6–9].

In this paper, a novel combination of the SCS and NMF algorithms is proposed. In Sect. 2 the joined (tandem) architecture is described. Simulation results on the two-bar problem are presented in Sect. 3. The simulations demonstrate the properties of the combined architecture, as with the properties of a hierarchy of such architectures. Conclusions are drawn in Sect. 4.

Architecture

SCS is a bottom-up transformation that aims to recover the original sources given only the mixed sources. In the case of independent sources covered by additive Gaussian noise, SCS can recover the original inputs. NMF, on the other hand, can be seen as a top-down generative algorithm that optimises its internal representation to minimise the reconstruction error between input and generated (reconstructed) input. These two unsupervised algorithms can be merged into a single parameter-free architecture, referred to as a 'loop'. This loop is depicted in Fig. 1(a). The learning capabilities and the working performance of this loop have been investigated by us [10] and by Hoyer [11], independently. One such loop exhibits good performance under the following conditions:

- *Bottom-up learning*: the bottom-up demixing matrix, **W** and the sparsifying SCS shrinkage function g are learned in this phase.
- *Top-down learning*: inputs, which are filtered by the bottom-up matrix **W** and by the SCS nonlinearity, are transformed by the pseudo-inverse $\mathbf{W}^+$ of matrix **W** (which is equal to $\mathbf{W}^T$ for our case [7]): $\mathbf{x}_{SCS} = \mathbf{W}^T g(\mathbf{W}\mathbf{x})$. In other words, the inputs are projected into the SCS subspace defined by the row vectors of matrix **W**, sparsification is performed on the projections (function g), and then these sparse components are transformed back to the original input space. These nonlinearly filtered inputs are shifted to the positive range to satisfy the requirements of NMF. The bounded and non-negative inputs are used in batch mode to compute the NMF basis set (the NMF matrix, **Q**).
- *Working phase*: inputs are denoised by projecting them into the SCS subspace, as before. Hidden components **h** are computed via the NMF iteration procedure (denoted by $\mathbf{I}_{NMF}$ in Fig. 1) keeping the **Q** matrix fixed. Reconstructed input $\hat{\mathbf{x}}$ is computed by multiplying the hidden vector **h** by the NMF matrix **Q**: $\hat{\mathbf{x}} = \mathbf{Q}\mathbf{h}$.

For further details, see the Appendix.

A two-layered hierarchy was built from a set of loops. The two layer architecture is shown in Fig. 1(b). The lower layer consists of parallel working loops, each corresponding to a subpart of the original input, i.e. each loop processes only a part of the input. The **h** hidden representations of the lower layer loops were concatenated into one block. This concatenated vector formed the input vector of the loop of the higher layer. To guide the eye, two-dimensional patterns were used in our simulations. For the same reason, loops of the lower layer corresponded to neighbouring regions of the 2D input space.

In the working phase, first, approximation of the hidden values are computed by the lower layer. The concatenation of these hidden values make up the input of the higher layer. In this higher layer, both learning and working occurs on the concatenated hidden representation of the lower layer. The reconstructed input of the second layer then over-runs the hidden values of the lower layers. Top-down reconstruction in the lower layers, in turn, corrects the reconstructed inputs $\hat{\mathbf{x}}$. Extension to more layers is straightforward.[3]

[3] Note that the only infection about the input scape, which is prewired into this hierarchy is restricted to the topographical arrangements of the loops of the lower layer. However, the 2D structure

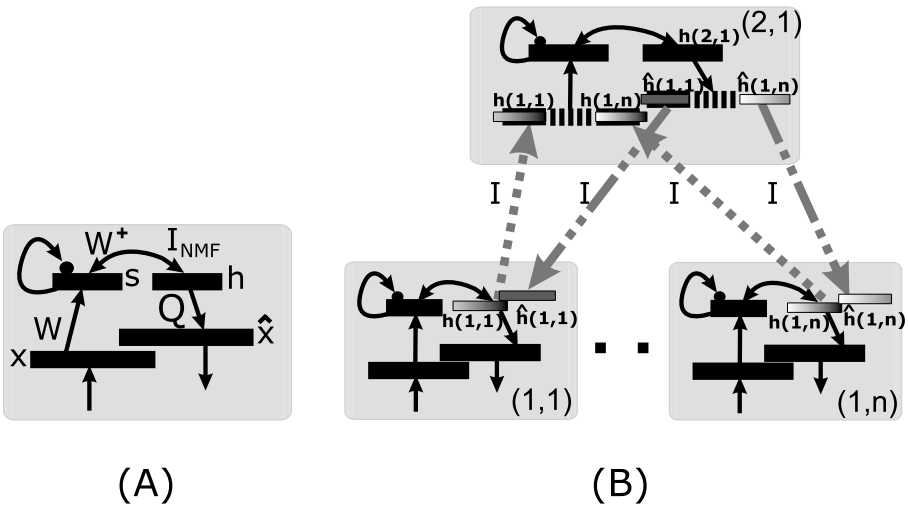

Fig. 1 (a) Graphical representation of the algorithm. One loop of the hierarchy **x, s, h** and $\hat{\mathbf{x}}$: input, shrunk (denoised) ICA components, hidden variable, and reconstructed input, respectively. **W**, $\mathbf{W}^+$, **Q** and $\mathbf{I}_{NMF}$ denote demixing matrix, pseudo-inverse of the demixing matrix, NMF matrix and NMF iteration, respectively. Arrow: linear transformation, arrow with black dot: linear transformation with component-wise nonlinearity (shrinkage kernel), lines with two arrow-heads: iteration. The algorithm was utilised in a two phase mode (see text for details). (b) Two-layer hierarchy. $\hat{\mathbf{h}}$: reconstructed input of second layer. Lower layers provide inputs (their hidden variables **h**) for higher layers. Reconstructed input of second layer overruns hidden variables at first layers. Dashed and dash-dotted grey arrows represent identity transformations (I)

Results

The bar example

The task is to learn and reconstruct bars on greyscale images while (i) each input is composed of two or more bars at vertical and horizontal orientations, (ii) the input is covered by additive Gaussian noise of zero mean, and (iii) in the hierarchical arrangement the input bars have missing parts. We studied the effect of the additional noise on the quality of the reconstruction. The 'original' noise free bars were represented by 1s (white pixels) for each vector component, whereas the background corresponded to 0 valued (black pixels) vector components. Zero mean Gaussian noise was added to the values of the noise-free inputs. These noise-corrupted inputs were shifted and clipped to the positive range while preserving 90% of the noise. A typical example from the set of the 'original' noise-free inputs, which contained two or more bars, is shown in Fig. 2(a). With regard to Hyvärinen [7], the SCS basis set was developed on noise-free inputs (e.g. Fig. 2(a)), but the noise-corrupted version of inputs Fig. 2(b) was input into the architecture in the working phase and the NMF learning phase. (See the Appendix for details.)

Simulation results are shown in Figs. 2, 3 and in Fig. 5. Fig. 2 (B) depicts one of the inputs used for training. Similar inputs were used for testing. The SCS, NMF, and combined SCS and NMF reconstructed inputs are shown in Fig. 2(C)–(E), respectively. Visual inspection reveals that ICA alone (Fig. 2(C)) is inefficient although tracks of the bars can be discovered. NMF (Fig. 2(D)) fails with such high noise content, the reconstructed input is very noisy no bar like structure can be seen. The combined method 'predicts' higher amplitudes for pixel values corresponding to the 1's and lower amplitudes for pixel values corresponding to the 0's of the original (noise-free) input, leading to a better reconstruction.

We can gain a better insight by examining the NMF basis vectors as well as the basis vectors of the combined method (Fig. 3). It was found that the SCS algorithm indeed discovers the bar structure (not shown in the figure), but fails to separate the individual bars. NMF is not able to cope with the high noise content: the learned basis set contains short tracks of bars covered by noise (Fig. 3(a)). The combined method shows superior performance (Fig. 3(b)): 11 basis vectors represent well distinguishable individual bars. Combinations of a few bars are represented by 12 basis vectors. These 23 basis vectors contain relatively small background noise. One basis vector contains mostly noise components.

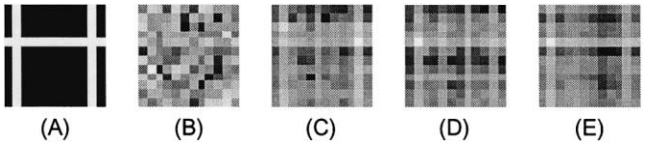

Fig. 2 Input and its reconstructed forms. (a) Perfect input without noise. Input size: 12 × 12. (b) Noise-free input covered with zero mean 1.5 STD Gaussian noise. (c) Reconstructed input (RI) with SCS. (d) RI with NMF. (e) RI with combination of SCS and NMF. Note the improved reconstruction for the combined method compared to single SCS or single NMF algorithms. The number of basis vectors for both algorithms (SCS, NMF) was set to 24

Quantitative noise dependencies are shown in Fig. 4. Root Mean Square (RMS) reconstruction error versus noise is depicted for the NMF and for the combined method. Figure 4(a) shows the noise dependence of the reconstruction error for a perfect single bar input covered with varying amount of noise. NMF (SCS+NMF) results are shown by dashed (dotted) lines. Another plot shows the analogous results for two bar inputs (Fig. 4(b)). Results are superior by a factor of two for both cases in favour of the combined method. The combined method improves noise resistance of NMF and, in turn, the combined method exhibits improved decomposition capability. Note that the RMS error is small but non-zero for zero noise inputs, because the filters developed on inputs with high noise content cannot be perfect.

The pattern completion property is demonstrated in a two-layered hierarchy, with the lower layer consisting of several loops (Fig. 1(a)), each working on localised subareas of the input. The input of the higher layer is concatenated from the hidden layer activities of the subnets of the lower layer over the entire input range, as shown in Fig. 1(b). Inputs, as in the previous experiment, were from two or more bars during both the learning test phases. Two different types of inputs were presented. The upper row of Fig. 5 shows an example when the original (noise-free) input has missing (zeroed) pixel components. The lower row (Fig. 5) depicts the case when some loops of the lower layer had no bar-like components and could transfer only noise. (This case is intended to simulate occlusions). The original noise-free and incomplete inputs are shown in column (a). Noise covered inputs are depicted in column (b). Reconstruction capabilities of the lower layer and of the lower and the higher layers together are shown in columns (c) and (d), respectively. Visual inspection reveals that input reconstruction in the hierarchy is considerably better.

The face example

The task here is to reconstruct faces. In these experiments the Feret [12] database was used. Frontal views of faces (2061 images) were used. Images were distorted and had identical eye-to-eye and eye-to-mouth distances. The edges of the images were removed and smaller images were used to lessen background related effects. Images were further transformed to have the same (average) intensity and the same variation around this average inten-

is represented only in the input set, and no explicit information about the 2D structure is prewired. The inputs, the NMF and the ICA basis sets are represented in 2D for visualisation purposes.

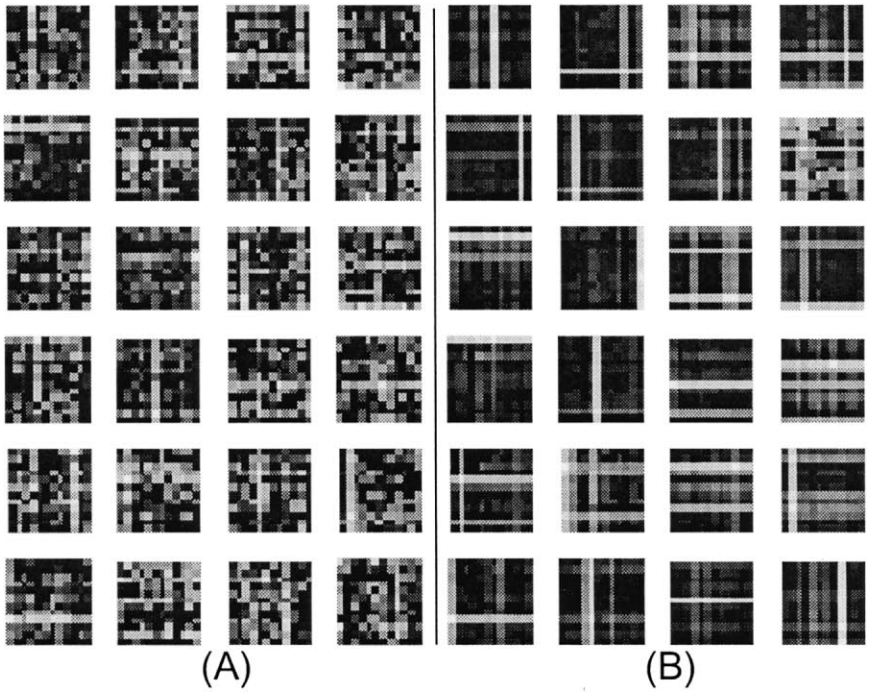

Fig. 3 Single layer basis vector sets for (a) NMF alone, and for the (b) NMF using SCS outputs. Inputs have 144 (= 12 × 12) dimensions. Number of filters: 24. For each method, all 24 filters are depicted. The additional standard deviation of the used Gaussian noise was 2.1

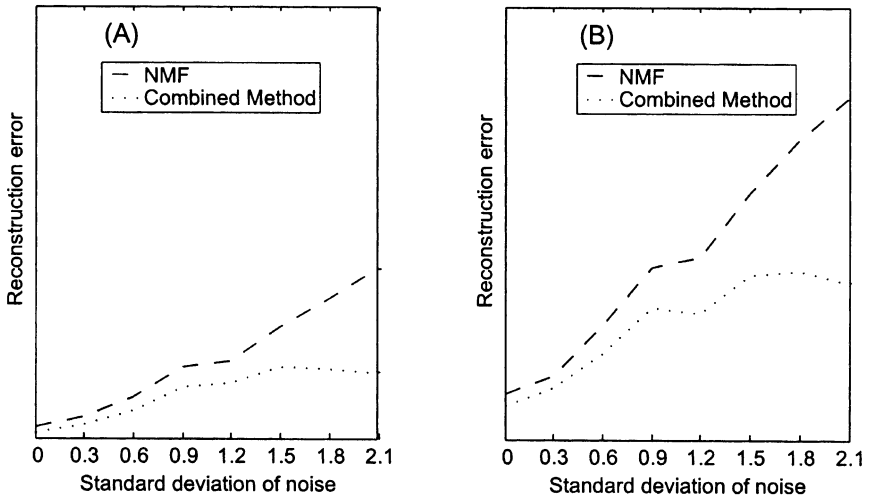

Fig. 4 Reconstruction errors. Markers denote different methods. NMF: dashed line, combined SCS and NMF: dotted line. (a) RMS reconstruction error for single bar inputs versus standard deviation of the noise. (b) Dependence of the RMS reconstruction error on standard deviation of the noise for double-bar inputs. In both cases, performance of the combined method is significantly better. Standard deviation

sity. Furthermore, these preprocessed images were downsized to 60 × 60 pixels.

In these experiments no noise was added and no ICA was used. The first layer contained four non-overlapping 60 × 60 pixel regions. The number of learned NMF basis vectors was 50 in each region. Taken together, the number of NMF basis vectors of the first layer was 200. The second layer was made of a single NMF network which contained and tuned 60 basis vectors.

Contrary to the learning phase, occluded inputs were used during the working phase. Covering the original inputs with a black triangle of varying size yielded the occluded inputs. Results are shown in Figs. 6) and 7. Figure 6) depicts one of the original faces (upper left) the occluded version of it (upper right), the reconstructed image produced by the first layer (lower left) and the reconstructed image produced by both layers (lower right). Figure 7 shows results on the reconstruction error as a function of the (linear) size of the occluded upper right triangle. Value 0.5 (0.25) corresponds to the case when half (one-eights) of the area is covered. The first layer provides better reconstruction for non-occluded inputs. This is because the supervisory information coming from the second layer is based on a smaller number of learned vectors (there are 60 vectors) and this compromises the reconstruction capabilities in the 60 × 60 (i.e. 3600) dimensions of the input. On the other hand, for large occlusions the advantage of the tandem architecture reveals, not only according to our subjective judgements (Fig. 6) that information in occluded areas are properly approximated, but also qualitatively, in the reconstruction error.

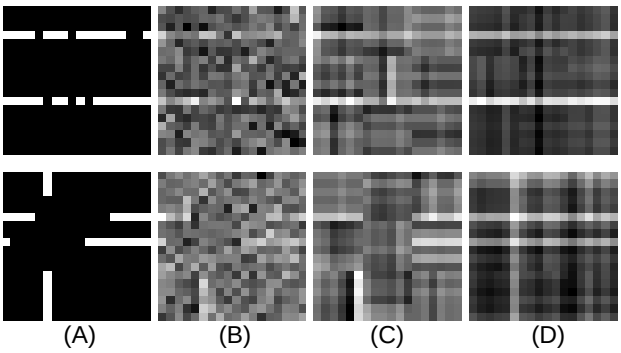

Fig. 5 Improved noise filtering and pattern completion in the hierarchy. Sub-architectures of the first layer now have $6 \times 6 = 36$ input dimension. Dimension of NMF hidden vectors of first layer units is 12 ($= 6 \cdot 2$). First layer is made of $3 \times 3 = 9$ units. Input of the second layer has 108 ($= 9 \cdot 12$) dimensions. Dimension of the hidden vector of the second layer is 36 ($= 3 \cdot 6 \cdot 2$). Upper row: Pixels of the inputs are missing. Bottom row: Pixels *and* sub-components of the inputs are missing. (a) Original input with missing pixels and missing sub-components, (b) noise corrupted input of the architecture, (c) reconstructed input produced by the lower layer, (d) reconstructed input produced by both layers. STD of applied noise is 0.9

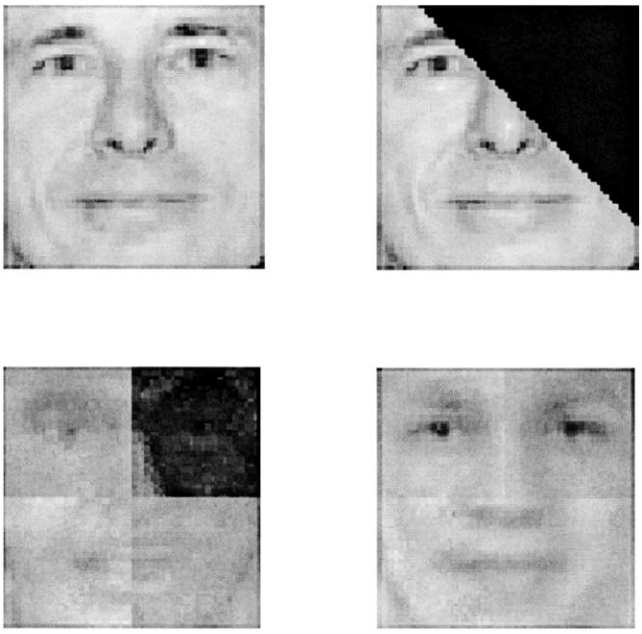

Fig. 6 Reconstruction of a face. Upper left: original face, upper right: occluded face, lower left: reconstructed image produced by the first layer, lower right: reconstructed image produced by both layers

This leaves us with the intriguing point that over-running the reconstruction process in lower levels is necessary if the higher level judges so (this would be the case for the occluded areas), but over-running the reconstruction process in general is not wise. A kind of inferencing is needed here, which imposes constraints onto the lower level and favours the lower level if constraints are mild. To give an example: uniform basis vectors could be included in the NMF basis sets of the lower level. If one of the sub-areas is occluded (e.g. by a black card), then this basis vector will allow perfect reconstruction. In spite of the perfectness of this reconstruction, it should be overrun by the higher layer. On the other hand, in areas, where higher layer reconstructions are close to their minimum, lower layer reconstruction should take place. This point is intriguing, because ICA is capable of estimating probabilities locally. In turn, neurons at the ICA layer could 'decide' what to do. These neurons have two types of input: the first is the ICA transformed input; the other, the curved arrows with dots in Fig. 1, are strictly determined by the ICA transformed reconstructed input. This input type serves sparsification. In the case of no or minor corrections imposed by the higher layer, the two types of input barely conflict with each other. However, if the internal representation is strongly modified by the higher layer, the input related to the reconstructed input will strongly conflict with the ICA transformed real input. In turn, we conjecture that optimisation of the reconstruction error may occur within the 'context' provided by the higher layer and the amount of modification 'suggested' by the higher layer. Small modifications may occur because the higher layer does not have full information. In this case, reconstruction should be left unmodified at the lower layer. This point deserves further study.

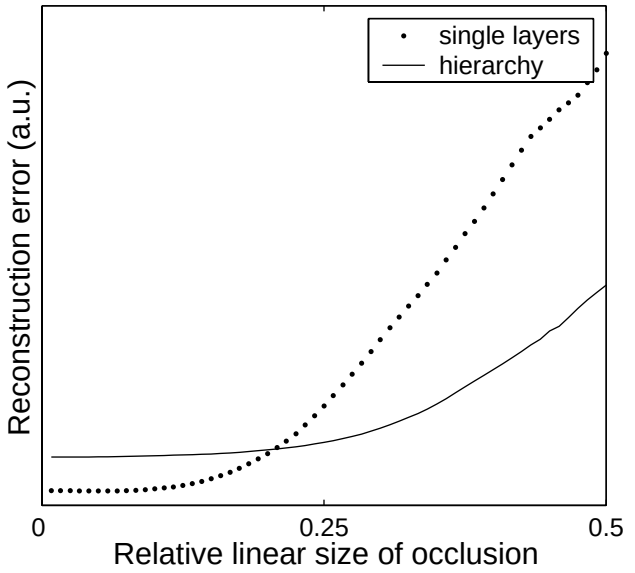

Fig. 7 Reconstruction error as a function of the linear size of occlusion. Dotted curve: single layer reconstruction error, solid curve: reconstruction error produced by the hierarchy. Reconstruction error is given in arbitrary units. Linear size 0.5 corresponds to a black area covering the upper right half of the image (For details, see text.)

Conclusions

A novel nonlinear reconstruction algorithm has been introduced in this paper. The algorithm combines the advantages of sparse code shrinkage and non-negative matrix factorisation by removing Gaussian noise (via the SCS algorithm), and by decomposing the inputs into sub-components (via the NMF algorithm). Favourable proper-

ties of the joined tandem algorithm have been demonstrated via computer simulation. Correlation-based NMF is capable of pattern completion, a feature which is strengthened when the joined algorithms are used in a hierarchical arrangement. Making use of the pseudo-inverse of the SCS transformation, mathematical theorems on convergence and the stability of both the SCS and the NMF algorithms are left intact in the combined architecture by construction.

Acknowledgements

Helpful discussions with Ata Kaban are gratefully acknowledged. We would like to thank the Editor-in-Chief, Professor Sameer Singh, for his kind help on suggesting a demonstrative example which suits the purposes of this journal.

This work was partially supported by Hungarian National Science Foundation (Grant No. OTKA 32487) and by Honda R&D Europe GmbH, Future Technology Research, Offenbach am Main, Germany. Portions of the research in this paper use the FERET database of facial images collected under the FERET program.

APPENDIX

The combined algorithm in one loop can be summarised as follows.

(1 a) *SCS Learning phase*:
Estimation of the sparse coding transformation $\mathbf{W}$ and shrinkage function g using noiseless inputs.
(1 b) *SCS Working phase*:
Loading the already calculated sparse coding transformation $\mathbf{W}$ and shrinkage function g.
(2) Computation of the projection on the sparsifying basis: $s = g(\mathbf{Wx})$, where $\mathbf{x}$ is the input.
(3) Estimation of the noise-free data: $\mathbf{X}_{SCS} = \mathbf{W}^T\mathbf{s}$.
(4 a) *NMF Learning phase*:
Batch estimation of NMF basis set $\mathbf{Q}$ and the hidden variables $\mathbf{H}$ using the $\mathbf{X}_{SCS}$ 'noise-free' inputs:
REPEAT UNTIL CONVERGENCE:

$$\mathbf{H}_{ij} \leftarrow \mathbf{H}_{ij} \frac{(\mathbf{Q}^T\mathbf{X}_{SCS})_{ij}}{(\mathbf{Q}^T\mathbf{QH})_{ij}}$$

$$\mathbf{Q}_{ij} \leftarrow \mathbf{Q}_{ij} \frac{(\mathbf{X}_{SCS}\mathbf{H}^T)_{ij}}{(\mathbf{QHH}^T)_{ij}}$$

Columns of matrix $\mathbf{H}$ represent the hidden representation vectors, columns of matrix $\mathbf{X}_{SCS}$ represent the $\mathbf{X}_{SCS}$ de-noised input vectors.
(4 b) *NMF Working phase*:
Using the already calculated NMF basis set $\mathbf{Q}$, the hidden variables ($\mathbf{H}$) are estimated.
(5) Estimation of the reconstructed input $\hat{\mathbf{x}}$ by multiplying the hidden variable $\mathbf{h}$ with the NMF basis matrix $\mathbf{Q}$.

References

1 Jutten C and Herault J. Blind separation of sources. Part I: an adaptive algorithm based on neuromimetic architecture. Signal Processing 1991; 24: 1–10
2 Comon P. Independent component analysis – A new concept? Signal Processing 1994; 36: 287–314
3 Bell AJ and Sejnowski TJ. An information-maximization approach to blind separation and blind deconvolution. Neural Computation 1995; 7: 1129–1159
4 Amari S-I, Cichocki A and Yang HH. A new learning algorithm for blind signal separation. Advances in Neural Information Processing Systems, Morgan Kaufmann, San Mateo, CA 1996, pp 757–763
5 Amari S-I. Natural gradient works efficiently in learning. Neural Computation 1998; 10: 251–276
6 Hyvärinen A. Survey on independent component analysis. Neural Computing Surveys 1999; 2: 94–128
7 Hyvärinen A. Sparse code shrinkage: Denoising of nongaussian data by maximum likelihood estimation. Neural Computation 1999; 11: 1739–1768
8 Lee DD and Seung HS. Learning the parts of objects by non-negative matrix factorization. Nature 1999; 401: 788–791
9 Lee DD and Seung HS. Algorithms for non-negative matrix factorization. Advances in Neural Processing Systems, vol 13. Morgan Kaufmann, San Mateo, CA, 2001
10 Szatmáry B, Pßczos B, Eggert J, Lörincz A and Körner E. Non-negative matrix factorization extended by sparse code shrinkage. In Proceedings 15th european conference on artificial intelligence (ECAI2002), IOS Press, Amsterdam 2002, pp 503–507
11 Hoyer PO. Non-negative sparse coding. In Neural networks for signal processing XII, Proceedings IEEE workshop on neural networks for signal processing, Martigny, Switzerland, 2002 (http:\\www.cis.hut.fi/phoyer/papers/)
12 Phillips PJ, Wechsler H, Huang J and Rauss P. The FERET database and evaluation procedure for face recognition algorithms. Image & Vision Computing 1998; 16: 295–306

Botond Szatmáry is at Eötvös Loránd University, Budapest. He has been studyting for a PhD in informatics in András Lörincz's group. He holds an MSc degree in mathematics and computer science. He was awarded a Pro Scientia and first prize in 2001 and second prize in 1999 at the National Competition in Sciences for University students of informatics. His main interest is in ANN, independent component analysis and vision research.

András Lörincz is at Eötvös Loránd University, Budapest. His background is in experimental and theoretical physics, physical chemistry, and he graduated in 1997. He has been involved in artificial intelligence research for about ten years. He is presently leading the Neural Information Processing Group at the Department of Information Systems of Eötvös Loránd University. Research fields in the group include the neurobiological foundation of AI and AI applications. He has over 150 refereed international publications.

Gábor Szirtes is a PhD student of informatics at Eötvös Loránd University, Budapest. He received excellent rankings at National Competitions in chemistry and biology for high school students. He holds a MSc degree in chemistry and has also taken courses in physics. His main interest is in brain modelling, ANNs, information processing and applications.

Edgar Körner received his Dr-Ing in 1977 in the field of biomedical engineering, and his Dr of Science in 1984 in the field of biocybernetics, both from the Technical University Illmenau, Germany, where he became full professor and head of the department

of neurocomputing and cognitive science in 1988. From 1992–97 he was a chief scientist at Honda R&D Co., Wako, Japan. In 1997 he moved to Honda R&D Europe (Germany) to establish the 'Future Technology Research Division'. His research focus is on brain-like artificial neural systems for image recognition and understanding, smooth transition between signal-symbol processing, and self-organisation of knowledge representation.

Julian Eggert is at the Future Technology Research Division of Honda R&D in Offenbach, Germany. He received his PhD from the Technical University of Munich in physics, where he was working at the theoretical biophysics department of Professor J. L. van Hemmen. His interests comprise the dynamics of spiking neurons and neuronal assemblies, large-scale models for the vision system, and gating in hierarchical neural networks via feedback and attention.

Originality and contribution

Pattern completion and noise filtering are partially contradictory requirements of pattern analysis. What is pattern and what is noise, how to distinguish them; these are the questions we are trying to address in our paper. The goal is to develop a representation, which, by means of a generative network, is able to compare external input and reconstructed input generated by internal representation. Two recent algorithms, sparse code shrinkage (SCS) and non-negative matrix factorization (NMF) are unified in our architecture. SCS performs bottom-up noise filtering. Correlation based pattern completion, however, can not be accomplished by SCS or by related independent component analysis (ICA) methods, given these methods attempt to minimize mutual information amongst transformed components. In our architecture, NMF complements SCS in order to find positively coded sub-structures. These sub-structures, that we shall call components, are ideal means to perform correlation based pattern completion. Pattern completion properties are demonstrated in an extended (hierarchical) architecture with two layers: Lower level sub-architectures monitor parts of a 2-dimensional input space. These parts are arranged in a topographical manner. Robust noise rejection and pattern completion properties are demonstrated in the two layered hierarchical network. The particular combination of the algorithms and the hierarchical data representation is the novelty of our approach. Similar tandem algorithms to our best knowledge, have not been suggested in generative networks. The particular blending combines bottom-up noise filtering and top-down pattern completion.