

Finding Correlations in Multimodal Data Using Decomposition Approaches

Daniel Dornbusch¹, Robert Haschke¹, Stefan Menzel² and Heiko Wersing² *

1- CoR-Lab, Bielefeld University
Universitätsstr. 25, D-33615 Bielefeld, Germany

2- Honda Research Institute Europe GmbH
Carl-Legien-Str. 30, D-63073 Offenbach/Main, Germany

Abstract. In this paper, we propose the application of standard decomposition approaches to find local correlations in multimodal data. In a test scenario, we apply these methods to correlate the local shape of turbine blades with their associated aerodynamic flow fields. We compare several decomposition algorithms, i.e., k-Means, Principal Component Analysis, Non-negative Matrix Factorization and Non-Negative Sparse Coding, with regards to their efficiency at finding local correlations and their ability to predict one modality from another.

1 Introduction

The relationship between local and global structural properties is a key issue in many complex problem domains, e.g., visual perception, motor planning, and data mining. Methods that are capable of identifying local decompositions of large problems into more manageable local elements are important in order to achieve structural robustness and generalization. Unsupervised decomposition algorithms have been investigated extensively in recent years in the field of object and pattern recognition. They find statistically relevant correlations in high-dimensional data sets. Hence, applying a decomposition algorithm to a data set comprising several modalities allows for the identification of inter-modal, semantically meaningful correlations. The simultaneous use of multimodal data in decomposition approaches can: (i) improve the interpretability of the extracted basis components of each single modality, and (ii) extract functionally relevant correlations between different modalities. In the next section, we provide a short introduction to the compared decomposition approaches. Section 3 investigates the application of these methods to a turbine blade test scenario comprising shape information and aerodynamical flow fields. In three studies we compare the abilities of the algorithms at finding local, relevant correlations between both modalities. Finally, Section 4 discusses the results and makes some conclusions.

2 Decomposition of Multimodal Data

The starting point of all decomposition approaches is a set of L vectors pooled in an input matrix, $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_L] \in \mathbb{R}^{M \times L}$. Each data vector, \mathbf{x}_i , can be

*Daniel Dornbusch gratefully acknowledges the financial support from Honda Research Institute Europe.”

regarded as an observation of M random variables. We aim for a more compact, approximate representation of this data matrix using a small set of $N < M \ll L$ meaningful components spanning a new vector space, $\mathbf{F} = [\mathbf{f}_1, \dots, \mathbf{f}_N]$. This assumes that the data vectors, \mathbf{x}_i , are highly correlated and span only a low-dimensional subspace of \mathbb{R}^M . In this case, the basis vectors, \mathbf{f}_i , will express typical correlations within the training set, and also correlations between different modalities of the data. Expressing the data vectors, \mathbf{x}_i , with respect to this new basis yields an approximation matrix $\mathbf{R} = [\mathbf{r}_1, \dots, \mathbf{r}_L]$. The coefficients of the linear combination – also known as encodings – form an $N \times L$ matrix \mathbf{G} . Formally, we can restate this approach as matrix factorization: $\mathbf{X} \approx \mathbf{F} \cdot \mathbf{G} \equiv \mathbf{R}$, $\mathbf{G} = \mathbf{F}^{-1} \cdot \mathbf{X}$, where \mathbf{F}^{-1} is the pseudoinverse of \mathbf{F} . The accuracy of the representation based on the basis, \mathbf{F} , is measured by the reconstruction error, e , between the original data, \mathbf{X} , and its approximation, \mathbf{R} : $e = \|\mathbf{X} - \mathbf{F} \cdot \mathbf{G}\|_2 = \|\mathbf{X} - \mathbf{R}\|_2$. Different decomposition approaches are distinguished by the restrictions imposed on \mathbf{F} and \mathbf{G} . For example, Non-negative Matrix Factorization (NMF) [1] constrains basis vectors and encodings to non-negative values to avoid cancellations of features and facilitate their interpretability. Non-Negative Sparse Coding (NNSC) [2] is a combination of Sparse Coding with the constraints of NMF. It enforces sparseness on the encodings, \mathbf{G} . k-Means clustering [3] represents observations by a set of prototypes, \mathbf{f}_i , resulting in an encoding, which is extremely sparse: only the coefficient associated to the nearest prototype is equal to one. Principal Component Analysis (PCA) [4] computes basis vectors that are pairwise orthogonal and point in the directions of the largest variances.

3 Application to a Turbine Blade Data Set

To investigate the practical abilities of the depicted algorithms at finding local inter-modal correlations, we compare these approaches in a test scenario using turbine blade shapes, which are of similar complexity to the ones used in an evolutionary optimization framework in [5]. The input data consists of 2D blade profiles – represented as binary images – and their associated aerodynamic flow fields – represented by pressure values calculated from a simulated draft (see Fig. 1). For the simulation of the fluid dynamic properties of the blade designs we used an in-house Navier-Stokes flow solver [6]. The set of turbine blades comprises shapes that are not necessarily suitable for practical application and are thus only of academic interest. Any arbitrary turbine shape induces a flow field which can be simulated using Computational Fluid Dynamics (CFD). Contrary to this, flow fields can exist that do not have a corresponding turbine blade shape. We conducted three studies to compare the decomposition approaches based on their efficiency at finding local, relevant correlations and their ability to predict one modality from another. To visualize the flow fields we use an adjusted jet color map (see Fig. 1(c)). It maps small absolute values to black in order to show their marginal influence in linear combinations. Larger positive and negative numbers are mapped onto ranges of red to blue colors respectively. Shapes are depicted on top of the corresponding flow fields using shades of gray.

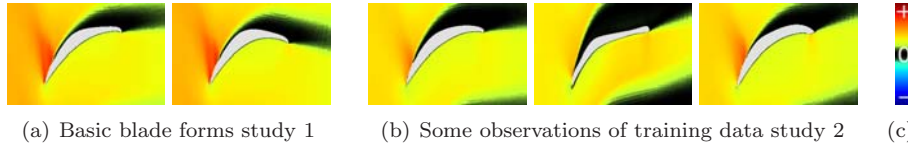


Fig. 1: (a) Training data of first study are limited to variations of two basic blade shapes differing in trailing edges. Associated aerodynamical flow fields represented by pressure values are calculated using CFD to simulate a draft. (b) Training data used in studies 2 and 3 have no shape restrictions. (c) Color map.

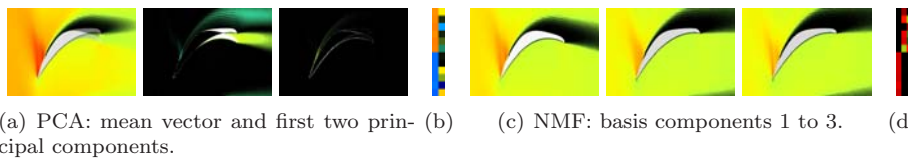
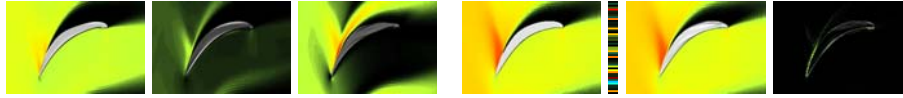


Fig. 2: Study 1 decomposition results: PCA: orthogonal basis components (a) and encodings (b) have arbitrary signs. Superposition of positive and negative parts. NMF: holistic, non-sparse basis components (c) and sparse encodings (d) caused by nature of training data, limited to non-negative values.

3.1 Study 1: Identifying Two Basic Blade Shapes

To verify the ability of the decomposition algorithms to identify common constituents of a data set, the first study considers a simple toy data set consisting of two basic blade shapes as shown in Fig. 1(a). Both shapes have a common front part and only differ in their trailing edges. The data set contains six small variations of each basic shape. We always compute three basis components, although two should suffice. Hence, we expect the third component to have much less of an influence on the reconstruction error than the first two. k-Means clustering produces prototypical basis vectors that are holistic and non-negative (due to the non-negativity of the training data). The encodings are maximally sparse and orthogonal. As depicted in Fig. 2(a) and 2(b), PCA extracts the mean vector of the training data and the two principle components related to the highest eigenvalues. All three features are holistic and the two principle components are orthogonal. The values of the encodings and the basis vectors have arbitrary signs. Using the mean vector and the first principle component, both basic blade shapes can be restored by summation or subtraction resulting in cancellations between positive and negative parts. The second principle component contributes to a further reduction of the reconstruction error. NMF (see Fig. 2(c) and 2(d)) is able to produce either sparse basis components or sparse encodings [7]. Due to the training data presented in the first study, NMF generates holistic, prototypical basis components and sparse encodings that consist of non-negative values. NNSC extracts prototypical, holistic basis components that represent the basic turbine blade forms. The encodings are extremely sparse.



(a) Representative holistic and sparse basis (b) Original, encodings, reconstruction, absolute difference.

Fig. 3: (a) Study 2: three out of the 20 basis components calculated by NMF on 250 observations. (b) Study 3: encodings were calculated only on the pressure values and then were used to reconstruct the missing shape modality.

Approach	(a) Study 2 Observed Data					(b) Study 2 Test Data					(c) Study 2				(d) Study 3	
	5	10	20	35	49	5	10	20	35	49	Ob. Data Sep.	Com.	Test Data Sep.	Com.	PR	SR
K-Means	52	36	29	23	21	87	88	78	78	77	19	21	70	77	80	104
PCA	38	28	19	13	10	74	53	44	38	34	9	10	32	34	38	57
NMF	40	31	22	17	15	74	59	48	38	36	13	15	32	36	51	71
NNSC	40	31	22	17	14	74	66	46	40	37	13	14	34	37	43	60

Table 1: Study 2: Normalized mean squared reconstruction errors (NMSE) for observed (a) and test data set (b) against number of basis components (BC). (c): 49 BC, Sep.: separate decomposition of modalities, Com.: combined decomposition of modalities. Study 3: Reconstruction of missing modalities (d), PR: pressure reconstructed, SR: shape reconstructed. Values scaled by factor 10^3 .

3.2 Study 2: No Shape Restrictions

In the second study, we appraise functional components that emerge from turbine blades not having any form restriction. In order to cover a higher variance of blade shapes we generate 250 diverse observations consisting of both modalities, shape and pressure. Figure 1(b) shows some of the training samples. We extract subspaces spanned by 5, 10, 20, 35 or 49 basis components. The variation of the number of basis vectors allows us to analyze the reconstruction errors in more detail. Additionally, we use a disjoint test data set comprising 65 observations to measure the ability to generalize and represent novel though similar data. The results of k-Means clustering are very similar to those of the first study. The basis components are prototypical and the encodings maximally sparse. The ability to generalize is limited to the selection of the nearest cluster centroid. The holistic principal components of PCA accomplish very small reconstruction errors, but are difficult to interpret due to positive and negative values. The basis components of NNSC are prototypical and holistic. They represent the basic blade forms of the training set. The encodings are sparse and in most cases observations are reconstructed by three or less activated features. NMF shows fundamental changes in comparison to the first study. For this data set most basis components and the encoding matrix are sparse. Figure 3(a) depicts three representative basis vectors calculated by NMF. The first illustrated component represents a holistic basic shape and its associated flow field. The second feature shows an interesting correlation between the lower, frontal surface and the pressure profile above the turbine blade and the third image displays a basis

vector focusing mainly on a single modality, the pressure profile. All NMF runs show qualitatively similar components, independent of their number. As can be seen from Tab. 1(a) and 1(b), the normalized mean squared reconstruction error, $NMSE = \frac{1}{L} \sum_{i=1}^L \frac{\|\mathbf{x}_i - \mathbf{r}_i\|^2}{\|\mathbf{x}_i\|^2}$, decreases for all methods uniformly, if the number of available basis components (BC) is increased from 5 to 49. It is shown that the subspace trained on the observed data is able to represent novel test data. However, in this generalization test the reconstruction errors are slightly increased. We also evaluated the impact of multimodality on the reconstruction error. To this end, Tab. 1(c) compares the reconstruction errors obtained from separate decomposition of both modalities and subsequent accumulation of the individual errors (column “Sep.”) with those reconstruction errors obtained from joint decomposition of both modalities (column “Com.”). Regarding the reconstruction error, PCA performs best. It is closely followed by NMF and NNSC. The relative generalization ability of NMF is unmatched. k-Means is outperformed by the other methods in reconstruction of the observed training and novel test data set. For all evaluated algorithms we find that the multimodality of the data slightly degrades reconstruction performance. Additional constraints such as non-negativity, sparseness or orthogonality further increase the reconstruction error. Nevertheless, these factors render it possible to extract more *meaningful* features and to find *local* correlations between modalities, which increase the quality of the basis vectors, i.e. their relevance and interpretability.

3.3 Study 3: Reconstruction of Missing Modalities

In the final study, we focus on the ability to reconstruct missing modalities based on the inherent correlations between the components’ subparts. The subspaces comprising 49 basis vectors computed for the second study on 250 observations are used to represent novel test data consisting of only shape or pressure information. The basis vectors are specialized to represent turbine blades. To project the test observations into the spanned subspace, we utilize the Moore-Penrose pseudoinverse. Some negative encodings are introduced, because the least squares solution for the linear combination is not restricted to non-negative contributions. However, this does not prevent the estimation of missing modalities, if we use those encodings to reconstruct all modalities, including the previously neglected ones. If the test data set is similar to the training observations used to calculate the features, the missing information is approximated successfully. Figure 3(b) depicts the reconstruction of a missing shape modality for a novel test observation. The reconstruction errors for the approximation of missing modalities can be seen in Tab. 1(d). “PR” stands for pressure reconstructed. This means that the encodings were estimated solely on the shape information and then used to reconstruct form and pressure. “SR” describes the counterpart: shape reconstructed. The encodings are calculated using only the flow fields and then utilized to estimate both modalities, including the former missing shape. PCA excels due to its strong capability at finding correlations. NNSC performs second best using its holistic, non-negative features. NMF represents a good

compromise between interpretability and reconstruction quality. K-Means' lack of flexibility caused by its limitation to only select the nearest cluster mean results in the highest reconstruction errors. Nevertheless, all the evaluated algorithms are able to approximate missing modalities successfully.

4 Conclusion

In this paper, we have proposed the application of standard decomposition approaches to find local correlations in multimodal data. In a test scenario, we applied these methods to correlate the local shape of turbine blades with their associated aerodynamic flow fields. We compared the algorithms with regards to their efficiency at identifying local, relevant correlations and their ability to predict one modality from the other. We found that the multimodality of the data and additional constraints such as non-negativity, orthogonality or sparseness can improve the algorithm's ability to extract meaningful basis components with a trade-off of a small decrease in the reconstruction performance. With regards to the reconstruction error, PCA performs best, because it decorrelates the input data and increases its statistical dispersion. NSC is strong at reconstructing missing modalities using non-negative features. NMF revealed an interesting inter-modal correlation between a shape feature on the lower side of the blades and the aerodynamical flow field on their upper side. Further, it was possible to predict one modality from the other, which is evidence for the learned inter-modal correlation and might turn out to be useful in design optimization processes. New design concepts could be generated by combining different local contributions to form a desired global flow field and selecting the associated shape primitives accordingly. Also, interchangeable local features might be identified that have similar influence on the air flow but have different shapes.

References

- [1] D. D. Lee and H. S. Seung. Learning the parts of objects by non-negative matrix factorization. *Nature*, 401(6755):788–791, October 1999.
- [2] P. O. Hoyer. Non-negative sparse coding. In *Proc. 12th IEEE Workshop on Neural Networks for Signal Processing (NNSP '02)*, pages 557–565, 2002.
- [3] J. B. MacQueen. Some methods for classification and analysis of multivariate observations. In *Proc. 5th Berkeley Symposium on Mathematical Statistics and Probability (BSMSP '67)*, pages 281–297. University of California Press, 1967.
- [4] K. Pearson. On lines and planes of closest fit to systems of points in space. *Philosophical Magazine*, 2(6):559–572, 1901.
- [5] S. Menzel and B. Sendhoff. Representing the change - free form deformation for evolutionary design optimisation. In T. Yu, D. Davis, C. Baydar, and R. Roy, editors, *Evolutionary Computation in Practice*, chapter 4, pages 63–86. Springer, Berlin, 2008.
- [6] T. Arima, T. Sonoda, M. Shirotori, A. Tamura, and K. Kikuchi. A numerical investigation of transonic axial compressor rotor flow using a low-reynolds-number k-e turbulence model. *ASME Journal of Turbomachinery*, 121(1):44–58, 1999.
- [7] A. Pascual-Montano, J. M. Carazo, K. Kochi, D. Lehmann, and R. D. Pascual-Marqui. Nonsmooth nonnegative matrix factorization (nsNMF). *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 28(3):403–415, 2006.