

# Abstract

This thesis with the title "**New extensions and perspectives on Slow Feature Analysis**", written by **Merlin Schöler**, focuses on the extension and understanding of Slow Feature Analysis (SFA), an unsupervised machine learning approach to identify low-dimensional latent factors from high-dimensional data by leveraging temporal consistency or, more generally, pairwise symmetric relationships between data points. The presented research covers different contributions in three areas:

**Gradient-based SFA** This work proposes a differentiable whitening modification to be added to differentiable parametric models, such as deep neural networks, that forces the model outputs to obey the constraints of the SFA optimization problem, without sacrificing the ability to train end-to-end. The efficacy of such augmented architectures for extracting interpretable SFA features from synthetic data, ego-visual time series, and images with graph structure are shown, followed by demonstrations how this approach can significantly improve the performance in nonlinear multilayer models when compared to classical SFA.

**Features from Directed Behavior** It is examined how goal-directed behavior influences optimal slow features. Using ergodic Markov chains as a simple model of spatially connected environments, optimal slow features under directed behavior are analytically derived, contrasting these with features obtained from homogeneous exploration, which are typically used for SFA. This work confirms previous findings and shows significant effects on feature scaling modulated by degree of goal-affinity, highlighting the impact on tasks such as value function approximation in reinforcement learning, and thereby possibly explaining previous negative results when learning slow features in parallel to behavior optimization through general misalignment of these objectives. To reduce detrimental effects, three corrective routes are proposed and their effect on approximation quality examined.

**A relationship between SFA and Variational Inference** Reinterpreting SFA through the lens of probabilistic modeling, this work reveals a connection between probabilistic inference and a soft variant of SFA. The latter can be derived as a special case of amortized variational inference under Gaussian mean-field assumptions for latent variables with dynamical linear process priors. As opposed to previous work, no assumptions on linearity are made for the mapping from latent to observed variables; however, for the linear case, a deterministic formulation of the evidence lower bound objective that does not require Monte Carlo estimation is provided including its necessary optimality conditions. Lastly, the implications of this perspective and avenues for further research are discussed.

Together, these contributions aim to advance the theoretical and practical understanding of SFA and to provide a solid foundation for future research focused on the slowness principle.

# Kurzbeschreibung

Diese Dissertation mit dem Titel "**New extensions and perspectives on Slow Feature Analysis**", verfasst von **Merlin Schüller**, beschreibt Erweiterungen und neue Perspektiven auf die Slow Feature Analysis (SFA), ein unüberwachtes Lernverfahren zur Identifikation niedrig-dimensionaler latenter Merkmale aus hoch-dimensionalen Daten, welches sich der Annahme der Langsamkeit, oder allgemeiner, paarweiser symmetrischer Beziehungen zwischen Datenpunkten bedient. Die vorgestellte Forschung behandelt unabhängige wissenschaftliche Beiträge in drei Ausprägungen:

**Gradienten basierte SFA** Diese Arbeit schlägt eine differenzierbare Modifikation als Ergänzung differenzierbarer parametrischer Architekturen, wie beispielsweise tiefer neuronaler Netze, vor. Diese erwirkt, dass transformierte Variablen die Nebenbedingungen des SFA-Optimierungsproblems erfüllen, ohne jedoch die Möglichkeit des Ende-zu-Ende Trainings einzuschränken. Die Möglichkeit solcher augmentierter Architekturen interpretierbare SFA Merkmale zu extrahieren wird anhand synthetischer Daten, ego-visueller Zeitreihen, und eines Bild Datensatzes mit Graphenstruktur demonstriert. Weiterhin wird eine Verbesserung der Modellperformance auch für klassische mehrschichtige SFA Architekturen aufgezeigt, sowie mit ähnlichen Ansätzen verglichen.

**Merkmale gerichteten Verhaltens** In diesem Teil der Arbeit wird untersucht, wie zielgerichtetes Verhalten optimale langsame Merkmale beeinflusst. Hierbei werden ergodische Markov-Ketten als einfaches Modell räumlich verbundener Umgebungen verwendet, optimale langsame Merkmale analytisch hergeleitet und diese Merkmale mit denen uniformer Exploration, wie sie für SFA üblich ist, verglichen. Die Ergebnisse bestätigen frühere Erkenntnisse und zeigen signifikante Einflüsse auf die Skalierung optimaler Merkmale auf, die durch die Stärke der Zielaffinität moduliert werden. Hierbei werden auch Konsequenzen für das Approximieren optimaler Wertefunktionen, wie im Reinforcement Learning üblich, hervorgehoben. Als konstruktive Maßnahmen werden drei mögliche Korrekturrouten vorgeschlagen und evaluiert.

**Zusammenhang zwischen Slow Feature Analysis und Variational Inference** Durch eine Neuinterpretation der Slow Feature Analysis im Rahmen probabilistischer Modellierung wird in dieser Arbeit ein Zusammenhang zwischen probabilistischer Inferenz und einer Variante der SFA mit Soft-Constraints aufgezeigt. Diese Formulierung von SFA wird als Spezialfall der amortisierten variationalen Inferenz hergeleitet, wobei eine stark eingeschränkte Mean-Field-Approximation für latente Variablen eines Gaußschen dynamischen Modells verwendet wird. Im Gegensatz zu vergleichbaren Ansätzen wird in dieser Herleitung auf Linearitätsannahmen in der Abbildung latenter zu observabler Variablen verzichtet, jedoch wird für den linearen Fall eine deterministische Formulierung der Zielfunktion vorgestellt, die ohne Monte-Carlo-Schätzung auskommt, und es werden kompakte Optimalitätsbedingungen hergeleitet. Abschließend werden die Implikationen dieser probabilistischen Perspektive sowie mögliche Ansätze für weiterführende Forschung diskutiert.

Diese Beiträge zielen darauf ab, sowohl das theoretische als auch das praktische Verständnis der Slow Feature Analysis zu erweitern und eine solide Grundlage für zukünftige Forschungsarbeiten, die auf dem Prinzip der Langsamkeit beruhen, zu schaffen.