**INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI**
**SHORT ABSTRACT OF THESIS**

| | | |
|---|---|---|
| Name of the Student | : | Neeraj Kumar |
| Roll Number | : | 11610210 |
| Programme of Study | : | Ph.D., EEE |

Thesis Title: A few algorithms for inverse problems in image processing.

| | | |
|---|---|---|
| Name of Thesis Supervisor(s) | : | Dr. Amit Sethi |
| Thesis Submitted to the Department/ Center | : | August 17, 2016 |
| Date of completion of Thesis Viva-Voce Exam | : | February 25, 2017 |
| Key words for description of Thesis Work | : | Super resolution, non-negative matrix factorization, sparse representation, polynomial neural networks, regression, etc. |

## SHORT ABSTRACT

Inverse problems are the problems that involve estimation of the parameters or data given certain observations. The available observations are often noisy and incomplete. Hence, the inverse problems are mostly ill-posed as there can be infinitely many solutions to a given inverse problem. Some priors based on the understanding of the physical phenomenon in the forward problem are used to estimate the solution of such problems. In this work, we provide algorithms for solving two inverse problems, (1) learning based single image super resolution and (2) reducing the solution space of non-negative matrix factorization. Single image super resolution (SISR) refers to the problem of estimating a high resolution (HR) from a single low resolution (LR) image. SISR is an ill-posed inverse problem as for SISR zoom by a factor of s in each dimension, there are $s^2$ unknown HR pixels for each known LR pixel. Natural image statistics and an understanding of human visual perception, such as importance of accurate edge reconstruction, have been used to formulate image priors that can constrain the solution space of the desired HR image pixels. In this work, we present two SISR approaches that exploit scale invariant statistics of natural images.

In the first approach, we present an algorithm to learn a direct mapping from LR space to the corresponding HR space. Our method not only tests (runs) and trains faster, it also requires a smaller number of training samples while achieving competitive reconstruction accuracy compared to other learning-based methods. We met these objectives by posing the HR reconstruction problem as an estimation of a function to predict the pixels of an HR patch using its corresponding LR pixels and their spatial neighborhood. We studied the impact of varying the input LR and output HR patch sizes and gained the following insights: Reconstruction accuracy for a given output HR patch size improves when input LR patch size is increased, but the improvement saturates after including a few extra layers of LR pixels. Moreover, HR reconstruction accuracy is the highest when the output HR patch is restricted to only that which corresponds to one LR pixel. We used zero component analysis as a pre-processing step to enhance the estimation optimization energy on perceptually salient features such as edges. We also tapped into the ability of polynomial neural networks to hierarchically learn refinements of a

nonlinear function to model the mapping from LR to HR patches. Accurate HR reconstruction with small input and output patch sizes not only makes learning more efficient, it also indicates that SISR is a highly local problem. In contrast, a recently-proposed and related technique using convolutional neural networks needs much larger training set and longer training time because of larger input-output patch sizes and a computationally expensive learning algorithm. Lesser training overhead (time and number of samples) and faster run time performance without compromising SISR reconstruction accuracy makes our technique quite attractive for real-time zooming applications even on resource-constrained computing platforms such as mobile phones and hand-held cameras.

In the second approach, we propose an SISR algorithm using wavelet decomposition and machine learning. While inter-scale wavelet properties have been exploited for SISR, we also use their intra-scale properties, which have not received much attention. We explore various sources of information within a wavelet transform that can be used to reconstruct the desired HR image. For SISR, the spatial neighborhood of known wavelet coefficients at the coarser scale, the co-parents located at the same coordinates in the other sub-bands and the finer scale approximation coefficients can be used to estimate the desired wavelet coefficients at the corresponding finer scale. Additionally, our algorithm uses a discriminative framework based on nonlinear regression, which provides a deterministic output with fewer training samples. Other methods based on generative models struggle to estimate the direction and magnitude of the wavelet coefficients at a desired level as they deliver probabilistic outputs even with a large number of training samples. The proposed algorithm fared better than the state-of-the-art wavelet-based and other SISR methods in terms of HR reconstruction accuracy, edge reconstruction quality, training overhead (time and number of samples), and testing time. We attribute the enhanced accuracy of our algorithm to incorporation of spatial context from the neighborhood of the input wavelet coefficients. Interestingly, the accuracy reached a maximum beyond which increasing neighborhood size has no advantage, suggesting that HR reconstruction is a local problem. Limiting the size of the neighborhood allowed our algorithm to learn fast from fewer samples. We also present a potential application of SISR in histology.

Another problem that we address in this work is reducing the solution space of non-negative matrix factorization. Non-negative source separation consists of estimating a set of unknown signals (called sources) from a set of observed mixtures, which are linear combinations of the sources with a constraint that the mixing coefficients and the source signals are non-negative quantities. Nonnegative matrix factorization (NMF) can be used to solve such problems. However, the main issue before designing a separation algorithm using NMF is to assess whether the problem admits a unique solution. If no unique solution is possible then there is a need to select a best plausible solution among the candidate ones. In this work we propose a data transformation algorithm to reduce the NMF solution space. Our algorithm is based on the geometrical insight that if the transformed data is sparse and lies on the axes and faces of the positive orthant in the transformed space, then it reduces the number of solutions in that space and even allows a unique solution under certain conditions. We propose that such a space is spanned by atoms of a data-derived over-complete dictionary if the original data is sparse with respect to the dictionary. The dictionary itself also serves as a transformation matrix between the original space and the transformed space. We call this one-side sparse NMF (OS-SNMF), and demonstrate it's several desirable properties. It facilitates the reduction of solution space and accurate recovery of true underlying basis. In this work we apply the developed methodology for solving source separation problem in hyperspectral images.