A NONLINEAR DICTIONARY FOR IMAGE RECONSTRUCTION

Mathiruban Tharmalingam
Department of Electrical Engineering& Computer Eng.
Ryerson University

Kaamran Raahemifar
Department of Electrical Engineering& Computer Eng.
Ryerson University

ABSTRACT

Complex signals such as images, audio and video recordings can be represented by a large over complete dictionary without distinguishable compromise on the representation quality. Large over complete dictionaries with more patterns can be used to increase the sparse coding as well as provide significant improvements in signal representation quality. The use of the over-complete dictionaries and sparse coding has been successfully applied in compression, de-noising, and pattern recognition applications within the last few decades. One particular dictionary, the Discrete Cosine Transform (DCT) dictionary has seen a great deal of success in image processing applications. However, we propose a novel non-linear over-complete dictionary that is sparser than the DCT dictionary while improving the quality of the signal representation. The proposed non-linear dictionary has demonstrated through experimental results to be superior to the DCT dictionary by achieving higher signal to noise ratio (SNR) in the reconstructed images.

Index Terms— Sparse coding, Non-linear dictionaries atoms, DCT dictionary, image reconstruction

1. INTRODUCTION

There has been a high level of interest in the recent decade in sparse coding. Sparse coding and over complete dictionaries can be applied individually or in conjunction in compression, feature extraction, pattern recognition, image reconstruction, image de-noising, and many more applications. In most image processing applications, the first step is to map the image into a transformed space where the signal can be represented in a sparse matter. One of the most frequently used transform for this application is the DCT due to its simplicity and robustness in signal representation. Once the transformation dictionary is known then the sparse coding algorithms solves the signal representation problem, where a signal $y \in R^m$ is represented using a small number of non-zero coefficients in the source vector $x \in R^n$ satisfying the linear model $\|Y - DX\|_2 = 0$ for exact representation and $\|Y - DX\|_2 < \varepsilon$ for an approximation of the signal. In the above linear model, D is the known dictionary, $D \in R^{mxn}$ and X is the sparse vector set and ε is the desired representation error. The sparse coding problem can be stated as an optimization problem shown in equation (1.1) using the L0 norm on the sparse vector [1]. The L0 norm counts the number of non-zero coefficients in the vector ignoring its magnitude.

$$\min_x \quad \|x\|_0 \quad \text{such that} \quad \|Y - DX\|_2 < \varepsilon \qquad (1.1)$$

There has been a growing interest and research into the type of dictionaries that can be used for sparse coding. There are two categories of dictionaries, the pre-defined dictionary such as the DCT, Gabor function dictionaries or the adaptive dictionaries. The adaptive dictionaries are dictionaries that adapt to the training samples and are shown to have excellent performance [2] but they are neither simple to create nor as robust as the DCT for some applications. The dictionary learning process is a time consuming process. In image reconstruction, a predefined dictionary is preferred to a learned dictionary because a learned dictionary will change with training samples, and the quality of the learned dictionary depends on the training samples. One of the biggest benefits of the predefined dictionary is it does not have to accompany the given sparse vector set to represent the data and thus it can be used to transfer images and data since the predefined dictionary can be recreated at the receiving side with minimal information. This is one of the reasons why the DCT dictionary has seen success in image reconstruction and compression application. However the question this paper asks, is the DCT the best predefined dictionary that can be used? We propose a novel pre-defined set of non-linear atoms to create a dictionary that outperforms the DCT dictionary.

The remainder of this paper is formatted as follows. Section 2 will contain a brief discussion on the available dictionaries such as the DCT dictionary, and adaptive dictionaries. Section 3 will present the formulation of the non-linear dictionary. Section 4 will be dedicated for experimental results and analysis. Finally, the paper will present its future work and conclusion in section 5 and 6 respectively.

2. TYPES OF EXISTING DICTIONARIES

The DCT dictionary is a predefined dictionary and is widely accepted because it can be easily implemented. The atoms of a DCT dictionary d_i are shown in (1.2) where m is the size of the discrete signal to be represented, and L is a fixed constant that controls the fixed frequency spacing of the dictionary [3]. The atoms of the dictionary are the column vectors of the dictionary. An over-complete dictionary is a dictionary where the number of column vectors is much more the number of row vectors. Given a signal $y \in R^m$, and a dictionary $D \in R^{mxn}$, the dictionary is said to be over-complete if n >> m. An over-complete DCT dictionary is used to evaluate the proposed hybrid dictionary's performance.

$$d_i = \cos\left(\frac{2\pi}{L} i\ k\right) \text{ where k = 1 , 2, 3 ... m} \quad (1.2)$$

The adaptive dictionaries are relatively new to the field. The learned dictionaries have shown promising results in experimental testing but the cost of implementation has been high. Most of the development in this field has been to employ parametric models in the training process to create a structured dictionary. Some of the techniques used to derive dictionaries are the lasso problem method, method of optimal directions (MOD) the generalized PCA [4], the Focuss-CNDL method [5], and the K-means, K-SVD [2]. Each algorithm has its benefits as it is a compromise between speed and the quality of the dictionary. These methods use different algorithms to solve either the problem stated in (1.3) or(1.4). The difference between the two optimization equations is the constraint on the sparseness. In equation (1.3) the L0 norm is taken on the vector, so all non-zero coefficients have equal weight since L0 norm counts the number of non-zero components, it imposes the sparse condition. In equation(1.4), the sparse restriction is relaxed, the L1 norm adds the weight of each coefficient, and therefore it can lead to a solution with many small non-zero components. The MOD, K-means and K-SVD all use equation(1.3). The MOD algorithm uses the Moore – Penrose pseudo-inverse to solve the problem $Y = DX$, even though it converges in a few iteration counts, the matrix inverse operation is very costly. The K-SVD again solves the optimization problem in (1.3) but instead of updating the entire dictionary, it updates a single atom at a time making the K-SVD a faster method then the MOD.

$$\min_{x,D} \quad \|Y - DX\|^2 \text{ subject to } \|x_i\|_0 < S \ \forall i \quad (1.3)$$

$$\min_{x,D} \quad \|Y - DX\|^2 \text{ subject to } \|x_i\|_1 < S \ \forall i \quad (1.4)$$

However for all methods, the quality of the learned dictionary depends heavily on the training data samples used; if the training samples are closely correlated then the resulting learned dictionary is not robust. If limited training data is used, the learned dictionary does not contain enough diverse atoms and will produce poor results when reconstructing other images. The time to run the learning algorithm grows with the number of training samples, so having a large training set is infeasible. The new developments in dictionary learning are attempting to combat these issues, making progress in online dictionary learning, unsupervised learning and so on to allow for large training data. However the idea of a simple predefined robust dictionary is an attractive option this paper is exploring.

3. NON LINEAR DICTIONARY ATOMS

The DCT dictionary is a simple, robust dictionary; however it is composed of only harmonic signals. Even though it is possible to reconstruct any signal with appreciable quality using the well-established DCT technique, it does not provide a sparse representation for all signals. Also recent work on dictionary learning has shown that the DCT dictionary is lacking some vital atoms. Intuitively the DCT atoms should be able to sparsely represent smooth harmonic signals but will have difficulty representing transients, asymptotic behavior, discontinuities and non harmonic signals such as rational, logarithmic and exponentials. Consider the task of representing a non-harmonic signal y = ax+b. The DCT dictionary is unable to represent this signal sparsely but by including polynomial atoms such as x, the new hybrid dictionary can represent the signal sparsely. Since the nature of the signal to be coded is unknown we propose to include a variety of functions some rational and logarithmic functions to handle asymptotic behavior, along with polynomial and exponential functions to provide support for smooth functions. The addition of new atoms obviously increases the size of the dictionary and the complexity of representing the dictionary. This paper sets out simple schemes to produce a structured hybrid dictionary with the above mentioned additional nonlinear atoms while maintaining the benefit of a predefined dictionary that is being able to recreate the dictionary without requiring each value of the dictionary. The proposed hybrid dictionary has an advantage over the adaptive dictionary because it does not require a training set. We expect the atoms with asymptotic behavior to be useful in representing natural signals with transients such as an image of a night sky with stars or percussion sounds in audio signals, while the addition of polynomials and exponentials will aid in representing smooth non-harmonic signals. However, the exponential and polynomial functions rapidly increase and become uncontained which make these ineffective, however if the region of the function, the start and end points, are selected carefully then it is shown to be effective.

The simplest dictionary extension is created by taking the cosine function to a higher order power, and making non-linear combinations. Given a small DCT dictionary DCT1 with atoms shown in equation (1.2) we perform atom multiplication and take the DCT to a higher order. We extend the given DCT dictionary by applying equation (1.5) to each of the atoms. We are effectively creating a third atom, cos(x)cos(y) = cos(x - y) - sin(x)sin(y). This can be viewed as a frequency modulator. The noteworthy part of this modification is that it can be fully described using the initial frequency, the fixed frequency spacing, the number of atoms in the dictionary and the number of times the dictionary extension was applied. This dictionary contains all the benefits of the DCT with only single additional information that is required by the receiving side to recreate the entire dictionary. The algorithm to extend the DCT dictionary is shown in Algorithm 1.

$$dext_i = d_i d_j = \cos\left(\frac{2\pi}{L} i\ k\right)\cos\left(\frac{2\pi}{L}\mathrm{j}\ k\right) \quad (1.5)$$

```
Given Dictionary DCT1 ∈ R^mxn
DCT_EXT = DCT1
for i = 1, i < n, i++
  for j = i, j < n, j++
    for k = 1, k < m, k++
        Atom[i,1] = DCT[k,i] × DCT[k,j]
    end
   DCT_EXT = [DCT_EXT   Atom]
  end
end
```

Algorithm 1 DCT dictionary extension

The discrete polynomial atoms are created using the formula in (1.6) where *m,* is dictated by the number of rows in the dictionary, the size of the signal. In the equation (1.6) k is a set of integers going from 1, 2, 3 ... m. There are also a predefined set of coefficients {a, b, c}. For our experimental testing we generated the coefficients using a fixed spacing scheme such that $a_i = a_1 + \Delta i$, where i determines the number of atoms from each order of the polynomial. The rational function set is created simply by selecting (c < 0) coefficients in equation(1.6). During our testing, the same set of coefficients for both (a and b) were used to generate both the polynomial and rational function set. We also limited the polynomial order, variable c, in equation (1.6), to be between (-3 and 3) with fixed spacing. This main purpose of using a fixed spacing scheme and the same constants for the polynomial and rational functions set is to reduce the complexity and the data required to recreate the dictionary.

$$p_i = (a_i x[k] + b_i)^{c_i} \quad (1.6)$$

The log function set was created by passing the polynomial set into the log transform. The exponential set was created by passing all the root functions to the exponential transform and the algorithm is shown in Algorithm 2. A similar algorithm is used to create the log dictionary using only the polynomial dictionary set. Once the dictionary is compiled, the generated atoms were scanned for discontinuities and invalid numbers and corrected by choosing the previous valid value or next valid value in the atom.

```
Given Dictionary D1 ∈ R^mxn
EXP_EXT = D1
for i = 1, i< n, i++
     for k = 1, k < m , k++
          Atom[k,1] = e^D[k,i]
     end
    EXP_EXT = [EXP_EXT Atom]
End
```

Algorithm 2 Exponential dictionary atom expansion

Once the atoms are corrected and all numbers are valid, the dictionary columns are normalized, and duplicate atoms are removed. This step is necessary for speeding up the sparse coding process. The final dictionary composition is summarized in Table 1 Nonlinear atoms of the hybrid dictionary which is a combination of the root atoms such as the DCT dictionary, higher order DCT extension, a set of polynomial functions, rational functions and logarithmic functions, and exponential functions set composed of the root atoms, which gives this dictionary composition a great variety of atoms that can represent signals efficiently.

Table 1 Nonlinear atoms of the hybrid dictionary

$\cos\left(\frac{2\pi}{L} i\ k\right)$
$\cos\left(\frac{2\pi}{L} i\ k\right)\cos\left(\frac{2\pi}{L}\mathrm{j}\ k\right)$
$\left(a_i x[k]+b_i\right)^{c_i}$
$\log\left(\left(a_i x[k]+b_i\right)^{c_i}\right)$ for $c_i > 0$
$e^{\cos\left(\frac{2\pi}{L} i\ k\right)}$
$e^{\cos\left(\frac{2\pi}{L} i\ k\right)\cos\left(\frac{2\pi}{L}\mathrm{j}\ k\right)}$
$e^{\left(a_i x[k]+b_i\right)^{c_i}}$

4. EXPERIMENTAL RESULT

The proposed dictionary was tested by reconstructing a test image of Lena using orthogonal matching pursuit (OMP) sparse coding technique. The OMP method was selected due to its speed when dealing with large dictionaries. The reader is referred to the literature on OMP and other greedy techniques for solving the sparse coding problem [6] [7] [8]. The 512x512 pixel Lena grey-tone image was broken down into 8x8 patches and sparse coding was applied on these patches. The original image and the reconstructed images with 7 non-zero constants are shown in Figure 1. The reconstructed image using the hybrid dictionary is sharper then the image reconstructed with the DCT dictionary. In Figure 2 the peak signal to noise ratio (PSNR) of the reconstructed image is shown. The PSNR is calculated using equation(1.8). The graph in Figure 2 shows that the non-linear dictionary performs better than the DCT dictionary. The difference in representation quality also increases quicker as the number of non-zero elements in the sparse vector is increased. The hybrid dictionary with 20 coefficients has an SNR of 43 dB while the DCT is at 36 dB which is a difference of 7 dB.

$$MSE = \frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}\left(Y(i,j) - Y\_recon(i,j)\right)^2 \qquad (1.7)$$

$$PSNR = 20\log_{10}\left(MAX\right) - 10\log_{10}(MSE) \qquad (1.8)$$

Figure 1 the original image of Lena is shown on top, the reconstructions using DCT and hybrid dictionary is shown on left right respectively.

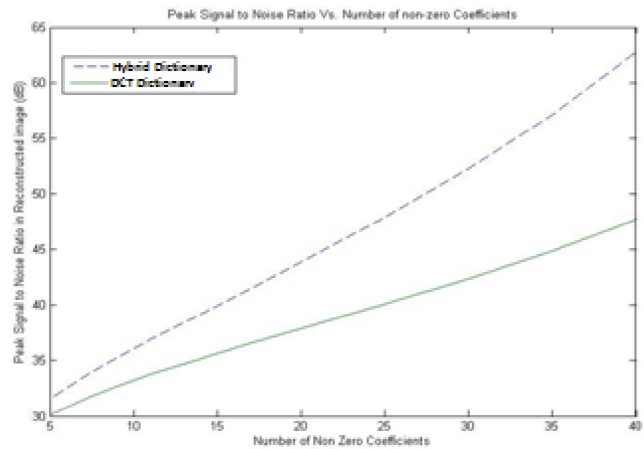

Figure 2 PSNR of Reconstructed Lena image

5. FUTURE WORK

The current results are as we expected, the addition of the non-linear, non-harmonic functions has improved the representation quality. Being inspired by the results, future work can focus on developing a parameter screening method to use the signals and select only the vital functions thus reducing the size of the hybrid dictionary. We expect the parameter screened hybrid dictionary to produce a predefined dictionary with some adaptive properties that will help improve signal reconstruction quality. Also, since these nonlinear atoms are not orthogonal, the OMP sparse coding algorithm may not be the best suited for selecting the proper atoms to represent the signals. The OMP algorithm was selected primarily for its speed, simplicity of implementation, and it performs well on a DCT dictionary, but future work should explore alternatives to see if there are any improvements to be gained.

5. CONCLUSION

The paper proposed a novel idea to create a predefined DCT-hybrid dictionary by extending the existing DCT dictionary. Instead of only including the DCT atoms, the paper suggested the idea of using other non-linear functions such as polynomials, rational, logarithmic, exponential, and phase shifted higher order cosine functions creating a diverse set of dictionary atoms that are able to reconstruct not only smooth but also transient signals. The performance of the hybrid dictionary was compared to an over complete DCT dictionary. The hybrid dictionary has performed as well or better than the DCT dictionary in image reconstruction; it was able to provide a reconstructed image with a higher SNR and a sparser representation. Further testing and a parameter screening method to reduce the size of the dictionary will make the hybrid dictionary efficient and practical.

7. REFERENCES

[1] Michal Aharon, Michael Elad, and Alfred Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," in *IEEE Transactions on Signal Processing*, 2006, pp. 4311-4322.

[2] Micheal Elad and Michal Aharon, "Image denoising via sparse and redundant representation over learned dictionary," in *IEEE Transaction on Image Processing*, 2006, pp. 3736-3745.

[3] James Bowley and Laura Rebollo-Neira, "Sparse image representation by discrete cosine/B-spline dictionaries," Aston University, Birimingham,.

[4] Ron Rubinstein, Alfred M.Bruckstein, and Micheal Elad, "Dictionaries for sparse representation modelling," in *Proceedings of the IEEE*, 2010, pp. 1045-1057.

[5] Joseph F. Murray and Kenneth Kreutz-Delgado, "Sparse image coding using learned overcomplete dictionaries," in *IEEE Workshop on Machine Learning for Signal Processing*, 2004, pp. 579-588.

[6] Stephane G Mallat and Zhifeng Zhang, "Matching Pursuit with Time-Frequency Dictionaries," in *IEEE Transactions on signal Processing*, 1993, pp. 3397-3415.

[7] Seokbeop Kwon, Byonghyo Shim, and Jian Wang, "Generalized orthogonal matching pursuit," School of Information and Communication, Seoul, Korea, 2011.

[8] Honglin Wu and Shu Wang, "Adaptive Sparsity Matching pursuit Algorithm for Sparse Reconstruciton," in *IEEE Signal Processing*, 2012, pp. 471-474.