IMAGE REPRESENTATION USING NON-CANONICAL DISCRETE MULTIWINDOW GABOR FRAMES

Nagesh K. Subbanna and Yehoshua Y. Zeevi
nagesh@tx.technion.ac.il zeevi@ee.technion.ac.il

Department of Biomedical Engineering,
Technion - Israel Institute of Technology,
Haifa, Israel.

ABSTRACT

We extend non-canonical multiwindow Gabor representations to quincunx sampling and discuss the advantages of representation of textures using non-canonical discrete multiwindow Gabor expansions, with the time-frequency plane sampled in a quincunx fashion. We show that the quincunx sampling is more efficient in terms of placement of the tiling zones in the time-frequency and achieves better time-frequency localisation when compared to the usual rectangular sampling. To emphasise the advantage of quincunx sampling, we compare images reconstructed using multiwindow Gabor frames using the canonical dual, a non-canonical dual, and finally a non-canonical dual under a quincunx sampling lattice and discuss our results. We show that the use of non-canonical duals permits a great deal of flexibility in the choice of analysis functions (and consequent generation of Gabor coefficients) and that duals can be optimised to the situation at hand.

1. INTRODUCTION

Gabor expansions are a popular choice of representation of signals in the time-frequency space and have been used in several practical applications. Gabor representations, with the Gaussian window function, achieve optimal localisation of signals in the time-frequency plane and this criterion is of importance in texture segmentation [8] where the authors have attempted to utilise the localisation of frequencies in the combined space to segregate textures. In [6], it was established that single windows are insufficient for proper localisation of frequencies and therefore, we choose multiwindow Gabor expansions as suggested in [15].

The multiwindow Gabor expansion of a rectangularly sampled, discrete time, finite energy signal \( f[k] \in \mathcal{C}^L \), of length \( L \) signal, is given by

\[
f[k] = \sum_{r=0}^{R-1} \sum_{m=0}^{b-1} \sum_{n=0}^{a-1} c_{r,m,n} \gamma_{r,m,n}[k]
\]  

where \( \gamma_{r,m,n}[k] = \gamma_r[k-na]e^{j2\pi mbk/L} \) is the \( r \)-th window function of the expansion frame, \( R \) is the number of windows used, \( a \) and \( b \) are the shifts along the time and the frequency axes, \( \bar{a} \) and \( \bar{b} \) are the number of shifts along the time and frequency axes respectively and \( c_{r,m,n} \) are the rectangularly sampled multiwindow Gabor coefficients given by

\[
c_{r,m,n} = \sum_{k=0}^{L-1} f[k]g_{r,m,n}^\ast[k],
\]

where \( g_{r,m,n} = g_r[k-na]e^{j2\pi mbk/L} \) are the analysis window functions. It is obvious that given the set of expansion vectors (set of frame vectors) \( \gamma_{r,m,n}[k] \), the task is to find the coefficients \( c_{r,m,n} \), and consequently, the analysis window functions \( g_{r,m,n}[k] \), so that good localisation of frequencies is achieved by the set of coefficients.

It has been established that both undersampling and critical sampling are generally unsatisfactory [16], [2] as the former is an incomplete representation and the latter suffers from lack of stability in reconstruction (shown by the Balian Low theorem). Further, the Gabor vectors are not orthogonal to each other, which necessitates the computation of duals to be able to reconstruct the signals. As oversampling is the only remaining alternative in Gabor expansions, we choose oversampling in this paper. Oversampling implies that there are more vectors than necessary and these vectors are linearly dependent, and therefore the dual is not unique. Any right inverse of the set of frame vectors constitutes a dual.

The traditional Gabor function based image reconstruction techniques fall into one of the two categories - the first technique is to use nearly orthogonal vectors so that reconstruction of the image can be made directly from the coefficients [5], [7]. The second is to utilise canonical duals [11] to find the biorthogonal functions to reconstruct the signal. The first technique suffers from a sharply restricted choice of analysis functions, since the near orthogonality implies that the ‘areas of significance’ have to be so placed as to have the least overlap [8], [5]. The limitation of analysis
functions is evident in certain examples in [7]. The latter choice of using the canonical dual [11], [9] which involves computation of the pseudoinverse of the set of expansion vectors, has two limitations. First, given an arbitrary choice of window functions, the computational difficulties of calculating the canonical may present several difficulties [12]. Second, the canonical dual provides the least square solution of the problem, which may not be desirable in cases of image reconstruction [1]. The human eye does not seem to work on the least square distance [7] and texture segmentation based on the square distance may not be accurate. Therefore, in order to alleviate these problems, we propose to use non-canonical duals to localise frequencies accurately. Further, we propose to utilise the denser and more uniform tiling of the quincunx lattice to alleviate the problems associated with the rectangular lattice.

The paper is organised as follows: In section 2, we discuss briefly non-canonical duals for rectangular sampling. In section 3, we extend the non-canonical duals to quincunx sampling and suggest techniques to obtain the dual easily. In section 4, we pick two natural images and two textures and reconstruct them using the most prominent coefficients obtained in the three methods and discuss them.

2. NON CANONICAL DUALS

Given a signal \( f[k] \in \mathbb{C}^L \), the Gabor expansion of the signal can be written in vector form as

\[ f = \Gamma c, \]  

where \( f \) is the signal \( f[k] \) of length \( L \), \( c \) is the vector of coefficients \( c_{r,m,n} \) of length \( Ra^0 \) and \( \Gamma \) is the set of expansion (or reconstruction) vectors given by

\[ \Gamma = \begin{bmatrix} \gamma_{0,0,0}[0] & \cdots & \gamma_{R_{,},,\bar{b}}[0] \\ \gamma_{0,0,1}[1] & \cdots & \gamma_{R_{,},,\bar{b}}[1] \\ \vdots & \ddots & \vdots \\ \gamma_{0,0,}\bar{L}-1 & \cdots & \gamma_{R_{,},,\bar{b}}[\bar{L}-1] \end{bmatrix}. \]  

Given the expansion frame, we need to compute the coefficients. The traditional way of generating coefficients is to use the canonical dual, which involves computing the pseudoinverse of the matrix \( \Gamma \), i.e.,

\[ c = (f, (\Gamma)\dagger) \]  

where \( (\Gamma)\dagger \) is the pseudoinverse of \( \Gamma \). However, the pseudoinverse has some disadvantages as stated in the previous section. Recognising that \( \Gamma \) is a rectangular matrix and has infinite possible right inverses, we introduce a dual of the form [4],

\[ \Gamma^*(\Gamma^*)^{-1}, \]  

where \( \Gamma^* \) is another matrix of the same form and dimensions as \( \Gamma \) but with a different set of window functions \( g_r \). The necessary and sufficient condition for the existence of duals of form (6) is that the matrix \( \Gamma^* \) be invertible. The conditions for the existence of the inverse have been investigated in detail in [13].

Briefly, we may state that the matrix \( \Gamma^* \) is a block circulant matrix and the inversion of block circulant matrices can be handled in several ways [3]. There are several ways of inverting the matrix \( \Gamma^* \) - primarily by using block discrete fourier transforms [13] and using Zak transforms [15]. However, the kind of oversampling (integer or rational) determines the techniques that can be used.

It was proved in [13] that, under the conditions of integer oversampling, if the set of window functions \( \gamma_r \) are positive definite\(^1\) sequences and \( g_r \) are of all of the same sign\(^2\), then the matrix \( \Gamma^* \) will always be invertible. Since the Fourier transforms of Gaussian window functions are positive definite and Gaussian window functions give optimum resolution in the combined space and are used most often in practice, we shall choose both \( \gamma_r \) and \( g_r \) to be Gaussian functions, and the time-frequency space oversampled by an integer factor.

The inversion of the matrix \( \Gamma^* \) can be accomplished in two ways in the case of integer oversampling. We have a simple technique [13], where the inversion of the matrix can be solved by the computation of Fourier transforms and inversion of a diagonal matrix. Another method proposed in [10] would have similar computational complexity. De-

\[^1\] Positive definite sequences are those sequences whose DFT is real and positive at all points.

\[^2\] Windows are positive or negative throughout the sequence.

Fig. 1. An illustration of the non-zero elements in the matrix \( \Gamma^* \) and the block circulant structure of the matrix. In integer oversampling, the block circulant matrices end up being diagonal or zero blocks.
pending on the circumstances, either may be better for the actual case. Similar results can be achieved using Zak transforms [15], which involves summation of a set of diagonal matrices and inverting them.

For the case of rational oversampling, we can still ensure that the complexity of inversion of the matrix is limited using Zak transforms [15], [13] or by using a combination of perfect shuffle matrices and block discrete Fourier transforms [13].

3. QUINCUNX SAMPLING

Quincunx sampling has been a popular alternative to the rectangular sampling for a number of reasons [14]. Briefly, two good reasons for using the quincunx sampling would be more uniform tiling of the time-frequency space and the denser packing, as can be seen in Fig. 2. The Gaussian function has a circular/elliptical tiling on the combined space, and it is obvious from the shape of the tiling that having a hexagonal (quincunx) shaped lattice would leave fewer ‘gaps’ in combined space than the rectangular sampling with similar rate of oversampling, as can be seen in Fig. 2. This has the added advantage of leading to a tighter frame [14].

3.1. Gabor vectors under Quincunx sampling

In rectangular sampling, the time and frequency shifts are independent of each other. In the quincunx sampling, we no longer have the independence of the time and frequency shift operators. In odd rows, we have the shift increased by \( a \) for every shift used and the actual shift is \( 2a \) between two adjacent time-shifted Gabor vectors. However, there is a simple way to reduce the quincunx sampling to a pattern similar to rectangular matrix. A close observation of the patterns shown in 3 indicates that the ‘rows’ of the quincunx matrix differs from the rows of the rectangular sampling only by a modulation factor. This can be achieved by modulating the rectangular sampling translation matrix \( T \) by a modulation matrix \( Q \) where \( Q \) is given by

\[
Q = \begin{bmatrix}
1 & 0 & \cdots & 0 \\
0 & e^{j2\pi b/L} & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & e^{j2\pi b/L}
\end{bmatrix},
\]

and

\[
T = \begin{bmatrix}
0 \\
1 \\
\vdots \\
L - 1
\end{bmatrix}.
\]

Given the easy conversion from the quincunx sampling to the rectangular sampling, we can use the non-canonical dual, only both the matrices \( G \) and \( \Gamma \) will be accompanied by the modulation matrix \( M \). Apart from this difference, the two matrices can be manipulated in the same way the rectangular sampling matrices were handled and inverted using the same techniques. A full discussion of the quincunx lattice and methods of handling the resultant matrices can be found in [14]. Briefly, we can say that both BDFT techniques and Zak transform techniques can be applied without any loss of generalisation to the quincunx lattice. The other way to handle the quincunx lattice is to reduce it into a rectangular lattice as shown in Fig. 3. The technique has been elaborated upon in [9].

4. SIMULATIONS AND DISCUSSIONS

We choose images of two types - natural images and texture images. In the former category, we choose two images, viz, Lena, and Nails. In the latter category, we choose two
images as well, viz, Honeycomb and Stripe. We desire to
reconstruct each of them using the least number of significant coefficients, with the expansion vectors \( \gamma_r \) already being set and the oversampling rate fixed. We define the rate of oversampling as the number of samples for reconstruction divided by the length of the signal, i.e., \( \frac{L}{\hat{L}} \) and in our case, the rate of oversampling is 4. In our work, we choose the absolute value of the coefficients \( |c_{r,m,n}| \) as the measure of the significance of the coefficients. We choose 10\% of the most significant coefficients as the ones necessary for construction. An alternate possibility would be to use the 'energy' of the coefficients as the measure for the choice of the significant coefficients [5]. In this instance, we have chosen four windows (all four windows are Gaussians, whose spreads\(^3\) are in arithmetic progression) for the representation of the image using the multiwindow Gabor coefficients. Here, we are aiming not at perfect reconstruction, but to measure the localisation of frequencies using non-canonical duals, with different sampling.

As may be seen in Figs. 5 and 6, the canonical dual tends to 'smear' the coefficients across frequencies and compromise the exact localisation of frequencies. The indicated areas in Fig. 5 show how the lines in the hat (region A), and the reflection in the mirror (region C), are both blurred while the exact texture in region B is lost in the canonical dual, and is less distinct in the non-canonical with rectangular sampling case. The observed results are in consonance with the observations of Chen et. al. [1] as shown in Fig. 4. Using non-canonical duals with rectangular sampling, we obtain better frequency localisation, but the best results are seen in the case of non-canonical duals with quincunx sampling. The result is fairly intuitive and bears out the observation that the uniform placement of areas of uniform spread and the better 'packing' by the quincunx sampling achieves better frequency localisation.

In the examples, we have limited ourselves to 10\% of the coefficients. When we choose 50\% of the coefficients, the reconstruction in both canonical and non-canonical techniques is nearly perfect. Further, for certain cases of windows, both non-canonical and canonical windows give good results. However, it is when the windows are arbitrary and the frequencies in the images diverse that the non-canonical solution really shines.

However, in the case of the Fig. 7, it is obvious that there is not too much change between the three outputs, especially in the honeycomb pattern itself. The lack of difference between the outputs is even more pronounced in the stripe image in Fig. 8. It must be remarked that the honeycomb and the stripe images have very little in the way of frequency variations - both of them are regular textures. The present scheme seems to indicate that the technique of using non-canonical duals is particularly well suited to a cadre of images where the expansion frame is fixed and there is a rather large variation in the set of frequencies in the image. The non-canonical dual performs better since it is better able to 'adapt' to the frequencies, by choosing duals other than the one which minimises the least square distance. It is of importance to mention that the choice of the norm is left to the user. We used the \( L_2 \) norm here, but any conceivable norm could be applied to the set of non-canonical duals. It affords the user a far greater freedom without inhibiting the ease of computational complexity to any great degree. The flexibility of the non-canonical duals is one of the most important features of the non-canonical Gabor scheme.

The adaptability of the non-canonical dual is enhanced by the choice of quincunx sampling, which improves upon the coefficients (both in terms of the uniformity and in terms of the greater density of packing of the tiling areas) when compared to the rectangular sampling. The optimal placement of the tiling areas in the case of the quincunx sampling alleviates the problems of frequencies falling in the 'blind spots' of the rectangular sampling.

4.1. Computational Complexity

The computational complexity of the localisation of frequencies using non-canonical duals is made up of two parts. The first part is the creation and inversion of the matrix \( \Gamma G^* \). The latter part consists of choosing the best coeffi-

---

\(^3\)The spread is the standard deviation of the Gaussian function.
Fig. 5. The original Lena image (top left) and the image reconstructed using 10% of the largest coefficients using the canonical dual (top right), the non-canonical dual with rectangular sampling (bottom left) and using the non-canonical dual with quincunx sampling (bottom right).

As discussed in [13], the creation of the matrix $\Gamma G^*$ is, in the best case, $2Rab$ multiplications and its inversion is of order $O(2L\log(b))$. The part about reconstruction is just choosing the set of coefficients to reconstruct from and this can be achieved by a good sorting algorithm. Reconstruction of the signal consists of a multiplication of a matrix of size $Rab \times L$ with a vector of length $Rab$.

5. CONCLUSION

From the examples, we can see that choosing a non-canonical dual can lead to more efficient representation in the time-frequency plane and helps in better localisation of frequencies. It is also apparent that non-canonical representation of signals can lead to better texture classification, with the choice of the metric remaining with the user, adding to the flexibility of the system. Other possible applications include speech processing and DNA sequence analysis. It is perhaps of interest to mention that the technique has already been used with considerable success in the latter application.

Fig. 6. The nails image from Brodatz album (top left) and the image reconstructed using 10% of the largest coefficients using the canonical dual (top right), the non-canonical dual with rectangular sampling (bottom left) and using the non-canonical dual with quincunx sampling (bottom right).

Fig. 7. The original honeycomb image (top left) and the image reconstructed using 10% of the largest coefficients using the canonical dual (top right), the non-canonical dual with rectangular sampling (bottom left) and using the non-canonical dual with quincunx sampling (bottom right).
Fig. 8. The original Stripe image (top left) and the image reconstructed using 10% of the largest coefficients using the canonical dual (top right), the non-canonical dual with rectangular sampling (bottom left) and using the non-canonical dual with quincunx sampling (bottom right)

6. ACKNOWLEDGEMENT

The authors are grateful to Prof. Y. V. Venkatesh for his suggestions. This research program has been supported in part by the Ollendorf Minerva Research Center, by the HASSIP Research Network Program HPRN-CT-2002-00285, sponsored by the European Commission, and by the Fund for Promotion of Research at the Technion.

7. REFERENCES