
Name of the Student: B. RAJATHILAGAM

Register Number: 0130404111

Year of Registration: 2005

Name of the Department: Centre for Excellence in Computational Engineering and Networking

Thesis Advisor: Dr. K. P. Soman
    Head, Center for Excellence in Computational Engineering and Networking
    Amrita School of Engineering, Ettimadai
    Amrita Vishwa Vidyapeetham
    Coimbatore – 641 112

Name of the School: Amrita School of Engineering, Ettimadai

Name and Address of the University: Amrita Vishwa Vidyapeetham,
    Amritanagar Post
    Ettimadai
    Coimbatore – 641 112
Synopsis of the Ph.D. Thesis

G-LETS: A NEW SIGNAL PROCESSING ALGORITHM

By

B. RAJATHILAGAM

Department of Information Technology

Amrita School of Engineering

Coimbatore – 641 105, Tamil Nadu, India

September, 2012
1 Introduction

New signal processing algorithms such as wavelets (Mallat S., 1989), multiwavelets (Chui, 1996), geometric wavelets (Chen, 2010), Gabor wavelets (Akan, 1997), ridgelets (Candes, 1999), and image representations are being developed to address the challenging aspects of signal and image processing such as local feature extraction, edge detection, and face recognition. All the above algorithms approximate the basis of the signal space, project the signal onto the approximate basis, and obtain its features. The performance of these algorithms depend critically on the approximation involved in determining the basis of the signal space. Further, when used for applications such as edge detection, additional approximations arise, ex. in the identification of the so-called edgels. In this work, a new algorithm for signal processing is presented, where the discrete signal is taken to be a member of a finite-dimensional group comprising a set of transformations, and the dimension of the signal space is obtained directly from group theory (Dreselhaus, 2008). Such an approach is seen to alleviate the approximations discussed above, and holds promise for better performance for the various applications of signal and image processing.

2 Scope of the Work

This thesis presents a new group-theory-based algorithm known as G-lets (Rajathilagam et al., 2012a; 2012b) for processing of discrete signals and images. From a group of transformations and representation theory, a finite basis of the signal space is obtained. The projections of the signal onto this basis are called G-lets. Amplitude resolution analysis and frequency analysis of signals and images have been performed. Features of signals have been localized and extracted successfully. Image compression has been performed. Edge detection of images has been performed using G-lets. The algorithm shows promise in application to other image processing problems such as face recognition, gesture recognition, pattern recognition, and 3-D reconstruction. Chapter 1 of the thesis provides a broad introduction to the area of signal processing and highlights the key contributions from the present work.

3 Group Theory and Signal Processing

A digital signal is finite-dimensional. Therefore, within the approximation of digitization, the signal exists in a vector space of finite basis. If it is possible to obtain this basis of the signal space, a signal processing algorithm may be developed wherein the signal is projected onto the basis of the signal space to obtain its individual pieces. Such an algorithm would provide a general framework for extracting various features of interest from the signal without loss of information, within the digitization approximation. This is the inspiration behind the development of G-lets. Group theory provides the crucial piece of the puzzle, namely the dimension of the vector space where the signal exists. Group theory (Lenz, 1990; Lenz et al., 2009; Vale, 2008; Foote et al., 2000) also shows the way for choosing any suitable transformation or set of transformations required for a particular application customized to capture specific characteristics. There are other ways in which group theory has been applied in the context of signal processing. They include generating Fourier basis of a signal space, performing convolution with irreducible representations, and generating projection matrices. These approaches are reviewed and the context of the present work is discussed in Chapter 2. In addition, Chapter 2 provides a primer on the group theory.
4 G-lets: Signal Processing using Transformation Groups

Chapter 3 presents the algorithm developed in this work, namely G-lets. It demonstrates how to construct G-lets and shows how an amplitude multiresolution analysis may be performed on 1-D signals and images using G-lets. Finally, Chapter 6 presents designing of amplitude filter banks using G-lets. All results are illustrated using a simple group comprising two types of transforms, called the dihedral group. Similar analysis may be performed with other finite-dimensional groups (Assefa et al., 2010). The dihedral transformation group contains rotations and reflections. The number of irreducible representations of a transformation group is equal to the number of conjugacy classes by Schur’s Lemma (Dresselhaus, 2008). For dihedral groups $G$ shown below, if ‘n’ is the dimension of the signal space, then the number of irreducible representations is given by $(n + 6)/2$ for even sized ‘n’ and $(n + 3)/2$ for odd sized ‘n’. Hence this number i.e., the number of conjugacy classes, becomes the actual dimension of the signal space.

$$G = \{R, R^2, R^3, \cdots R^n, S, SR, SR^2, SR^3, \cdots SR^{n-1}\}$$

An $i^{th}$ G-let matrix looks like

$$G - let_i = \begin{pmatrix} IR_i & 0 \\ IR_i & IR_i \\ 0 & IR_i \\ \vdots & \vdots \end{pmatrix}$$

where $IR_i$ is the $i^{th}$ irreducible representation. G-let matrices may also be constructed by a combination of $IR_1$, $IR_2$, etc within a single matrix. More examples of G-let matrices can be seen in the reference (Rajathilagam et al., 2012a). The oscillations of G-lets produces an amplitude resolution for the chosen signal. The ‘Lena’ image and its amplitude resolution images are shown in the Fig. 1. The ‘Lena’ image is of size 256 x 256.

**ALGORITHM G-let Decomposition** (sig, D)

1. Calculate the dimension ‘n’ of the signal ‘sig’
2. Choose a transformation group, for example the Dihedral group D
3. Generate basis matrices for G-let G(sig) from 1 to $2n$ using ‘Generate_G-let_Matrices’ function given below. The $i^{th}$ G-let is taken as $G_i(sig) = R_i$
4. Project the signal onto the G-let matrices to oscillate the signal
5. Each G-let is a different resolution in terms of amplitude produced by the corresponding oscillations
6. Generate G-let frequencies using a dilation operation
7. For each amplitude resolution, obtain low and high frequencies using dilation that separates oscillatory portion
8. The signal is now a linear sum of the G-let coefficients

$$\text{sig} = (G_1(sig) \oplus G_2(sig) \oplus G_3(sig) \cdots \oplus G_n - 1(sig)) \times (-I_n)$$

$$\text{sig} = (G_{n+1}(sig) \oplus G_{n+2}(sig) \oplus G_{n+3}(sig) \cdots \oplus G_{n-1}(sig)) \times (-I_n)$$

These are the first and second basis sets corresponding to rotations and reflections respectively

**ALGORITHM Generate_G-let_Matrices**()

1. $angle = 360^\circ/n$ taken in radians
2. Calculate number and dimension of irreducible representations
3. Identify the conjugacy class of matrices
4. Make the transformation matrices $R$ using, one and two dimensional irreducible representations placing them diagonally
5. For rotations and reflections, in the case of Dihedral groups
6. return $R$
5 Frequency Analysis of Signals and Images using G-lets

Frequency analysis and extraction of spatially localized features in signals and images are discussed in Chapter 4. Developing frequency filter banks is demonstrated in Chapter 6. A linear transformation matrix $T(R)$ of a vector space $V_n$ has its diagonal blocks as irreducible representations $IR_1, IR_2, \ldots IR_{(n+z)/2}$, and produces oscillations $O(s)$ of width equal to the dimension of the corresponding block, $\mu_r$, when transforming a $n$-tuple signal. In a dihedral transformation matrix $D(R)$, the oscillations $O(s)$ happen at alternate values due to the nontrivial two dimensional irreducible representations. In this work (Rajathilagam et al., 2012c), frequency of a discrete signal at a position is considered to be proportional to the difference between the amplitudes of consecutive positions. The difference $G_d(x)$ between consecutive high $G_h(x)$ and low $G_l(x)$ transform coefficients is proportional to the difference $S_d$ in the amplitude between consecutive points of the signal in the same position $x$ (indicative of the frequency at $x$).

$$G_d(x) \propto S_d(x) \tag{5.1}$$

Depending on the choice of the transformation $T(R)$, maximum frequency $f_{max}$ in a signal results in either maximum $G_{d_{max}}$ or minimum $G_{d_{min}}$ difference between the consecutive high and low transform coefficients $G(x)$. A dilation operation $L(x)$ may be defined to further transform the signal so as to highlight the highest frequency $f_{max}$ of the signal, while suppressing the lower frequencies $f_{min}$.

$$f_{max}(or f_{min}) = L(x)T(R)S \tag{5.2}$$

After dilation, at a specific amplitude resolution $A(t)$ (ex. specific transformation angles for a dihedral group), the maximum frequency $f_{max}$ stands out more clearly than at other resolutions. The result after dilation is given from $T(x)$ and $L_m(x)$ by

Figure 1: (a) ‘Lena’ Image ; Lena Rotation G-lets: (b) 10th G-let (c) 25th G-let (d) 89th G-let (e) 96th G-let (f) 162nd G-let
\[ DG(x) = G(x)L_m(x) \]  

Since the \( j^{th} \) coefficient is the highest, only that coefficient stands out while the rest are suppressed. Thus, a dilation operation has been constructed to highlight the maximum frequency of the signal. This is depicted in Fig. 2(a) and the results are shown in Fig. 2(b), Fig. 2(c), Fig. 2(d) and Fig. 3. Also from the resultant equation, it is seen that the dilation operation is localized. Hence the G-let transform has the ability to retain the position of each frequency in the G-let unlike a Fourier transform (Bayram et al., 1996; Hess-Nielsen et al., 1996; Stankovic, 2010; Neill, 1999; Hlawatsch et al., 1992).

\[
DG(x) = G_1(x) \left| \frac{G_{j-1}}{G_{max,j}} \right| \oplus G_2(x) \left| \frac{G_{j-1}}{G_{max,j}} \right| \oplus \cdots \oplus G_m(x) \left| \frac{G_{j-1}}{G_{max,j}} \right| \oplus \cdots
\]  

(5.4)

Figure 2: (a) Frequency Analysis of G-lets; ECG Signal: (b) Signal (c) G-let Frequencies (d) Low Resolution form of the ECG Signal

6 Local Feature Extraction using G-lets

Using G-let frequencies, spatially local features of signals and images may be extracted. For an image, every frequency is calculated with its horizontal and vertical neighborhood. First the highest frequencies are noted as the potential places of local features. For every high frequency, initially the horizontal neighborhood is calculated row wise. Then its vertical neighborhood is found out column wise. For each frequency belonging to the vertical neighborhood, the horizontal neighborhood is followed. All these frequencies put together make a local feature around
the chosen high frequency marking an irregular outline as the local feature. The boundaries of the local neighborhood can also be cutoff to have a specific length on all sides, so that there is regular shape for the borderline, for example, a box. Thus every local feature is spatially localized with respect to a base high frequency. Fig.4 shows the features the eyes and nose within a boxed limit and another local feature, the cap, as frequencies.

7 Edge Detection in Images using G-lets

Detecting edges in an image is a fundamental problem in image processing, critical to further applications in computer vision. The use of G-lets to detect edges in images is presented in Chapter 5. The amplitude resolution of G-lets improves the identification of edges by deepening the intensity variations. First G-lets are obtained by applying G-let matrices to an image. Then the sum of left and right G-let coefficients are produced. The edges are calculated as gradient magnitudes of the resultant image. The sum of left and right G-let coefficients actually amplifies the brightness of the edges by a factor of $2\cos \theta$ for dihedral groups and improves the contrast between the depth and inner texture edges proportionally. The G-let matrices when applied on the resultant image column wise, highlight vertical edges in the image. G-let matrices when applied on the resultant image row wise, highlight horizontal edges in the image. Their individual thresholded complements are calculated and thus four types of edges put together sketch all the edges of an image - vertical edges, horizontal edges and their respective complements.

Depth and illumination edges are darker than texture edges unlike edgels (Canny, 1986; Raskar, 2004; Ahmad, 1999). Due to this inner texture edges can be removed easily by thresholding.
Therefore the shapes of objects can also be recognized. Object boundaries or depth edges in an image correspond to high frequency areas which are not affected by the oscillations of G-lets suppressing low frequencies. Edges of textures appear brighter only when low frequencies are slowly suppressed creating better intensity variations in the texture regions. For the same reason, straight lines, junctions and corners are preserved by G-let edges. The steps involved in edge detection is shown in the block diagram of Fig. 5.

![Diagram of edge detection using G-lets](image)

Figure 5: Edge Detection using G-lets

The results of edge detection are shown from Fig. 6 to Fig. 9. In Fig. 6 it is seen that the edges appear according to their depth in the scene with respect to the angle of the camera. When more G-lets are used deeper edges appear in the scene. In Fig. 7 there are straight lines and circles in

![Edge Detection of 'Flower' Image](image)

Figure 6: Edge Detection of ‘Flower’ Image: (a) Original Image (b) Edges with 2 G-lets (c) Edges with 4 G-lets (d) Canny Edges
the original image but there are gaps at the points crossing circles and other lines in the Canny edges. G-let edges are continuous and hence the junctions of lines are not broken. Without the use of a Hough transform G-let edges are directly able to identify the corners in the image. The depth edges of G-lets are naturally sharper than texture edges, demonstrated by the clarity of G-let edges which are better than the edgels of Canny edges shown in Fig. 8. In Fig. 9, the ‘Mandril’ image, looking closely at the area near the nostrils of Canny edges, there are some unwanted texture edges due to the higher illumination in that area of the original image. The G-let edges of the same image clearly diminishes that portion without affecting the actual edges in the area and does not show extra edges in other similar areas.

Figure 7: Edge Detection of ‘LineCircle’ Image: (a) Original Image (b) Edges with 3 G-lets (c) Canny Edges

Figure 8: Edge Detection of ‘Butterfly’ Image: (a) Original Image (b) Edges with 3 G-lets (c) Canny Edges

Figure 9: Edge Detection of ‘Mandril’ Image: (a) Original Image (b) Edges with 2 G-lets (c) Canny Edges
8 Face Detection using G-let Edges

The advantage of G-let edges in suppressing texture edges and highlighting depth edges can be utilized in detecting faces within an image. The features of a face - the eyes, nose and mouth are not identified as textures by G-lets, they are outlined with respective depth edges. Due to this discrimination the shape of the face is sketched neatly by G-let edges. This can be verified in the ‘Face Array’ image in Fig. 10, where the Canny edges show additional texture edges that make unclear the identity of the objects in the image. It is observed that the light falling on the faces from a specific angle in the scene is presented as actual edges in the texture area by Canny edge detector. In the corresponding G-let image, the texture areas in the faces have no prominent edges. Instead smooth low intensity pixels cover the light exposed areas which can be removed by suitable thresholding. The same effect can be seen in another image showing a group of faces in the ‘Children’ image of Fig. 11. The G-let edges are able to distinguish eyes, nose and mouth of each face in the image with just one G-let. Compared to the G-let features, Canny edges show less distinct features.

![Image of Face Detection](image.png)

**Figure 10:** Face Detection of ‘Face Array’ Image: (a) Original Image (b) Faces with 2 G-lets (c) Canny Edges

![Image of Face Detection](image.png)

**Figure 11:** Face Detection of ‘Children’ Image: (a) Original Image (b) Faces with 1 G-let (c) Canny Edges

9 Summary and Conclusions

Some of the salient features of this algorithm are discussed below:

- With every G-let, the signal is portrayed by representation matrix from different perspectives. The irreducible pieces might be big or small and that depends on the size of the signal. When a signal shows an irreducible representation of a large block size, it is possible to create an equivalent representation through a linear transformation with the help
of its subspaces (null space or range space) such that the block becomes further reducible. This technique allows to focus on and explore further that part of the signal captured by the irreducible block as shown in (Rajathilagam et al., 2012c). The features of the signal under consideration are found to be spread across conjugacy classes, i.e., here they are spread across two conjugacy classes.

- The number of irreducible representations is equal to the number of conjugacy classes. Irreducible representations form the orthonormal basis. Conjugacy classes represent the structural symmetries for the signal. So, whenever a change occurs in the signal at a particular point, the above relationship propagates this change across different conjugacy classes.

- When a representation is chosen for a signal, the different basis of the underlying vector space gives different representations as if viewing through a kaleidoscope. Thus a powerful method for matching the features of two signals especially in applications like pattern matching and face recognition with a variety of basis sets is obtained.

- Since multiple G-lets are involved, the algorithm is ideally suited for a multi-processing system.

- Expanding any function \( f(x) \) in terms of the basis functions of irreducible representations requires that the representation group is obtained by applying to the function \( f(x) \) all the transformations. Then the function is linearly expressible in terms of the basis functions in the various irreducible representations. One may also choose one of the irreducible representation functions \( g(x) \) with its projection of the signal and generate a new set of basis functions which will include itself in the set of basis functions. This feature allows us to customize the representation for any signal.

- As the size of the signal increases, the number of conjugacy classes is almost 50% less. This promises a lossless compression of the signal, though further attention is needed in this regard.

- In determining irreducible representations, a set of reducible representations may be divided into lower angle matrices and higher angle matrices. The lower angle matrices may then be explored using a completely different transformation group to further extract more features from that part of the signal.

- Group representation theory states that, new representations can be generated by direct product or tensor product of any two representation groups. Then their irreducible representations also turn out to be a direct product or tensor product.

- It is also possible to use complex representation matrices so long as a suitable transformation group is chosen. This quality could lead to the discovery of a new set of features in a signal.

- Local features: The local features of a signal is present around the frequencies of a G-let. For an image, every frequency is calculated with its horizontal and vertical neighborhood. For each frequency belonging to the vertical neighborhood, the horizontal neighborhood is followed. All these frequencies put together make a local feature around the chosen high frequency marking an irregular outline as the local feature. The boundaries of the local neighborhood can also be cutoff to have a specific length on all sides, so that there is a regular shape for the borderline as shown in (Rajathilagam et al., 2012c).
• The number of G-lets required for sketching the edges of an image is very much less than though proportional to $n'$, the dimension of the signal.

• In the current study, only amplitude resolution of G-lets and gradients are used for calculating edges. G-let frequencies has not yet been explored to identify the edges.

• Since lines and corners are preserved by G-let edges, 3-D reconstruction of multiple images may be feasible.

• Texture edges are handled well by G-lets when compared to Canny edges which has many spurious inner texture edges.

• Edge thinning and linking need not be done using non-maximum suppression, instead low intensity edges that appear with more G-lets can be utilized.

• The G-let edges carry the distance of the edge and the low light areas of an object with respect to the camera or projector, see Fig. 6 for example.

• Complexity: In general, the computational complexity of performing frequency analysis is proportional to the dimension of the chosen signal $n'$. Since the G-let matrices are sparse matrices, the order of multiplication boils down to the sum of $\#$ of two dimensional irreducible representations $\times 4$ (the size of one two dimensional representation) and $\#$ of one dimensional irreducible representations. The computational complexity also depends on the number of viewpoints that need to be explored, the choice of transformation group and size of the signal. The matrices obtained in all of the representations are sparse orthonormal block-diagonal matrices and this adds an advantage for large signal calculations. One member of every conjugacy class put together shows a complete and different view of the signal. The same can be achieved by the alternate members of the conjugacy classes. In this way, the number of G-lets required for reconstruction of the signal can be cut down to the number of conjugacy classes.

• The transformations used in wavelets are scaling, dilation and translation. Since translations cannot be finite, scaling and dilation can form a transformation group. Wavelet frequencies are both temporally and spatially localized because they look for similar structures in a signal at a particular time. Within the rules of group theory, the same effect may be achieved with transformation groups. Geometric wavelets segregate fine and coarse details of the input signal using a plane similar to the choice of a transformation group and the corresponding dilation operation done in G-lets. The computational complexity involved in repeating the process for the sub-regions to create a dictionary of multi-scale analysis of data can be compensated by the use of the G-let matrices mapped to the transformation group. Chapter 7 compares G-lets with existing signal processing algorithms such as wavelets and Fourier analysis.

10 Scope for Future work

G-let basis is limited by the choice of the transformation group. Though dihedral groups have been used in the current implementation, which transformation group and its characteristics suit a particular application has not been examined. The G-let basis of a discrete signal is limited by the discretization method used with a suitable sampling technique. But within this discretization, the transformation group provides a basis very close to its natural basis. Chapter
of the thesis presents the conclusions and discusses in more detail the possible future work that can arise from this research.

Thus a different approach to signal processing with the choice of a group of transformations is proposed in this work. The results provide a method of multiresolution analysis in terms of amplitude and frequency separately without an approximation like a mother wavelet. G-lets can also be extended to a continuous signal in the Hilbert Space. This method may be useful for edge detection, face detection, face recognition, denoising, object recognition, 3-D reconstruction, shape detection, text identification in images, gesture recognition, target recognition and other applications in image processing and robotics.

References


11 **List of Publications (Based on Ph.D. research work)**


