Iterative Pyramidal Filtering Method for Improved Signal Recognition in Radio Spectrograms

Miguel López-Benítez, *Senior Member, IEEE*

***Abstract*—Spectrograms are an essential time-frequency representation tool that has been used to address several important problems in wireless communication systems. However, most existing techniques based on the processing of radio spectrograms require a relatively high Signal-to-Noise Ratio (SNR), performing poorly at low/moderate SNR. In this context, this paper proposes an iterative radio spectrogram filtering method based on a novel pyramidal convolution kernel. The obtained results demonstrate that the proposed technique improves the recognisability of signal components in radio spectrograms. Two illustrative examples are provided to show how this method helps to extend noticeably the SNR operational range of techniques for wireless communications based on the processing of radio spectrograms.**

***Index Terms*—Time-frequency signal representations, radio spectrogram, spectrum analysis, spectrum awareness.**

I. Introduction

RADIO spectrograms describe the temporal evolution of the power spectral density of wireless communication signals and provide essential information such as the received signal strength, carrier frequency, occupied bandwidth, spectral mask and transmission pattern. Spectrograms have been used to address various problems in wireless communications such as automatic blind modulation classification (based on heuristic algorithms [1]–[3] and convolutional neural networks [4]–[6]), radio technology identification [7], extraction of frequency hopping signal parameters [8], [9], spectrum sensing [10], detection of radar signals [11] and characterisation of the Signal-to-Noise Ratio (SNR) and Doppler shift [12].

Most techniques based on the processing of radio spectrograms implicitly assume that the signal components have a sufficiently high SNR such that they can be clearly differentiated from the background noise, determined by the receiver noise floor. However, their performance degrades significantly in the low SNR regime. To address this practical limitation, this work proposes a relatively simple iterative filtering method that enhances the recognisability of signal components in noisy radio spectrograms. The proposed method applies multiple times the same type of filter to the radio spectrogram using an increasing filter size in each iteration and then adds each filtered contribution with an adequate weighting factor. The overall effect of this process is an improved response and signal detectability at low SNR without sacrificing the signal detection accuracy at high SNR. In addition to classical filters such as the Gaussian and box blur filters, a new filter type with a pyramidal convolution kernel is proposed, which provides a better trade-off between accuracy and computational cost in practical implementations. The immediate benefit of the proposed method is an enhanced recognisability of signal components that extends the SNR range of operation of techniques based on the processing of radio spectrograms.

The author is with the Department of Electrical Engineering and Electronics, University of Liverpool, Liverpool, L69 3GJ, United Kingdom, and also with the ARIES Research Centre, Antonio de Nebrija University, 28040 Madrid, Spain (email: m.lopez-benitez@liverpool.ac.uk).

Final author version (March 2022).

II. System Model and Problem Formulation

A radio spectrogram can be mathematically represented as a matrix $\mathbf{P} \in \mathbb{R}^{M \times N}$ of $M \times N$ power values. The convention adopted in this work is that the vertical and horizontal axes of $\mathbf{P}$ are associated with time and frequency, respectively. Thus, each element of $\mathbf{P}$, denoted by $P[m,n] \in \mathbb{R}$ ($m = 1, \ldots, M$ and $n = 1, \ldots, N$), is the power level measured at the mth time instant and nth frequency point, the vertical size M is the number of time samples and the horizontal size N is the number of frequency bins over the target frequency span.

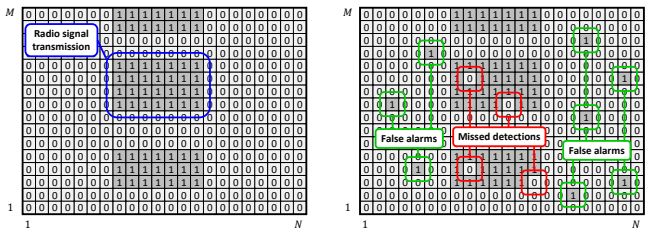
In order to decide on the presence of signal components in the spectrogram, an energy decision threshold $\lambda \in \mathbb{R}$ is employed, which can be calculated based on different methods [13]. This work assumes that λ is set based on the receiver noise floor [14] to achieve a constant false alarm rate [15]. The threshold is used to convert the matrix of continuous powers $\mathbf{P}$ at the receiver into a binary matrix $\mathbf{B} \in \mathbb{B}^{M \times N}$ (with $\mathbb{B} = \{0, 1\}$) where each element $B[m,n] \in \mathbb{B}$ is calculated as:

$$B[m,n] = \begin{cases} 0, & P[m,n] < \lambda \quad (1a) \\ 1, & P[m,n] \geq \lambda \quad (1b) \end{cases}$$

indicating where each spectrogram element is believed to contain a signal component ($B[m,n] = 1$) or not ($B[m,n] = 0$).

Let $\mathbf{T} \in \mathbb{B}^{M \times N}$ be a matrix whose elements $T[m,n]$ contain the true states of the elements $B[m,n]$ of matrix $\mathbf{B}$ at the receiver (see Fig. 1). Some of the elements of $\mathbf{B}$ will unavoidably be incorrect due to errors in the signal transmission and detection process, which can be characterised in terms of the false alarm probability $P_{fa} = P(B[m,n] = 1 | T[m,n] = 0)$ and detection probability $P_d = P(B[m,n] = 1 | T[m,n] = 1)$. $\mathbf{B}$ can be seen as a degraded version of $\mathbf{T}$ where random errors are introduced with probabilities P_{fa} and $1 - P_d$. In an ideal scenario of infinite (sufficiently high) SNR it is possible to set a threshold λ such that $P_{fa} = 0$ and $P_d = 1$ so that $\mathbf{B} = \mathbf{T}$. However, in many practical scenarios $\mathbf{B} \neq \mathbf{T}$ in general.

The target of this work is to develop a method to process a binary spectrogram $\mathbf{B}$ obtained by thresholding a continuous-power radio spectrogram $\mathbf{P}$ in order to provide an estimation as close to the (unknown) ground truth $\mathbf{T}$ as possible that can help identify radio transmissions within the spectrogram.



(a) Matrix $\mathbf{T}$ (b) Matrix $\mathbf{B}$

Fig. 1: Matrix model for radio spectrograms: (a) ground truth at the transmitter (matrix $\mathbf{T}$), (b) states observed at the receiver (matrix $\mathbf{B}$).

III. Proposed Method

A. Motivation and Overview

Notice from Fig. 1 that each radio signal transmission can be precissely characterised by a rectangular cluster of points that determines its occupied bandwidth and start/end times[1]. The proposed method exploits this clustered structure. The main underlying idea is that the density of points per unit area inside such rectangular cluster, referred to as Signal Area (SA), is in general greater than outside. For sufficiently high SNR, an SA density of one point per spectrogram element is obtained and $\mathbf{B} = \mathbf{T}$. However, the SA density will decrease as the SNR decreases since some points will be missing due to signal misdetections, which can occur with probability $1 - P_d$, where P_d decreases with SNR [15, eq. (1)] (the density outside SAs will remain constant when λ is set for a constant P_{fa}).

Since the gaps left by missing SA points will be surrounded by other SA points, it may still be possible to get a detectable response within those gaps by applying a filter to the degraded spectrogram $\mathbf{B}$ so that the response at each point depends not only on the point itself but its neighbourhood as well. The output $\mathbf{F}_r$ of filtering $\mathbf{B}$ with a filter $\mathbf{h}_r$ of radius[2] r is obtained from the two-dimensional convolution $\mathbf{F}_r = \mathbf{h}_r * \mathbf{B}$, where each element $F_r[m,n]$ of $\mathbf{F}_r$ is obtained as:

$$F_r[m,n] = \sum_{x=-r}^{r} \sum_{y=-r}^{r} h_r[x,y] \cdot B[m-x, n-y]. \quad (2)$$

As the SNR decreases the density of SA points will decrease and the filter radius should increase to improve the sensitivity. However, increasing the filter radius will also have a stronger blurring effect on the spectrogram data, which will make it more difficult to accurately identify the true boundaries of SAs. To address this sensitivity-specificity trade-off, the proposed method filters $\mathbf{B}$ several times using filters with different radii and weights each individual output $\mathbf{F}_r$ by the inverse of the employed filter radius. Thus, the outputs for smaller filter radii will have higher weights to allow an accurate identification of the SA edges, while the outputs for larger radii will have lower weights to avoid excessive blurring of the SA dimensions but will also be accounted for to help increase the chances of obtaining a detectable response within the SA gaps. The output obtained after K filtering iterations, denoted by $\mathbf{G}_K$, is:

$$\mathbf{G}_K = \sum_{r=1}^{K} \frac{\mathbf{F}_r}{r^p} = \sum_{r=1}^{K} \frac{\mathbf{h}_r * \mathbf{B}}{r^p}, \quad (3)$$

where parameter $p \in \mathbb{R}^+$ is introduced to allow a fine tuning of the weight carried by each contribution $\mathbf{F}_r$.

The final step of the proposed method is to binarise matrix $\mathbf{G}_K$ with a properly set threshold χ in order to produce a final binary output matrix $\mathbf{I}$ whose elements are calculated as:

$$I[m,n] = \begin{cases} 0, & G_K[m,n] < \chi \quad (4a) \\ 1, & G_K[m,n] \geq \chi. \quad (4b) \end{cases}$$

The final output $\mathbf{I}$ should be more similar to $\mathbf{T}$ than $\mathbf{B}$.

A complete definition of this method requires an appropriate selection of the filter $\mathbf{h}_r$ and the configuration parameters p, K and χ, which are discussed below.

B. Filter Selection

A common choice for data averaging and smoothing is the Gaussian filter, whose unnormalised coefficients are given by[3]:

$$\tilde{h}_r[x,y] = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right), \quad (5)$$

where $x, y \in \{-r, \ldots, 0, \ldots, r\}$ and the filter standard deviation σ is selected to ensure that a sufficiently representative portion of the Gaussian function is captured (a typical choice is $\sigma = r/3$, which embraces 99.73% of the Gaussian function). Since the shape of the Gaussian function changes with σ, the relative weights of the Gaussian filter coefficients will also change with the filter radius r. In particular, the Gaussian filter will approach a box blur filter with flat coefficients as the radius increases. This will unavoidably have an impact on the contribution of each component $\mathbf{F}_r$ in (3), which will be weighted explicitly by the factor $1/r^p$ and also implicitly by the filter shape itself. For a fixed value of p, this means that in practice the first $\mathbf{F}_r$ component ($r = 1$) will have a much higher total weight than the others ($r > 1$). For $r > 1$, the Gaussian filter shape will be wider and lower, which will lead to negligible $\mathbf{F}_r$ contributions when weighted by the factor $1/r^p$, thus removing the desired iterative filtering effect. This can be compensated by selecting $p < 1$, however, as it will be shown, in practice it is difficult to find a suitable fixed value of p that performs well over the whole SNR range.

This motivates the proposal of a novel filter characterised by a pyramidal convolution kernel given by (see the Appendix):

$$\tilde{h}_r[x,y] = \min(r+1+x,\ r+1-x,\ r+1+y,\ r+1-y), \quad (6)$$

[1]Certain radio transmissions may not lead to rectangular clusters, in particular electromagnetic emissions from systems that are usually not intended for wireless communications, including some types of radars [11, Figs. 1 and 9], microwave ovens [16, Fig. 2b], and several sources of man-made noise [17, Fig. 4]. While the method here proposed should also work in these cases, such specific signal formats require a tailored study that is beyond the scope of this work, whose focus is on wireless communication signals, usually characterised in radio spectrograms by rectangularly-shaped clusters.

[2]This work follows the common convention of defining the radius of the filter as the apothem of its square kernel. Thus, a filter with radius r will have a total side length of $2r+1$ and contain $(2r+1)^2$ filter coefficients.

[3]The notation $\tilde{\mathbf{h}}_r$ for a filter, or $\tilde{h}_r[x,y]$ for its coefficients, indicates an unnormalised filter. The normalised version employed in (3) is obtained as $\mathbf{h}_r = \tilde{\mathbf{h}}_r[\mathrm{tr}(\tilde{\mathbf{h}}_r\, \mathbf{J}_{2r+1})]^{-1}$, where $\mathbf{J}_l$ denotes the $l \times l$ all-ones matrix.

where $x, y \in \{-r, \ldots, 0, \ldots, r\}$. This filter preserves its shape (and hence the relative weights of the filter coefficients) for any value of the filter radius r. Moreover, this filter has square symmetry, which intuitively seems more appropriate to detect the square shape of signal components in radio spectrograms.

A box blur filter $\tilde{h}_r[x, y] = 1$, $\forall x, y \in \{-r, \ldots, 0, \ldots, r\}$ is also considered for comparison. The Gaussian and box blur filters are the standard choices for data averaging, which is the effect necessary for the proposed iterative filtering method, and are suitable for comparison to the proposed pyramidal filter.

C. Iterative Stopping Algorithm

A simple way to set K in (3) would be according to the maximum filter size that can be fitted within the spectrogram dimensions, i.e. $2K + 1 \leq \max(M, N)$. However, this approach would likely lead to a high computational workload. A smarter approach can be devised by noting that every new contribution $\mathbf{F}_r$ in (3) carries a smaller weight $1/r^p$ than the previous one and therefore there must be a certain number of iterations K beyond which any new contributions will be negligible. The number of iterations K can be automatically determined by comparing the successive outputs of (3) and stopping when their difference is below a certain threshold.

The similarity between two successive outputs of (3) can be quantified in terms of the Peak Signal-to-Noise Ratio (PSNR). The PSNR in dB at the kth iteration of (3) is given by [18]:

$$\mathrm{PSNR}_k\,[\mathrm{dB}] = 10\log_{10}\left(\frac{(G_k^{\max})^2}{\mathrm{MSE}_k}\right), \tag{7}$$

where $G_k^{\max}$ is the maximum possible element value of $\mathbf{G}_k$ and MSE_k is the mean squared error between $\mathbf{G}_k$ and $\mathbf{G}_{k-1}$.

Since every iteration of (3) adds a new contribution to the total sum, the maximum value of $\mathbf{G}_k$ will increase with k. $G_k^{\max}$ can be calculated from (2) and (3) as:

$$\begin{aligned} G_k^{\max} &= \max_{m,n}\Big(G_k[m,n]\Big) = \max_{m,n}\left(\sum_{r=1}^{k}\frac{F_r[m,n]}{r^p}\right) \\ &= \sum_{r=1}^{k}\frac{\max_{m,n}\Big(F_r[m,n]\Big)}{r^p} = \sum_{r=1}^{k}\frac{1}{r^p} = H_k^{(p)}. \end{aligned} \tag{8}$$

Recalling that $\mathbf{B}$ is a binary matrix, this maximum possible value is obtained for an element $F_r[m, n]$ where it holds that $B[m-x, n-y] = 1$ for all x, y within the filter template. The value of that element $F_r[m, n]$ then becomes, according to (2), the sum of the filter coefficients, which for a normalised filter is equal to one. According to (8), $G_k^{\max}$ is the kth generalised harmonic number of order p, typically denoted as $H_k^{(p)}$ [19].

MSE_k aims to quantify the variation of $\mathbf{G}_k$ with respect to the output $\mathbf{G}_{k-1}$ from the previous iteration as:

$$\begin{aligned} \mathrm{MSE}_k &= \frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}\left(\underbrace{\sum_{r=1}^{k}\frac{F_r[m,n]}{r^p}}_{G_k[m,n]} - \underbrace{\sum_{r=1}^{k-1}\frac{F_r[m,n]}{r^p}}_{G_{k-1}[m,n]}\right)^2 \\ &= \frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}\left(\frac{F_k[m,n]}{k^p}\right)^2 = \frac{\mathrm{tr}(\mathbf{F}_k^{\mathsf{T}}\mathbf{F}_k)}{MNk^{2p}}. \end{aligned} \tag{9}$$

Introducing (8) and (9) into (7) yields:

$$\mathrm{PSNR}_k\,[\mathrm{dB}] = 10\log_{10}\left(\frac{MNk^{2p}\big(H_k^{(p)}\big)^2}{\mathrm{tr}(\mathbf{F}_k^{\mathsf{T}}\mathbf{F}_k)}\right). \tag{10}$$

Based on (10), the iterative process in (3) proceeds until:

$$\Big|\,\mathrm{PSNR}_k\,[\mathrm{dB}] - \mathrm{PSNR}_{k-1}\,[\mathrm{dB}]\,\Big| < \varepsilon\,[\mathrm{dB}], \tag{11}$$

where ε is a properly set threshold in decibels.

The algorithm presented decides when to stop the iterative process in (3), i.e. the value of K. The value of p determines the convergence rate of this iterative process and therefore the required number of iterations K. An adequate choice of p will be discussed in Section IV based on simulation results.

D. Thresholding Method

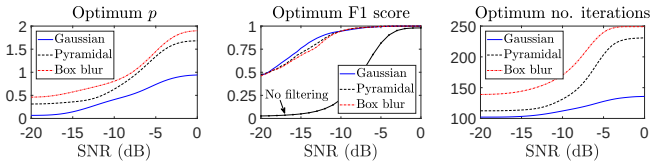
The final step in (4) requires an adequate selection of the threshold χ used to classify the elements of $\mathbf{G}_K$ as idle/busy points. To this end, the Otsu algorithm is selected [20].

IV. Simulation Results

The proposed method was evaluated using the same simulation approach and configuration parameters as [15]. Based on the conclusions of [15], the energy decision threshold λ was set based on the receiver noise floor [14] to achieve a constant false alarm rate of $P_{fa} = 0.01$ (thus, the resulting P_d depends on the experienced SNR). Spectrograms with one central radio channel were generated with dimensions $M \times N = 250 \times 500$ points. This spectrogram resolution was found to keep simulation times at reasonable levels without impacting the obtained results (a detailed study on different spectrogram resolutions can be found in [15]). Given a ground truth spectrogram $\mathbf{T}$, the corresponding received spectrogram $\mathbf{B}$ was generated by introducing random errors with probabilities P_{fa} and $1 - P_d$, which was then fed to the proposed iterative filtering method to produce the final output spectrogram $\mathbf{I}$. The proposed method was applied with a threshold $\varepsilon = 0.1$ dB and the maximum number of iterations was restricted according to $2K+1 \leq \max(M, N)$; the iterative filtering process in (3) was stopped if this limit was reached before the iterative stopping algorithm converged.

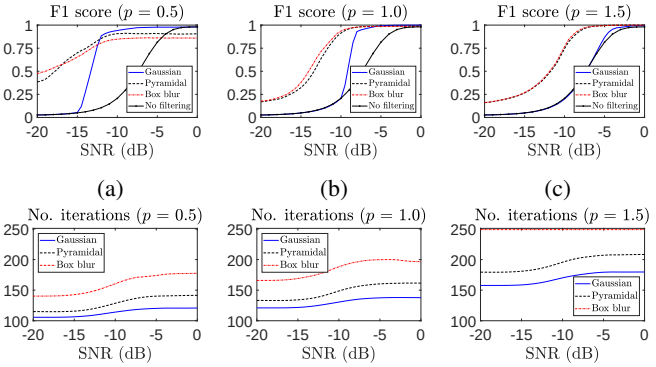
Fig. 2 shows the performance of the proposed method in an ideal best-case scenario where p is set to the optimum value[4] that maximises the accuracy for each SNR (Fig. 2a). The accuracy (Fig. 2b) is assessed in terms of the F1 score [21] by comparing $\mathbf{I}$, the output of (4), to the ground truth $\mathbf{T}$. The '*No filtering*' curve is obtained by comparing $\mathbf{B}$ and $\mathbf{T}$ and is included as a reference to show that the resulting accuracy is significantly improved when the proposed iterative filtering method is employed. The best accuracy is obtained with the Gaussian filter at low SNR (below –10 dB) and the proposed pyramidal filter at high SNR (above –10 dB). However, the Gaussian filter provides a similar level of accuracy at high SNR and would therefore be a preferred choice owing to its lower computational cost as quantified by K, the number of

[4]The optimum p can be found in simulations by means of exhaustive search over a sufficiently large interval ($p \in [0, 2]$ was found to be sufficient here).



(a) (b) (c)

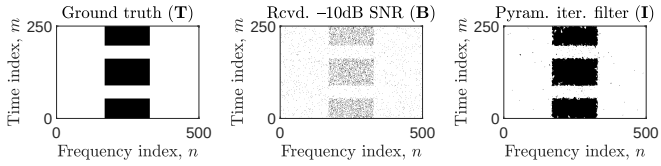
Fig. 2: Performance in an ideal scenario (with optimum p).



(a) (b) (c)

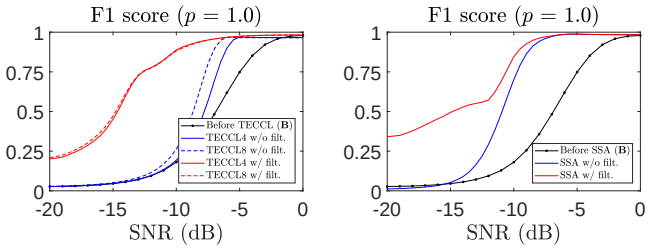
(d) (e) (f)

Fig. 3: Performance in a realistic scenario (with fixed p).



(a) (b) (c)

Fig. 4: Sample radio spectrograms: (a) ground truth at the transmitter (matrix $\mathbf{T}$), (b) observed at the receiver at –10 dB SNR (matrix $\mathbf{B}$), and (c) output after iterative pyramidal filtering, $p = 1$ (matrix $\mathbf{I}$).



(a) TECCL method (b) SSA method

Fig. 5: Performance improvement for SA estimation methods with the proposed iterative pyramidal filtering method.

iterations required for (3) to converge (Fig. 2c). Unfortunately, finding the optimum value of p for each SNR requires the knowledge of the ground truth at the transmitter ($\mathbf{T}$), which is unknown in a practical system implementation and indeed the data to be estimated. Therefore, the results shown in Fig. 2 represent an ideal upper bound to the real performance that could potentially be attained in a practical implementation.

In a practical system implementation a pre-set fixed value of p needs to be selected beforehand and used over the whole SNR range. Fig. 3 shows the performance achieved for various fixed values of p. With an adequate configuration (e.g., $p = 1$), the pyramidal and box blur filters outperform the Gaussian filter (for the reasons discussed in Section III-B) over the whole SNR range, in particular in the low SNR regime, at the expense of a higher computational cost. The box blur and pyramidal filters achieve a similar level of accuracy, however the latter requires a lower number of iterations. Therefore, the proposed pyramidal filter provides the best trade-off between accuracy and computational cost among the three considered filters in a realistic scenario (i.e., with fixed p).

Some sample spectrograms are shown in Fig. 4 to illustrate the accuracy improvement provided by the proposed method. The ground truth at the transmitter (Fig. 4a) is seen very differently by the receiver when SNR = –10 dB, for which $P_d = 0.11$ (Fig. 4b). However, the application of the proposed iterative pyramidal filtering method enhances significantly the visibility of the transmitted signals at the receiver (Fig. 4c).

The benefit that this enhancement can bring to techniques for wireless communications based on the processing of radio spectrograms is illustated in Fig. 5. This figure compares the performance of two SA estimation methods from the literature with and without the method proposed in this work. The considered methods are the Transmission Encapsulation based on Connected Component Labeling (TECCL) method proposed in [22] and the so-called Simple Signal Area (SSA) estimation method proposed in [23]. As it can be appreciated, the TECCL and SSA methods can provide accurate estimations of the SAs in the spectrogram based on the received matrix $\mathbf{B}$ (blue lines), however, the resulting accuracy of both methods is noticeably improved when matrix $\mathbf{B}$ is first processed with the proposed iterative pyramidal filtering method and the resulting matrix $\mathbf{I}$ is then provided as the input to the TECCL and SSA methods (red lines). The effect of this is an improvement of the sensitivity achieved at low SNR and therefore an expansion of the operational SNR range for both methods.

V. Conclusions

This work has proposed an iterative pyramidal filtering method to enhance the recognisability of signal components in radio spectrograms. The obtained results have demonstrated that the proposed method can extend noticeably the operational SNR range of techniques for wireless communications based on the processing of radio spectrograms.

Appendix
Pyramidal Convolution Kernel

The filter coefficients of the proposed pyramidal convolution kernel can be obtained based on the model in Fig. 6. Each filter coefficient is obtained as the height w of the pyramid's lateral face over that point. Notice that the edges of the pyramid's base are located one element beyond the actual filter radius r because the height at those edges is $w = 0$; thus the distance between the centre of the pyramid's base and any of the base edges is $r+1$. This ensures that for a filter of radius r, whose total side length is $2r+1$, there is a total of $(2r+1)^2$ non-zero filter coefficients as expected.

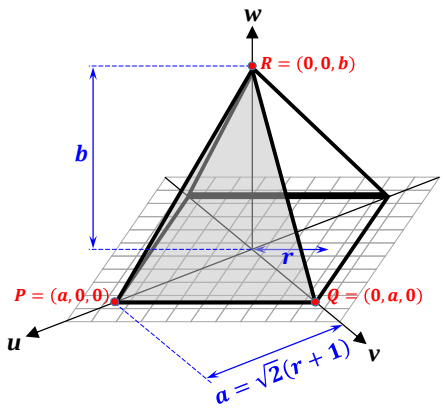


Fig. 6: Pyramidal convolution kernel.

The filter coefficients are obtained from the equations of the four planes that define the pyramid's lateral faces, which can be calculated based on basic geometry principles. Let's consider the plane on the front lateral face of the pyramid (shaded in Fig. 6). This plane contains the three vertex points $P = (a, 0, 0)$, $Q = (0, a, 0)$ and $R = (0, 0, b)$ and the edge vectors $\overrightarrow{PR} = \langle -a, 0, b \rangle$ and $\overrightarrow{QR} = \langle 0, -a, b \rangle$. A normal vector to the plane can be obtained as $\overrightarrow{PR} \times \overrightarrow{QR} = \langle ab, ab, a^2 \rangle$. Evaluating the general equation of a plane, which is given by $ab(u - u_0) + ab(v - v_0) + a^2(w - w_0)$, at any of its points (u_0, v_0, w_0), for example P, Q or R, then the plane equation $bu + bv + aw = ab$ is obtained. Repeating the same procedure for the other three planes and solving for w the following general expression is obtained for the four planes on the pyramid's lateral faces: $w = b \pm (b/a)u \pm (b/a)v$.

As shown in Fig. 6, $a = \sqrt{2}(r + 1)$, while the pyramid's height b can be freely choosen. To preserve the shape of the pyramid (i.e., the relative weight of each filter coefficient) regardless of the filter radius, which is one of the drawbacks of the Gaussian filter in the proposed method, the height can be chosen as $b = r + 1$. The plane equations then become $w = r + 1 \pm (1/\sqrt{2})u \pm (1/\sqrt{2})v$.

Finally, the pyramid (and hence the filter template) can be aligned with the axes with a rotation of $\pi/4$ radians, which can be obtained with the change of variables $x = (u - v)/\sqrt{2}$ and $y = (u + v)/\sqrt{2}$ or, equivalently, $u = (\sqrt{2}/2)(x + y)$ and $v = (\sqrt{2}/2)(y - x)$. The plane equations for the lateral faces then become $w = r + 1 \pm (x + y)/2 \pm (x - y)/2$, which leads to the cases $w = r + 1 \pm x$ and $w = r + 1 \pm y$.

The intersection of these four planes defines the surface of the pyramid's lateral faces. The height w of the pyramid surface at every point (x, y), which corresponds to the minimum of the four planes' heights at that point, provides the value of the filter coefficients $\tilde{h}_r[x, y]$ as given by equation (6).

REFERENCES

[1] Y. Jie, Y. Chenzhou, and Z. Yue, "Modulation classification based on spectrogram," *Journal of Systems Engineering and Electronics*, vol. 16, no. 3, pp. 475–488, Sep. 2005.

[2] T. J. Lynn and A. Z. Sha'ameri, "Automatic analysis and classification of digital modulation signals using spectogram time-frequency analysis," in *Proc. Int'l. Symp. Commun. and Information Technol. (ISCIT 2007)*, Oct. 2007, pp. 916–920.

[3] A. Shklyaeva, P. Kovar, and D. Kubanek, "Classification of digital modulations mainly used in mobile radio networks by means of spectrogram analysis," in *Proc. 12th IFIP Int'l. Conf. Personal Wireless Commun. (PWC 2007)*, Sep. 2007, pp. 341–348.

[4] S. Jeong, U. Lee, and S. C. Kim, "Spectrogram-based automatic modulation recognition using convolutional neural network," in *Proc. 10th Int'l. Conf. Ubiquitous Future Netw. (ICUFN 2018)*, Jul. 2018, pp. 1–3.

[5] Y. Zeng, M. Zhang, F. Han, Y. Gong, and J. Zhang, "Spectrum analysis and convolutional neural network for automatic modulation recognition," *IEEE Wireless Comms. Letters*, vol. 8, no. 3, pp. 929–932, Jun. 2019.

[6] W. Jiang, X. Wu, Y. Wang, B. Chen, W. Feng, and Y. Jin, "Time-frequency-analysis-based blind modulation classification for multiple-antenna systems," *Sensors*, vol. 21, no. 231, pp. 1–19, Jan. 2021.

[7] F. A. Bhatti, M. J. Khan, A. Selim, and F. Paisana, "Shared spectrum monitoring using deep learning," *IEEE Trans. Cognitive Commun. and Networking*, vol. 7, no. 4, pp. 1171–1185, Dec. 2021.

[8] X. Mankun, P. Xijian, L. Tianyun, and X. Mantian, "A new time-frequency spectrogram analysis of FH signals by image enhancement and mathematical morphology," in *Proc. 4th Int'l. Conf. Image and Graphics (ICIG 2007)*, Aug. 2007, pp. 610–615.

[9] Y. He, Y. Su, Y. Chen, Y. Yu, and X. Yang, "Double window spectrogram difference method: A blind estimation of frequency-hopping signal for battlefield communication environment," in *Proc. 24th Asia-Pacific Conference on Communications (APCC 2018)*, Nov. 2018, pp. 439–443.

[10] J. Kim and J. G. Andrews, "Sensitive white space detection with spectral covariance sensing," *IEEE Trans. Wireless Commun.*, vol. 9, no. 9, pp. 2945–2955, Sep. 2010.

[11] W. M. Lees, A. Wunderlich, P. J. Jeavons, P. D. Hale, and M. R. Souryal, "Deep learning classification of 3.5-GHz band spectrograms with applications to spectrum sensing," *IEEE Trans. Cognitive Commun. and Networking*, vol. 5, no. 2, pp. 224–236, Jun. 2019.

[12] S. Kojima, K. Maruta, Y. Feng, C.-J. Ahn, and V. Tarokh, "CNN based joint SNR and Doppler shift classification using spectrogram images for adaptive modulation and coding," *IEEE Trans. Commun.*, vol. 69, no. 8, pp. 5152–5167, Aug. 2021.

[13] M. López-Benítez and F. Casadevall, "Methodological aspects of spectrum occupancy evaluation in the context of cognitive radio," *European Trans. Telecommun.*, vol. 21, no. 8, pp. 680–693, Dec. 2010.

[14] M. López-Benítez, J. Lehtomäki, K. Umebayashi, and D. Patel, "Accurate noise floor calibration based on modified expectation maximisation of Gaussian mixture," in *Proc. IEEE Wireless Comms. and Networking Conf. (WCNC 2019)*, Apr. 2019, pp. 1–6.

[15] M. M. Alammar and M. López-Benítez, "Evaluation of the impact of thresholding and frequency/time resolution on signal area estimation methods," in *Proc. 93rd IEEE Vehic. Tech. Conf. (VTC 2021-Spring), 7th Int'l. Works. Smart Spectrum (IWSS 2021)*, Apr. 2021, pp. 1–7.

[16] H. Li, M. Syed, Y.-D. Yao, and T. Kamakaris, "Spectrum sharing in an ISM band: Outage performance of a hybrid DS/FH spread spectrum system with beamforming," *EURASIP Journal on Advances in Signal Processing*, vol. 2009, no. 834527, pp. 1–11.

[17] M. López-Benítez and F. Casadevall, "On the spectrum occupancy perception of cognitive radio terminals in realistic scenarios," in *Proceedings of the 2nd IAPR International Workshop on Cognitive Information Processing (CIP 2010)*, Jun. 2010, pp. 99–104.

[18] U. Sara, M. Akter, and M. S. Uddin, "Image quality assessment through FSIM, SSIM, MSE and PSNR – A comparative study," *Journal of Computer and Communications*, vol. 7, no. 3, pp. 8–18, Mar. 2019.

[19] M. W. Coffey and N. Lubbers, "On generalized harmonic number sums," *Applied Maths. and Comp.*, vol. 217, no. 2, pp. 689–698, Sep. 2010.

[20] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Syst., Man, & Cyber.*, vol. 9, no. 1, pp. 62–66, Jan. 1979.

[21] D. M. W. Powers, "Evaluation: From precision, recall and F-measure to ROC, informedness, markedness and correlation," *J. Machine Learning Technol.*, vol. 2, no. 1, pp. 37–63, Dec. 2011.

[22] J. Kokkoniemi and J. Lehtomäki, "Spectrum occupancy measurements and analysis methods on the 2.45 GHz ISM band," in *Proc. 7th Int'l. ICST Conf. Cognitive Radio Oriented Wireless Networks and Commun. (CROWNCOM 2012)*, Jun. 2012, pp. 285–290.

[23] K. Umebayashi, H. Iwata, J. J. Lehtomäki, and M. López-Benítez, "Study on simple signal area estimation for efficient spectrum measurements," in *Proc. 26th European Conf. Networks and Commun. (EuCNC 2017)*, Jun. 2017, pp. 1–5.