The Use of Linear Projections in the Visual Analysis of Signals in an Indoor Optical Wireless Link

Joe Faith and Sujan Rajbhandari
School of Engineering, Computing, and Information Sciences
Northumbria University, Newcastle upon Tyne, UK.
e-mail: joe.faith@northumbria.ac.uk, sujan.rajbhandari@northumbria.ac.uk

***Abstract*— Wavelets can be used to transform a signal into a multi-dimensional signal where each dimension represents one wavelet resolution, such that a machine learning classifier, such as artificial neural network (ANN), may be used to then classify the received signal and recover the transmitted information. Since there is no upper limit for wavelet resolution and wavelet resolution produces highly redundant coefficients, computational difficulties are when signal need to be classified. Here we demonstrate the use of dimension reduction techniques to visualise indoor optical wireless communication (OWC) signal in the presence of artificial light interference, scale reduction technique to for efficient classification and the resulting decoding errors.**

Data and software used are available from http://computing.unn.ac.uk/staff/cgjf1/cndsp.

I. INTRODUCTION

There is established interest in using the optical medium as the basis of indoor high-speed communication networks [1, 2]. Secure licence free operation in near infrared (IR) range for wireless commutation offers a bandwidth in the range of THz meaning a number of applications can be accommodated within a small region without need for complex multiplexing techniques. Unfortunately, the IR medium does have severe drawbacks, one of the most important being the intense ambient noise is generated from both natural and artificial light sources (ALSs). The common household lighting sources have a strong spectral presence in near-IR range. The cheap transceivers, on the other hand, are available in this range leading to low signal-to-noise ratio (SNR) at the receiver. In fact, if not alleviated, the interference due to ALSs is the most dominating source of the system impairment [3, 4]. A number of receiver design; optical bandpass filter and an electrical high pass filter, orthogonal polarisers coupled with a differential detector [5], differential optical filtering [6], angle diversity techniques [7] to name few; had been studied. A novel signal detection technique based on the continuous wavelet transform (CWT) and an ANN for indoor optical wireless communication (OWC) link was first proposed by Dickenson *et al* [8]. The receiver design was based on a concept of feature extraction and pattern recognition using the CWT and the ANN. The received signal in the presence of the ALI is analysed using the predefined CWT scale and classification to binary '0' and '1' is carried out using ANN. . Wavelet analysis is used to extract features from the received signal that relate to time and frequency. The proper selection of wavelet can be used to reduce the effect of the noise at the receiver. ANN can be trained to classify the received signal to binary '0' and '1' based on the wavelet coefficient. The previous study showed that CWT-ANN based receiver offer significant improvement over to basic threshold detection and can outperform maximum likelihood sequence detection (MLSD) [8].

One of the key problems using the CWT-ANN technique is the selection of proper CWT scales. The CWT analysis produces highly redundant coefficient and there is no upper limit for wavelet resolution which needs high computational requirement. Though ANN can be trained to select the proper scales, it will lead to inefficient and complex ANN structure and longer training time. Each scale corresponds to an input node of the ANN, and hence reducing scale can lead to the simpler ANN structure and reduce training time. However, the selection the wavelet scales that offers the optimum classification is challenging. Selection of the CWT scales by visual observation were carried out in [8] in which CWT of the signal with and without interference is compared visually (using naked eye) to find the optimum wavelet scales. The visual observation is always prone to human error and may not produce optimum results as the optimum scale tend to change with the data rate.

Visualizing data in a high dimensional space (here each wavelet scale corresponds to a space) is difficult task. In this study, a different technique to visualize data using of dimensionality reduction is studied so that CWT scales that offer the best classification can be identified. Though the technique can be applied to any signal classification application, the focus here is on the indoor OWC system, where optical signal is corrupted by intense ambient light noise. The paper is organised as follows: the concept of multidimensional data visualisation is discussed in Section II, followed by targeted projection pursuit (PPP) in Section III. Experimental evaluation of the proposed techniques for indoor OWC is explorer in Section IV. Finally conclusions are drawn based on the observations in Section V.

II. MULTIDIMENSIONAL DATA VISUALISATION

Signal decoding can be considered as a form of classification problem in which the received signal must be classified as carrying a 1 or 0. But when considering such problems, simple region of convergence (ROC) curves or generalisation performance statistics are not enough to fully understand the nature or source of errors. As an illustration, consider Fig. 1 which shows the performance of a hypothetical receiver on four different

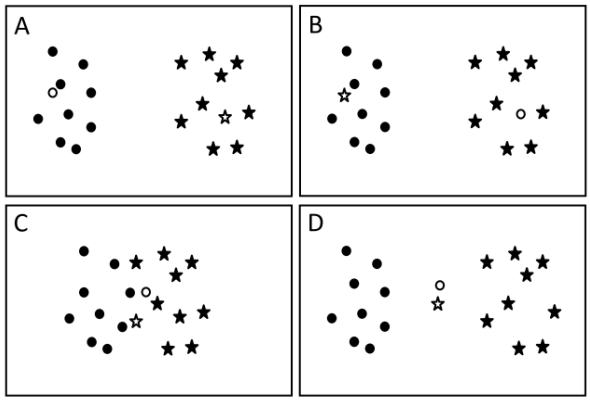


Figure 1: Scatter plots showing the performance of a classifier on four hypothetical sets of binary signals, with empty markers representing detection errors.

hypothetical sets of signals, each containing ten 1s and 0s (represented by stars and circles). (Here we assume that the signals are represented using two numerical independent variables and so can be visualised in a scatter plot.) The solid markers represent signals that were correctly classified, and the unfilled markers represent errors. Although the generalisation performance in each case is the same, the nature of the problems are different. In (A) the signal seem clearly distinguishable and we would suspect that the decoding algorithm were not suited to this particular problem, and we would choose another (or possibly that there were a bug in its implementation). In (B) it would seem that the receiver was working correctly but possible that there was some error in recording the original transmission. In the case of (C) the signal does not appear to be clearly separable and we would expect any detector to show error. And in (D) it looks like the bulk of the signal is clearly separable apart from a small number of outliers, suggesting we add an outlier detection step to our detector.

The examples illustrated here are simple, involving signals that can be adequately represented using two real attributes; but most signals, such as those discussed above, are more complex. The problem of visualising high-dimensional data is primarily one of dimensionality reduction (DR); i.e. representing the data in a low- (typically two- or three-) dimensional space such that it can be presented visually to the user. Many DR techniques are available, and most have been applied to the problem of visualisation, including non-linear methods such as Multi-Dimensional Scaling (MDS), Sammon Mapping, Self-Organising Maps (SOM), and IsoMap; and techniques based on linear projections of the data, such as Principal Components Analysis (PCA), Singular Value Decomposition (SVD), and Projection Pursuit – see Enqvist *et al* [9] for a review of such methods.

Koren and Carmel [10] argue that there are a number of advantages to using linear projections for visualising higher-dimensional data:

1. The resulting view is an accurate, if partial, representation of the data. Non-linear techniques, such as MDS, inevitably deform the topology of the data (such as the relationship between classes) in their arrangement of the data points.
2. Unlike MDS, for example, the view does not have to be re-generated every time new data is added: new data (such as a classification testing set) can be compared with the old (such as a training set).
3. Linear methods tend to be computationally efficient.
4. The resulting axes carry useful information about the data, in particular about which independent variables are the best discriminators of the independent variable.

However, unless the data contains a great deal of redundancy, any dimension reduction process involves the loss of information. The resulting view may emphasise one aspect of the original data; but other aspects are inevitably lost as the space of possibilities is reduced. Therefore one possibility is to present a range of different plots, each involving a different pair of coordinates. The scatter-plot matrix, for example, shows the data pictured against all possible pairs of coordinates, resulting in $1/2n(n - 1)$ plots, where n is the dimensionality of the original data. One dynamic alternative to such static views is Asimov's Grand Tour [11] -- described as an attempt to look at the data 'from all possible angles'. A Grand Tour is a video sequence in which each frame shows the result of a single projection of the data, with the sequence as a whole including all possible projection planes. However, the Grand Tour replaces the quality of projection pursuit with quantity: a grand tour in high dimensional space may be long and mostly uninformative.

Ideally we would have some way of allowing user to guide the tour, to use their perception of the data to find projections of interest. Cook and Buja [12] proposed and implemented an interface that allows the user to not only pause and rewind a given Grand Tour, but also to amend the resulting view by controlling the input from each dimension independently. The problem is that projection component manipulation is an opaque interface in the sense that it is rarely possible for the user to anticipate the effect of their actions. Where the user has strong intuitions about the nature of the structure of interest in data, and its relationship with the underlying coordinate system, then it may be possible for them to determine how best to use component-based controls to reveal structure in the data more clearly. In other words, once the user knows what they are looking for then such an interface will help them find it, but it is unsuited to true exploration of the data. The user has n controls to manipulate (one for each dimension of the original data set), the effect of each will be unknown and which will have unpredictable effects in combination. The user can do little more than random search -- which has its place, but is of little use when faced with a truly large dimensionality set.

III. TARGETED PROJECTION PURSUIT

Targeted projection pursuit (TPP) is a technique for exploring complex data which allows the user to manipulate their view of the data through an intuitive and transparent interface [13].

Suppose an experimenter uses a linear projection such as PCA to visualise some high-dimensional data in a two-dimensional scatter plot and they notice what seems to be a pattern, such as some clustering or a trend. The immediate question is whether another view of the data

would show the pattern more clearly, or whether it is just the result of noise. In this situation, the natural impulse is to try to "grab" those points that fail to fit the perceived pattern and try to move them into place. TPP allows the user to do just this, by playing with the data to explore possible views. Using TPP the user can try to move the data points to better fit that pattern by using a simple rubber-band drag-and-drop interface, such as selecting one of the clusters to separate it from the other data points. If there is a projection in which these changes can be found, then the points move as the user drags them. If not, then the points stick.

The overall process is thus one of hypothesis-formation and testing: by attempting to move some points, the experimenter is suggesting a hypothesis about the structure of the data; if the data fits the hypothesis, then the result is shown.

The principal advantage of such an interface is that it is transparent, in the sense that the response of the system is intuitive and predictable. If the user spots a partial pattern then they manipulate the elements of that pattern directly, rather than controls whose effects on the pattern are unpredictable. Thus this tool uses the human visual system to do what it is best at – spotting patterns – while the computer based tool is left to do 'dumb' linear algebra.

In this context we wish to use TPP to find views of the data that maximise the separation between classes (in this case representing 1 and 0 signals). This can be done with the following algorithm:

Suppose the data $\boldsymbol{X}$ is comprised of samples partitioned into k classes, such that $x_{i,j}$ is the row of $\boldsymbol{X}$ made up of the j^{th} sample of the i^{th} class. The aim is then to find, or 'pursue', a projection $\boldsymbol{P'}$ of the data that separates the classes within the two-dimensional view produced by applying that projection to the data, $\boldsymbol{V'} = \boldsymbol{XP'}$. $\boldsymbol{P'}$is found iteratively as follows.

Given an initial projection, $\boldsymbol{P^0}$, and initial view, $\boldsymbol{V^0} = \boldsymbol{XP^0}$, a target, $\boldsymbol{T^1}$, is then derived from the current view, $\boldsymbol{V^0}$, by using a simple repulsion-attraction model: the centroid of each class of points within the view is moved away from every other with a velocity inversely proportional to their separation, and each point is moved towards the centroid of its class. If the centroid of each class within the view is $\bar{v}_i^0$, the mean of $v_{i,j}^0 = x_{i,j}.\boldsymbol{P^0}$, the image of $x_{i,j}$ under $\boldsymbol{P^0}$, then the new centroid for the i^{th} class is given by:

$$\bar{t}_i^1 = \bar{v}_i^0 + k_0 \sum_j \frac{\left(\bar{v}_i^0 - \bar{v}_j^0\right)}{\left|\bar{v}_i^0 - \bar{v}_j^0\right|}. \tag{1}$$

And the new target for the j^{th} sample in the i^{th} class:

$$t_{i,j}^1 = v_{i,j}^0 + k_1\left(\bar{t}_i^1 - v_{i,j}^0\right). \tag{2}$$

The new targets for each point $t_{i,j}^1$ together form new target matrix, $\boldsymbol{T^1}$, which is then 'pursued'. That is, a new projection, $\boldsymbol{P^1}$, and view, $\boldsymbol{V^1} = \boldsymbol{XP^1}$, is found that minimises $|\boldsymbol{XP^1} - \boldsymbol{T^1}|$, using a perceptron learning algorithm [4].This process is then iterated until a convergence condition is met; in this case that the net movement as a result of projection pursuit is below a threshold.

$$\frac{|\,\boldsymbol{P}^{n+1} - \boldsymbol{P}^n\,|}{|\boldsymbol{P}^{n+1}|} < \theta. \tag{3}$$

IV. Experimental Evaluation

A. Data

The signal under analysis is an on-off keying (OOK) non-return-to-zero (NRZ) data stream that is subjected to the artificial light interference (ALI) and the Gaussian noise as shown in Fig.2. The transmitter consists of a data source and transmitter filter $p(t)$ with a unit-amplitude impulse response of one bit duration T_b. The output of transmitter filter is scaled by a peak transmitted power P_c to achieve the required average optical power. The transmitted signal is then subjected to the additive white Gaussian noise (AWGN) $n(t)$ and the fluorescent light interference (FLI) $m_{fl}(t)$. The modelling of FLI based on the experimental measurement is given in [14] and is beyond the scope of the paper. The receiver front-end consists of a photodetector with a reponsivity of R followed by a matched filter $r(t)$, matched to the transmitter filter $p(t)$ and sampler, sampling at the rate of $1/T_b$. The CWT of the sample output y_i is carried out at predefined scales, which are the input to the classifier.

The system described in Fig. 2 is simulated using Matlab and the CWT of the received OOK-NRZ signal at 100 Mbps in the presence of the FLI is given in Fig.3. For the simplicity, it is assumed that signal is free from AWGN (note that in all simulations, AWGN is assumed to be present). The CWT of signal $s(t)$ at a scale $a > 0$ and translational value $b \in \mathbb{R}$ is given by [15]:

$$F(a,b) = \int_{-\infty}^{\infty} s(t)\frac{1}{\sqrt{a}}\psi^*\left(\frac{t-b}{a}\right)dt \;; \tag{4}$$

where $\psi(t)$ is a continuous function in both the time domain and the frequency domain called the mother wavelet and * denotes complex conjugate.

The CWT provides time-frequency resolution (see Fig. 3). The wavelet coefficients between different scales are highly redundant and removing some of these scales doesn't affect the performance of the system. However, the difficult task is to selecting the optimum scales.

In this study, for the one-to-one comparisons of the visualization method, the input to the classifier is the CWT of the signal y_i at scales 1 to 150. A multilayer perceptron ANN with 6 hidden nodes and 1 output node with tan-sigmoid and linear transfer function in the hidden and output layers respectively is utilized as a classifier. Note that the number of neurons in the input is equal to number of scales in the CWT.

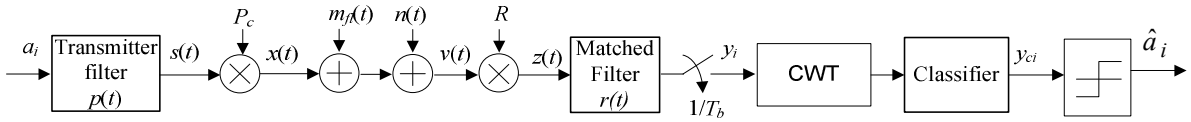


Figure 2: The block diagram of generalized OWC system.

The performance of the ANN classifier to correctly detect binary '1' and '0' for OOK-NRZ schemes at 100 and 200 Mbps are demonstrated in Fig. 4. It can be observed that there is certain overlapping between the constellation of '1' and '0' at 100 Mbps. Moreover, mis-classification due to overlapping can be clearly observed. On the other hand, the constellations are well separated at 200 Mbps and mis-classification is not observed. This clearly demonstrates that the effect of ALI is more pronounced at lower data rate. The classification at higher data is easier as the FLI has most of its spectrum at low frequencies.

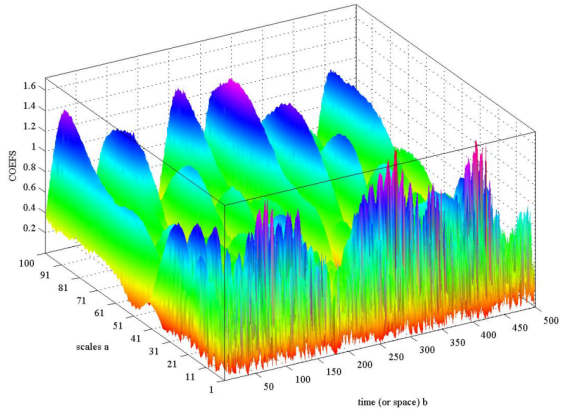
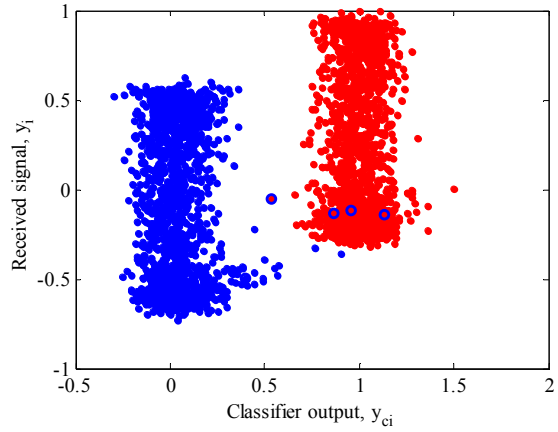
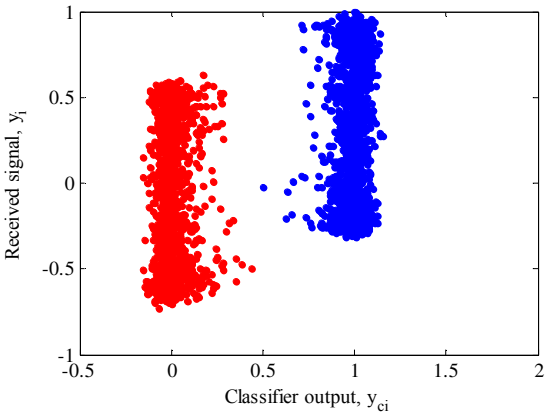


Figure 3: The CWT of received OOK signal at 100 Mbps.



(a)



(b)

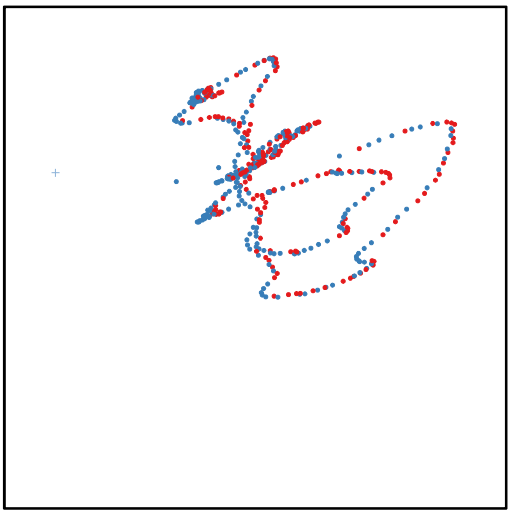
Figure 4: Scatter plots showing the performance of a classifier on (a) 100 Mbps OOK data at SNR of 15 dB and (b) 200 Mbps OOK data at SNR of 15 dB. (Note, the color denotes the transmitted data (red = 1, blue = 0) and empty markers representing detection errors).

B. *Method*

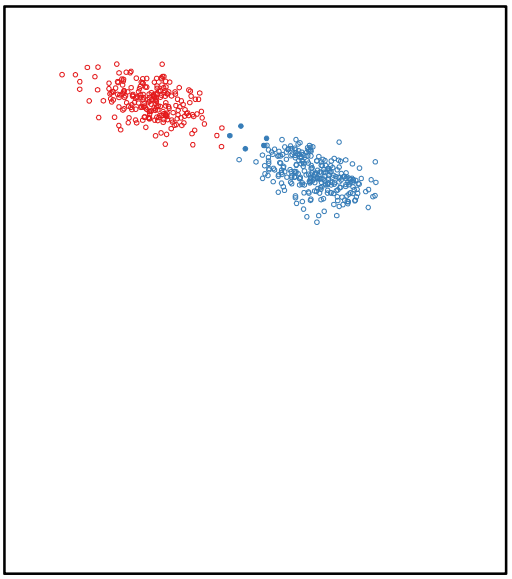
TPP was used to visualise the 100 Mbps OOK-NRZ data at SNR of 15 dB, comprising the first 150 wavelet components for each signal. All 5 misclassified signals were included, along with a control group of 500 randomly selected correctly classified signals. A projection was found to show the separation between the values of the transmitted signal. The 20 most significant components in the projection were taken and used to train and test an ANN classifier.

C. *Results*

Fig. 5a shows a view of the received signals plotted against the first two principal components. It is



(a)

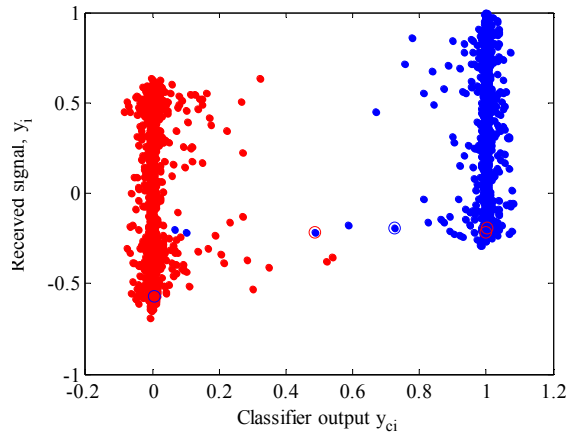


(b)

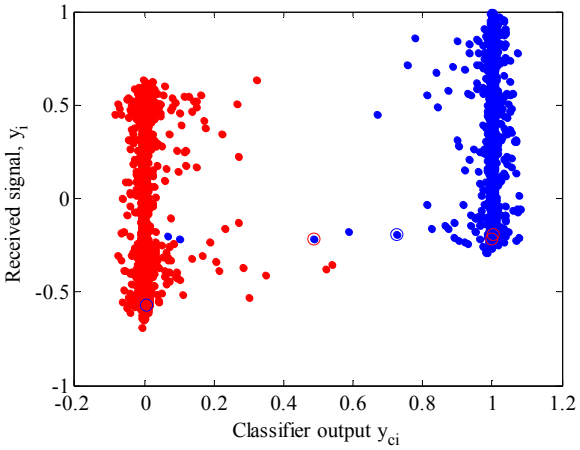
Figure 5: Scatter plots of 150 wavelet components of the 100 Mbps OOK data at SNR of 15 dB produced using linear projections. Color denotes the transmitted data (red = 1, blue = 0). (a) shows the data plotted against the first two principal components. (b) shows the view produced using TPP with empty markers denoting correctly classified instances and filled markers representing mis-classifications.(Note that images are scalable, and are best viewed at a larger size).

immediately apparent that the signal is dominated by the low frequency ALI, with the points representing the received signals forming a trajectory tracing the changing effect of interference. This is confirmed by Table I which shows the largest components with the largest weights in the respective projections. In the case of PCA, the largest components are all in the range 121-140, corresponding to the dominant frequencies of the ALI. Fig. 5b shows the view of the same signals that was found by TPP to maximise the separation between the values of the transmitted bits. Here we can see that the signal is not completely separable, with some confusion at the plane of separation. The few signals that were misclassified by the ANN are shown as filled markers and these are found, as we would expect, near the plane of separation. However we can also see that all the errors in this particular set were due to 0's being misclassified as 1's, so it may be possible to 'hand-tune' the ANN to remove these errors by increasing the threshold slightly. (Note that the perceptron convergence theorem [16] does not guarantee that an ANN converges to an optimal solution in all cases; only where the data is linearly separable. The use of online, rather than batch, training may also prevent an optimal solution being found.)

As Koren and Carmel suggested, the components of the projections underlying these views also contain useful information. The principal components are dominated by the dominant frequencies of the ALI, but the components



(a)



(b)

Figure 6: Scatter plots of classifier output of the 100 Mbps OOK data at SNR of 15 dB produced using linear projections (a) full scales from 1-150 and (b) reduced scale given in Table I. Color denotes the transmitted data (red = 0, blue = 1).

TABLE I.
THE DOMINANT SCALES FOR EACH OF THE LINEAR PROJECTIONS IN FIG. 5

PCA	128,127,130,132,139,129,135,133,134,137,126,131,140,138,123,136,124,125,121,122
Separation by transmitted bit	1,2,3,5,4,8,10,6,14,7,16,150,12,9,20,46,149,48,44,26

of the separated view indicate which may the best components for use in an ANN. The components shown in Table I indicate that these include the highest frequency components (1-20), as expected, but also that some much lower frequencies (44, 48, 149, 150) may be useful in classifying these signals (note that scale in wavelet domain are proportional to the inverse of the frequency). The classification of the received signal based on the 20 dominant scales and scales from 1:150 is given in Fig. 6.The striking similarities between the classification based on all scale and dominant scales (given in Table I) can be observed. In fact, the number errors in the classified signal are equal. Thus, the classification of the received signal can be done with much reduced scale (20 scales instead of 150) which significantly reduces the overall complexity. The visualisation technique also can also give good indication of the dominant scale and possibility of error position.

V. CONCLUSIONS

Visualisations play an essential role in the scientific process because they reveal aspects of the data that may not be apparent by standard statistical means, but visualisation is difficult where the data is high-dimensional. It has been suggested that an advantage of linear projections for reducing the dimensionality of data for the purposes of visualisation is that the components of the resulting projections may themselves include useful information about the data. This has been confirmed by showing how 2 dimensional scatter plots of the data produced using linear projections show the effect of interference on an optical wireless signal; show the error in classifying that signal and suggest possible solutions; and suggest ways to increase the computational efficiency of the receiving classifier.

REFERENCES

[1] G. Ntogari, T. Kamalakis, and T. Sphicopoulos, "Performance analysis of space time block coding techniques for indoor optical wireless systems," *IEEE Journal on Selected Areas in Communications,* vol. 27, pp. 1545-1552, 2009.

[2] M. D. A. Mohamed and S. Hranilovic, "Optical impulse modulation for indoor diffuse wireless communications," *IEEE Transactions on Communications,* vol. 57, pp. 499-508, 2009.

[3] A. J. C. Moreira, R. T. Valadas, and A. M. d. O. Duarte, "Performance of infrared transmission systems under ambient light interference," *IEE Proceedings -Optoelectronics,* vol. 143, pp. 339-346, 1996.

[4] R. Narasimhan, M. D. Audeh, and J. M. Kahn, "Effect of electronic-ballast fluorescent lighting on wireless infrared links," *IEE Proceedings - Optoelectronics,* vol. 143, pp. 347-354, 1996.

[5] S. Lee, "Reducing the effects of ambient noise light in an indoor optical wireless system using polarizers," *Microwave and Optical Technology Letters,* vol. 40, pp. 228-230, 2004.

[6] A. J. C. Moreira, R. T. Valadas, and A. M. Duarte, "Reducing the effects of artificial light interference in wireless infrared transmission systems," in *Proceedings of IEE Colloquium on Optical Free Space Communication Links*, London, U.K, 1996, pp. 5/1-5/10.

[7] R. T. Valadas, A. M. R. Tavares, and A. M. Duarte, "Angle diversity to combat the ambient noise in indoor optical wireless communication systems," *International Journal of Wireless Information Networks,* vol. 4, pp. 275-288, 1997

[8] R. J. Dickenson and Z. Ghassemlooy, "A feature extraction and pattern recognition receiver employing wavelet analysis and artificial intelligence for signal detection in diffuse optical wireless communications," *IEEE Wireless Communications,* vol. 10, pp. 64-72, 2003.

[9] N. Elmqvist, P. Dragicevic, and J. D. Fekete, "Rolling the Dice: Multidimensional Visual Exploration using Scatterplot Matrix Navigation," *IEEE Transactions on Visualization and Computer Graphics,* vol. 14, pp. 1539-1148, 2008.

[10] Y. Koren and L. Carmel, "Robust linear dimensionality reduction," *IEEE Transactions on Visualization and Computer Graphics,* vol. 10, pp. 459-470, 2004.

[11] D. Asimov, "The grand tour," *SIAM Journal of Scientific and Statistical Computing,* vol. 6, pp. 128–143, 1985.

[12] D. Cook, A. Buja, J. Cabrera, and C. Hurley, "Grand tour and projection pursuit," *Journal of Computational and Graphical Statistics,* vol. 4, pp. 155–172, 1995.

[13] J. Faith, "Targeted projection pursuit for interactive exploration of high- dimensional data sets," in *11th International Conference Information Visualization, 2007*, 2007, pp. 286-292.

[14] A. J. C. Moreira, R. T. Valadas, and A. M. d. O. Duarte, "Optical interference produced by artificial light," *Wireless Networks,* vol. 3, pp. 131-140, 1997.

[15] S. G. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 11, pp. 674 - 693, 1989.

[16] A. B. Novikoff, "On convergence proofs on perceptrons," *Proceedings of the Symposium on the Mathematical Theory of Automata,* pp. 615–622, 1962.