

Identification Of Conveyor Belt Splices And Damages Using Neural Networks

M Alport^a, P Govender^a, S Plumb^b and L van der Merwe^c

- ^(a) Applied Physics Group, School of Pure and Applied Physics, University of Natal, Durban, South Africa.
- ^(b) CheckIt Systems, CheckIT Systems, Suite 307, Instinct House, 116 Stamford Hill Rd Durban, South Africa.
- ^(c) Richards Bay Coal Terminal, Richards Bay, South Africa.

Introduction

Richards Bay Coal Terminal (*RBCT*), situated 150 km north of Durban is the largest single such terminal in the world. The second largest is Port Waratah Coal Services (*PWCS*) at Newcastle in Australia. *RBCT* receives coal by railroad where it is stored before being loaded onto bulk shipping carriers for export. *RBCT* exported 62,3Mt in 1997. Within the terminal, the coal is transported exclusively by 85 km of polyester core conveyor belts with an annual belting budget of \$2.1M. Unscheduled belt stoppages due to breakages also incur heavy penalties due to ship demurrage, stacker delays and tippler standing. Over the past two years, preventative belt maintenance has been given a high priority.

The *Applied Physics Group* was commissioned to develop a high-speed video imaging solution that could be used to survey 7-8 critical belts on a monthly basis. The belts themselves are typically 2.2 m wide, high-tension polyester cloth-based, and run at a speed of 6 m/s. Since the terminal runs for 24 hours per day, it was necessary to image the belt whilst it was in motion. This is achieved with a high-speed digital video camera. The video clip is subsequently analysed to evaluate the quality of splices and identify other surface damages. The essential problem here is to automatically locate splices and damages, without having to manually examine the typically 3 000- 15 000 frames in each video clip.

Although superficial video imaging of the surface covers only gives a partial indication of the condition of the interior, this information has been successfully used to identify precursors to failure. However the image processing algorithms that we have developed may be readily applied to x-ray surveys of steel core belts and some preliminary measurements have been taken with a linear diode array.

Video Surveys

Initially, the first solution (*BeltQC I*) implemented at *RBCT* consisted of a permanent video camera mounted under conveyor #631 with the live video being returned to the top of the control tower located about 0.5km away via a microwave link. The capture of the video stream and the analysis of the belt images were done on a PC connected to the *RBCT* LAN. However, this turned out to be an inappropriate choice of technology for that very harsh and corrosive industrial environment. Essentially the camera, its optics, the microwave and telemetry links required an onerously high level of routine maintenance after being subjected to blown coal dust, high temperatures and high pressure hosing. In addition, the scaling required to monitor additional belts multiplied the maintenance responsibility and capital commitment linearly with the number of belts. Finally, after about 4 months of operation, this phase of the project was terminated when a total failure of belt 630 destroyed the camera enclosure.

BeltQC II was then implemented, based on a more appropriate service model, which involved monthly video surveys of the 8 critical belts. Each survey typically takes 1-2 days depending on lighting conditions and belt stoppages etc. The video tapes were then returned to the laboratory, digitised and analysed with a hardcopy report being sent to the client within two weeks. The belt splice and damage images and analyses is also updated every month on a web site. Fortunately, the timescale of belt splice deterioration is in excess of one month so

the sampling frequency of one month has been adequate. We have been operating in this mode for over two years now, with the renewed 2001 contract being serviced by CheckIT Systems, a company which has spun off from the Applied Physics Group and in which the University of Natal has an interest. Although the abrasive coal dust environment still takes a toll on video recording equipment, routine servicing and underwater enclosures have ensured minimal downtime due to equipment failure.

Video Clip Software

Typically we film the conveyor belt by setting up the video camera on a tripod about 4m from a roller. The roller curvature tends to open up the splice. Initially we used a Sony Hi8 video camera but more recently have switched to a digital format. At each belt, sufficient video footage is obtained to span more than one revolution of the belt. These video clips are then returned to our laboratory where they are first digitised by a frame grabber into .AVI format files with a typical frame resolution of 640×480 pixels. Initially these were then converted into sequential .jpg images, but now we are able to work directly from the .AVI files. Fortunately, there is very little colour information in the belt images so the three colour planes can be reduced to one greyscale plane. We then run the splice/damage detecting routines on each image by single stepping through the AVI file. The results are written to an Access database. Since we only store the frame number, AVI filename and some other data relating to the survey date etc, the database is a manageable 3 Mbytes per year. Initially, we were storing the extracted BMP images of splices and damages in the database as well but this quickly produced an unwieldy, fragile .mdb file. We use CD writers for all our backups and archiving, finding them to be relatively cheap and very reliable. We still await the move to DVD writers.

A number of Visual Basic programs have been written that step through the Access database, apply relevant filters such as belt number, survey date etc and then produce an MSWord document containing tables of splice/damage positions, images etc. In spite of attempts to distribute our report on the web, we still find that the client prefers to consult the hardcopy version.

Fortunately, the PAL frame rate of 50 frames per second produces overlapping images in consecutive frames since at 6m/s, the belt travels a distance of $6\text{m/s} \times 0.04\text{s} = 0.24\text{m}$ in one frame time. Due to interlacing, each video frame actually consists of two alternating sets of odd and even lines. Since the odd and even fields are captured 20ms apart, the output from the frame grabber does need to be processed before any imaging algorithms can be applied. This involves extracting the odd and even lines of the digitised image and then assembling these into two separate odd and even fields. These odd and even images have a horizontal resolution equal to that of the original composite image but a vertical resolution which is half that of the original. Additionally these sets of images are captured 1/50s apart i.e. at twice the PAL frame rate. We can thus use these odd and even images independently, depending on which has the better display of the damage or splice.

The camera itself needs to be set to a high (~1/2000s) effective shutter speed to freeze the belt without excessive blurring. Here the criteria is that during the time when the shutter is open, the belt must not move by a distance which is less than the scale length of the smallest desired spatial feature.



Figure 1: An example of a single video frame showing a conveyor belt splice which exhibits some delamination.

Figure 1 shows an example of one of the frames extracted from a video clip of the conveyor with a diagonal delaminating splice running from the bottom left to the top right of the roller. This example illustrates a number of common characteristics of the images. Since the video clips are taken under available ambient lighting conditions, often the full length of the belt is not optimally illuminated. In a single frame there are wide variations of light intensity from the bright sky to the underside of the roller. High contrast shadows from supporting structures also often fall across the region of interest.

The essential problem here is to automatically locate subtle variations in the belt due to splices and damages, without having to manually examine the typically 10,000 frames in each video clip.

Image Processing

The onerous task of manually single stepping through ~15,000 frames for each of 8 belts every month quickly persuaded us that more efficient automatic algorithms needed to be developed. We first needed to identify what properties of the image were associated with the features of interest namely the splices and damages. The splices were generally narrow, darkened regions running at an angle of about 15° to the horizontal. Damages could be almost anything, but for convenience, were classified into one of the following categories:

- a) Vertical gauge
- b) Horizontal gauge
- c) Rip
- d) Abrasion

For the purposes of this paper, we will only be discussing the problem of distinguishing “splice” from “belt”. And for these purposes, we explored a number of techniques that targeted a specific defining characteristic of these features. However, to retain the generality of the techniques, it was necessary to ensure that the different damage features as well splices could be identified.

Non-spectral techniques

A conveyor belt in good condition has very little grey scale variation over its surface, whereas a damage or splice has a spatial variation in its greyscale. We first used this discriminating characteristic to consider non-spectral techniques such as the δ -discriminator, contrast and Radon transforms. Non-spectral techniques do not explicitly use the spatial scale lengths of an image. The δ -discriminator technique uses the sum of the edge enhanced belt image as a measure of the magnitude of spatial variations. Suitable image gradient operators are the Sobel or Roberts algorithms[1]. Figure 2 shows the output from the δ -discriminator used to identify the splices in the video clip which are marked with circles. There are a number of false indicators.

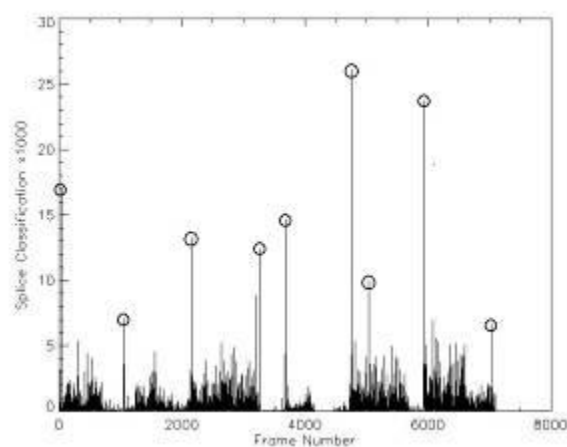


Figure 2. The output from the δ -discriminator used to identify the nine video frames containing images of splices.

Contrast, calculated as the difference between the maximum and minimum pixel values in the region of interest, also worked fairly well, but gave a large number of false indicators where the belt was uniformly rough for example. We also explored the Radon transform which may be used to detect linear features within an image. The Radon transform[2] of an image $A(x,y)$ is defined as:

$$R(\rho, \theta) = \int_{-\infty}^{\infty} A(\rho \cos \theta - s \sin \theta, \rho \sin \theta + s \cos \theta) ds$$

This equation calculates the integral along a line s through the image, where ρ is the distance of the line from the origin and θ is the angle from the horizontal. Thus a single line through the image, along a dark splice for example, is transformed to a low valued single point in the ρ, θ plane. This succeeded in identifying (linear) splices, but there was no clear extension to identify irregularly shaped damages.

Spectral techniques

Damages such as abrasions, horizontal and vertical gauges have characteristic spatial sizes, so it seemed reasonable to also examine techniques that could discriminate in terms of the horizontal and vertical scale lengths.

The Fast Fourier Transform (FFT) is able to provide spatial information from two dimensional images, and are currently the de facto industry standard for the JPEG format. However, *Wavelet Transform* (WT)[3], is becoming an increasingly accepted replacement for the JPEG image compression standard. Wavelets are more suited to compressing data with local structure and specifically have local support unlike the extended sine and cosine bases used for the FFT. Figure 3 shows that at compression ratios in excess of 70, the Daubechies [4] wavelet algorithm produces coefficients that more closely approximate the original image than the FFT. This would suggest that the FFT or WT coefficients could be used to characterise the damages and splices on the belt images and that we might expect the WT to perform better. The next problem was to determine which combination of coefficients code for spatial structures of interest. One way to discover this mapping is to train a Neural Network[5] using the FFT or WT coefficients as inputs and the specific feature classification as an output.

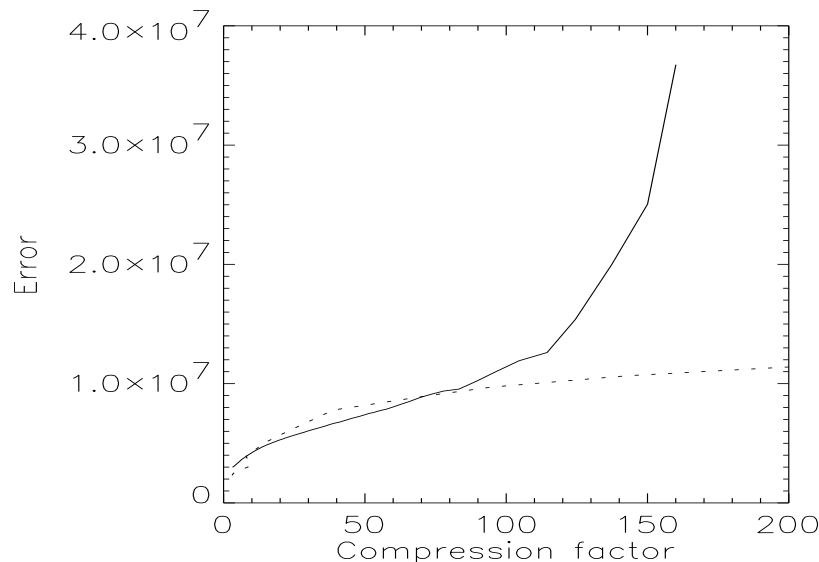


Figure 3 Comparison between the error achieved by the JPEG and wavelet compression of a greyscale image as a function of the compression ratio.

Artificial Neural Networks

Artificial Neural Networks (ANN) have been used in a wide variety of applications over the past 10-15 years, and have been fairly successful in dealing with real world complexity. Commonly, physics and the scientific method specifically, prefer to deal with a simple isolated arrangement such as a single charge in a vacuum. The interaction of two or more particles quickly degenerates into the unsolvable many-body problem. Indeed, some may point out that within the quantum mechanical world, vacuum fluctuations make the situation of empty space intractable in itself!

An ANN maps a series of inputs to a series of outputs as shown in Figure 4. There may be a number of layers including one or more hidden layers with the outputs of one forming the inputs of the following layer. At each of these layers, the inputs (x_i) and outputs (y_i) are related by a linear mapping

$$y_i = \sum_j w_{ij} x_j$$

with the w_{ij} 's being determined by the desired outputs using an initial training set.

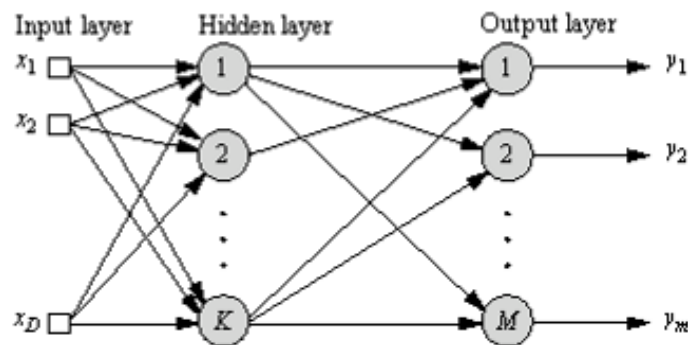


Figure 4: Topology of an artificial neural network (ANN).

The presence of one or more hidden layers allow the ANN to define non-linear relationships. There are a number of different ANN configurations and architectures. The Multilayer Perceptron (MLP) is most suited to our application and we use the WTN or FFT coefficients as inputs and the classification of the structure (damage or splice) as an output for the ANN. As a first step, we have restricted our investigation to classify input belt images as either “good belt” or “splice belt”.

To reduce the dimensionality of the input data set, we used the first 72 major FFT coefficients and an equal number of the WT coefficients. Furthermore, since there is a lot of structure outside the region of interest defined by the belt, the average image, obtained by averaging the previous 15 images was first subtracted from the target image before the WT or FFT coefficients were computed.

The process is as follows:

1. Present a training set of input coefficients and output classifications to the ANN. A splice is given the arbitrary output value of “0” and the belt a value of “1”. The training set comprises 20% of the available data set.
2. Adjust the w_{ij} 's, minimising the error between the ANN output and the desired classification of 0 or 1.
3. During the training phase, 20% of the available data is reserved for cross validation. The training process is terminated when either the error of the training set and cross validation diverge or the training error decreases to an acceptable value – typically 0.0001.
4. After training is completed, the weights are frozen and the remaining 20% of the data set is presented to the ANN as unseen randomly chosen test set.

An extremely important part of the training process is to ensure that a clean data set is presented to the ANN. For this reason we classified the input data consisting of examples of splice and belt as shown in Table 1 according to a system which graded the splice and belt images.

Table 1. Classification codes and criteria used to filter the database images. The image archive for 8 conveyor belts spanned a period of 7 months in 2000.

Classification Code	Classification Criteria
0	Bad belt image – unusable.
1	Very good “good belt” image – little noise present & no structures present.
2	Good “good belt” image – some noise & no structures present.
3	Average “good belt” image – lots of noise & no structures present.
4	Poor “good belt” image – small indistinct splice structure present.
5	Average “good belt” image – structure other than splice present in belt image.
6	Very bad “good belt” image – distinctive splice structure present in belt image. Can be mistaken for splice structure.
7	
8	No splice structure present.
9	Very poor “splice belt” image – very small angled lines & lots of noise present.
10	Poor “splice belt” image – very small angled lines.
11	Negative splices – splice structure is angled line of zero pixels.
12	Average “splice belt” image – angled line < ¼ belt length, obscured.
13	Average “splice belt” image – angled line < ½ belt length & > ¼ belt length, obscured.
14	Average “splice belt” image – angled line > ½ belt length, obscured.
15	Good “splice belt” image – angled line < ¼ belt length, partially obscured.
16	Good “splice belt” image – angled line < ½ belt length & > ¼ belt length, partially obscured.
17	Good “splice belt” image – angled line > ½ belt length, partially obscured.
18	Very good “splice belt” image – angled line < ¼ belt length, very distinct
19	Very good “splice belt” image – angled line < ½ belt length & > ¼ belt length, very distinct
20	Very good “splice belt” image – angled line > ½ belt length, very distinct
21	Belt image contains damage other than splice
22	Unclassified belt image

Photo 1 shows two examples of a subtle obscured splice image and a more distinct splice image to illustrate the variability of this feature.

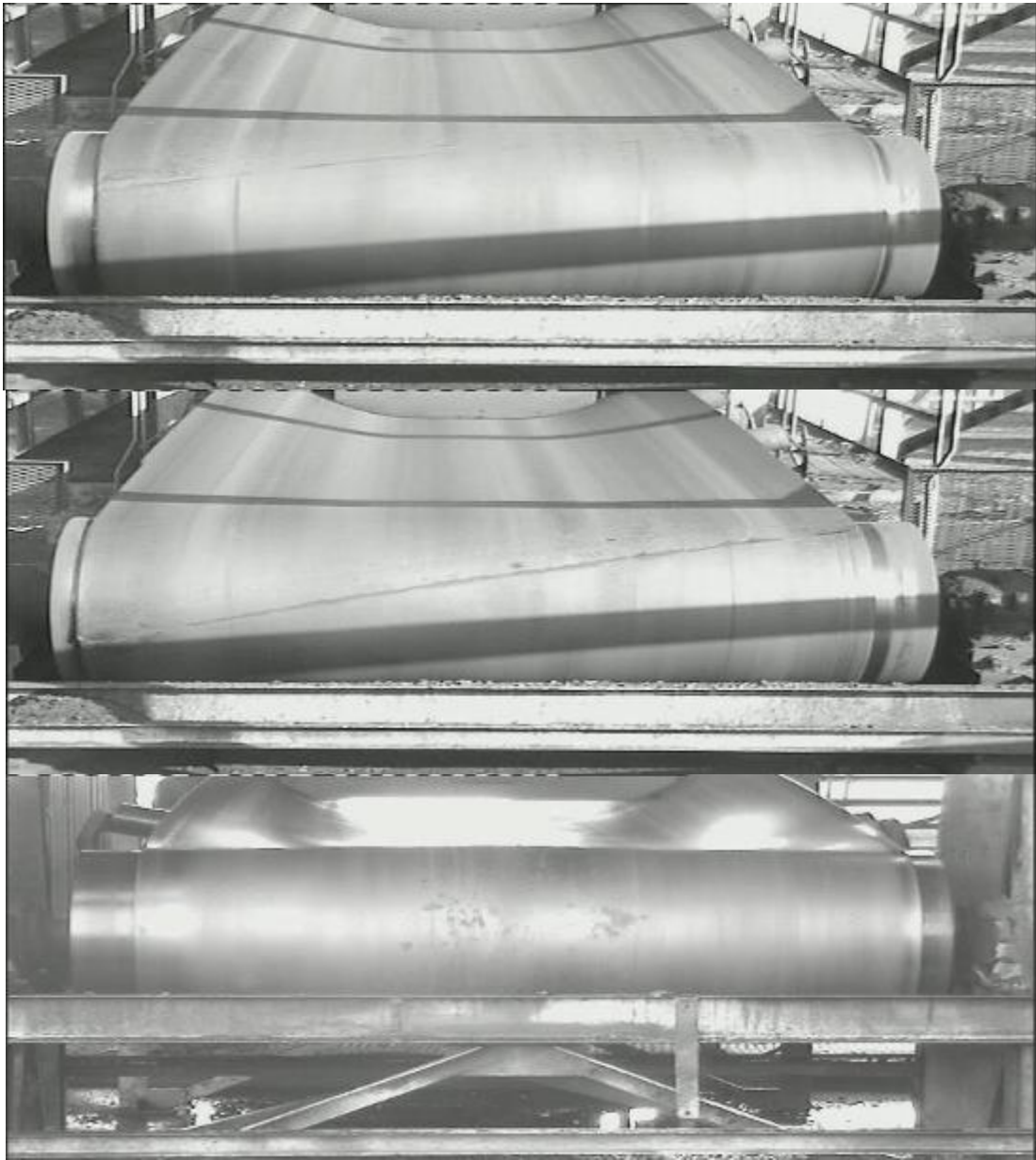


Photo 1. Examples of classification categories used to rank the different splice and belt images.

Top: Category 12 - Average “splice” image – angled line $< \frac{1}{4}$ belt length, obscured.

Middle: Category 20 - Very good “splice” image – angled line $> \frac{1}{2}$ belt length, very distinct.

Bottom: Category 21 with damage other than splice present.

The classifications shown in Table 1 were then grouped into six input data sets NN1, NN2...NN6 as shown in Table 2.

Table 2. Composition of each of the six data sets obtained by taking various combinations of the images classified as splices and belt from Table 1.

Data Set	Splice Codes	#	Belt Codes	#
NN1	12-20	240	1-3	608
NN2	15-20	137	1-3	608
NN3	9-20	349	1-3	608
NN4	12-20	240	1-5	702
NN5	15-20	137	1-5	702
NN6	9-20	349	1-5	702

The number of examples in each set is shown in the column to the right. For example, Data set NN1 was constructed by combining all images with splice classification codes in the range 12 to 20 and all belt classification codes in the range 1 to 3 resulting in a data set of 240+608=848 exemplars.

Results

Following training and testing, the ANN results for the WT and FFT coefficients were tabulated for each of the input data sets NN1, NN2 ...NN6 as shown by the confusion matrixes in Table 4 and Table 5.

The confusion matrix is a useful measure of the ability of the ANN to classify the inputs into the appropriate outputs and, for our example of two possible categories, has the form shown in Table 2: Each entry in the 2x2 confusion matrix is expressed as a percentage.

Table 3. Structure of the Confusion matrix.

	Splice	Belt
Splice	Splice classified as splice	Splice classified as belt
Belt	Belt classified as splice	Belt Classified as belt

Table 4. Confusion matrixes for each of the three Training, Cross Validation and Testing subsets using the Wavelet SOSQ input files.

Data Set	Training		Cross Validation		Testing	
NN1	97.04	2.96	97.04	2.96	93.75	6.25
	0.0	100.0	0.0	100.0	10.42	89.58
NN2	100.0	0.0	100.0	0.0	100.0	0.0
	0.0	100.0	0.0	100	3.70	96.30
NN3	97.61	2.39	71.64	28.36	75.36	24.64
	12.18	87.82	5.26	94.74	14.81	89.86
NN4	97.08	2.92	98.18	1.82	95.83	4.17
	0.58	99.42	9.52	90.48	8.3	91.67
NN5	100	0.0	100.0	0.0	88.89	11.11
	0.0	100	0.0	100.0	3.70	96.30
NN6	92.89	7.41	88.33	11.67	89.86	10.14
	4.76	95.24	33.33	66.67	23.19	76.82

Table 5. Confusion matrixes for each of the three Training, Cross Validation and Testing subsets using the FFT coefficient input files.

Data Set	Training		Cross Validation		Testing	
NN1	95.59	4.41	82.14	17.86	83.33	16.67
	8.14	91.86	35.0	65.0	27.08	72.92
NN2	98.73	1.27	62.07	37.93	59.26	40.74
	1.03	98.97	27.27	72.27	14.81	85.19
NN3	96.52	3.48	61.54	38.46	54.29	45.71
	21.54	78.46	30.30	69.70	20.00	80.00
NN4	98.52	1.48	59.65	40.35	66.67	33.33
	21.97	78.03	22.22	73.68	12.5	87.50
NN5	98.70	1.30	77.42	22.58	77.78	22.22
	1.01	98.99	22.22	77.78	3.70	96.30
NN6	96.52	3.48	60.26	39.74	44.29	55.71
	22.36	77.64	27.27	72.73	20.00	80.00

Considering the test confusion matrixes in Table 4 and Table 5, it is seen that the WT coefficients correctly identify splices with a frequency of 89-100%, depending on the diversity of the input data set. This is to be compared with the significantly reduced success rate for the FFT coefficients of 44-83%. This ranking may also be seen by plotting the actual ANN output on a scale of 0 to 1 (0=splice, 1=belt) for the WT and FFT inputs as shown in Figure 5. Clearly, the WT coefficients (top panel) produce a more reliable robust classification.

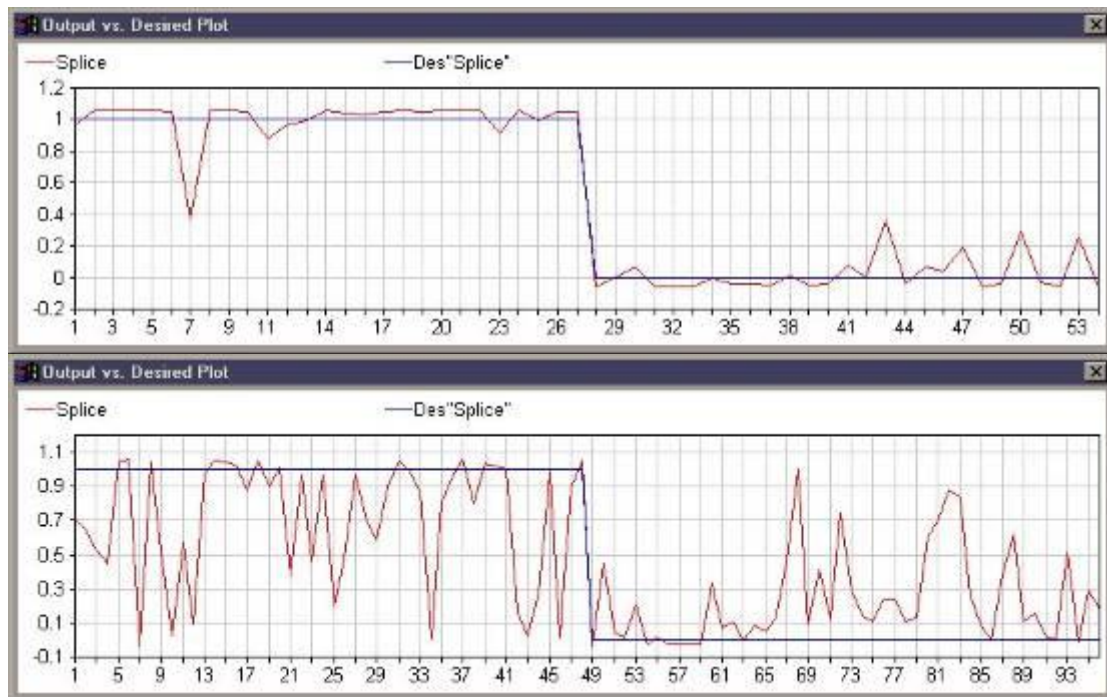


Figure 5. Plot of output vs desired for the Wavelet (top) and FFT (bottom) MLP neural network using data set NN1.

The overall performance of the WT and FFT ANN's is shown as a histogram in Figure 6. Apart from the WT performing better than FFT for most splice and belt images, it is interesting that they have similar magnitudes for the belt images and data sets NN4, NN5 and NN6. This suggests that the FFT coefficients are equally able to characterise smooth images.

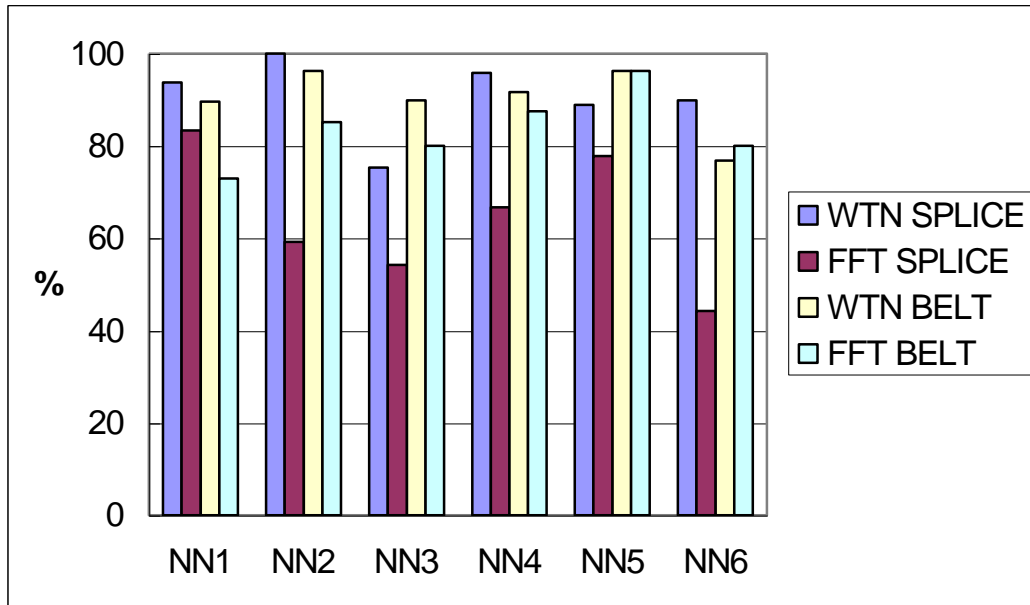


Figure 6. Comparison of the % correct splice and belt ANN classifications using the WT and FFT coefficients.

It is often said that more can be learnt from failures than successes and this is also true when examining the wrongly classified images. Photo 2 shows an example of a splice image that was incorrectly classified by the ANN as a “belt” image. The grey-scale variation across the splice is indeed subtle, and without the use of oblique lighting, it is perhaps not unreasonable that such a mistake was made. One possible improvement could be to perform a histogram equalization on the image before the coefficients are calculated.

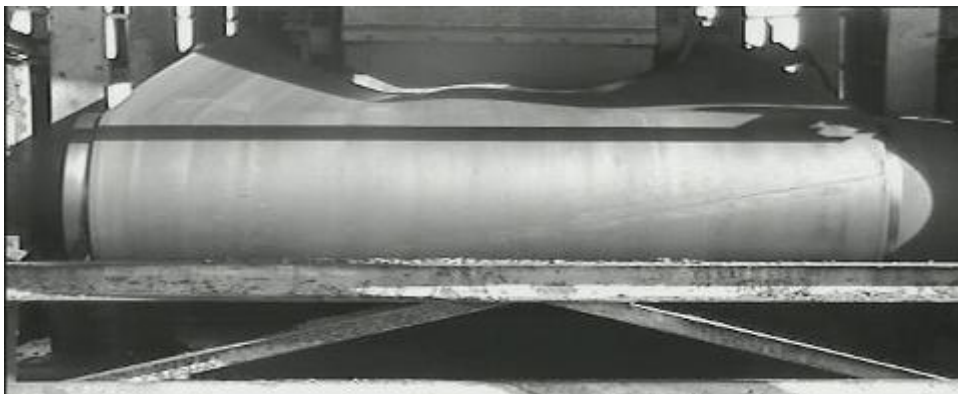


Photo 2. Example of a “splice” image that was wrongly classified as a “belt” image.

The performance of the belt classification could be improved by the following:

1. Using larger sample sets in the training of the neural network. The sample sets used were relatively small, with the biggest set of ANN input files contained 698 records in total. It is always necessary for a wide range of input data to be presented to the ANN during the training phase.
2. Using longer training cycles. Relatively short training cycles of 1,000 to 5,000 iterations were used.

3. Improving the quality of the input images. This is mainly an operational request and using an analysis of miss-classified images, specific confusing artefacts such as shadows can be avoided in the future. It may also be possible to use histogram equalisation to enhance the structure of subtle features.

Keograms

Having located the splices, we also generate a special view of each belt that allows the client to view the whole belt in one image called a Keogram[6]. A Keogram or BeltGraph, as shown in Photo 3, is constructed by extracting a single horizontal line of pixels for each image. These lines are then stacked side by side so that the full belt can be displayed in a single image. This image is then annotated with the splice positions. The Keograms are used to correct tracking of belts since this behaviour is almost impossible to observe these dynamics from a fixed point. Skew splices are detected and internal carcass deviations (for example folds in fabric) become visible as lines on the rubber and can be used for warranty purposes and /or product improvement with suppliers.

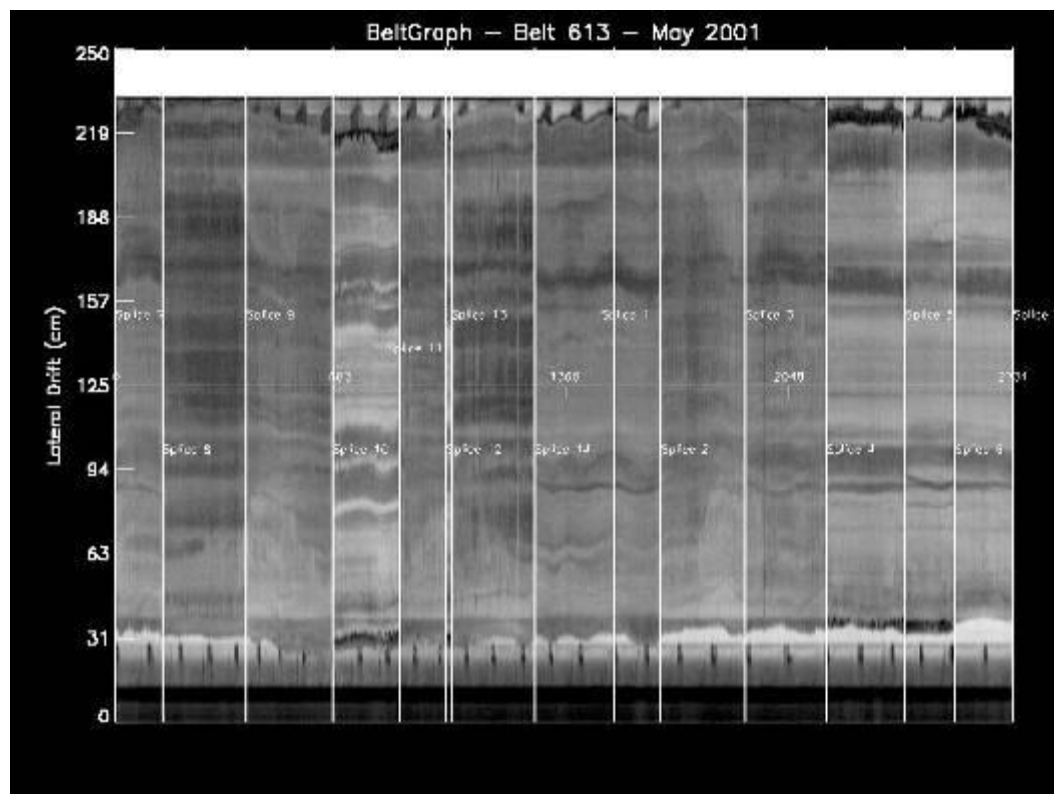


Photo 3. An example of a Keogram for Conveyor #613. The splices, together with their positions are shown as an overlay.

Conclusions and Discussion

The automatic identification of splices from recorded video clips of a moving high speed conveyor belt has been achieved with a promising degree of accuracy using wavelet coefficients as inputs to an MLP artificial neural network. The low number of wavelet coefficients used effectively reduces the dimensionality of the input image pixel space.

We have further presented a comparison of WT and FFT coefficients which show that the wavelet algorithm more accurately discriminates between splices and belt features. The output from the neural network (scaled between 0 and 1) gives a direct measure of the confidence of that particular classification. In practice, this would only leave a small proportion of the total number of input images that would require manual classification.

The use of wavelet coefficients as inputs to the ANN has the advantage that the algorithms may be readily extended to classifying other features. Specifically the added classification of

vertical and horizontal gauges, rips and abrasions should be possible and this is currently being investigated.

Finally it should be pointed out the neural network based imaging techniques developed here are not limited to the optical video surveys, but are equally applicable to x-ray imaging of steel core conveyor belts.

From a maintenance perspective, RBCT personnel have found the BeltQC video surveys to be very helpful. The BeltQC reports provide additional visual material that is used for decision making on conveyor belt replacement and planning of maintenance schedules. Minor damage to the belt can in some cases be traced to the cause of the damage. The annotated Keograms also provide an easily understood single view of the whole conveyor belt. When this information is used as a part of belt management a reduction in downtime as well as cost can be achieved.

Acknowledgements

The authors gratefully acknowledge the financial support received from Richards Bay Coal Terminal and the National Research Foundation, THRIP BeltQC3 Project 1271: GUN 2047503.

APG graduate students Jaysen Naicker, Aubrey Mhlongo and Thavashen Padayachee also made numerous contributions to the software routines and algorithm development.

References

-
- [1] Numerical Recipes W. Press, B. Flannery, S. Teukolsky and W Vetterling, Cambridge University Press, 1990.
 - [2] AK Jain, Fundamentals of Digital Image Processing. Prentice Hall, 1989.
 - [3] C. Torrence and G Compo "A practical Guide to Wavelet Analysis" Bull. American Meteorological Society, V 79 1, p61-78 (1998) <http://paos.colorado.edu/research/wavelets/>
 - [4] I. Daubechies, "Ten Lectures on Wavelets", Society for Industrial and Applied Mathematics, 357pp (1992).
 - [5] JC Principe NR Euliano WC Lefebvre "Neural and Adaptive Systems: Fundamentals Through Simulations" John Wiley and Sons, ISBN: 0471351679, (2000).
 - [6] RSA patent: Application number 2000/4739 Monitoring moving objects M. Alport, S. Plumb.