Using a micro to process 30 line Baird television recordings

Early television recordings on gramophone records – Phonovision – were crude in the extreme. The author describes a method for improving picture quality by correlation and digital filtering

by D. F. McLean
B.Sc. (Hons)

In the late 1920's, J. L. Baird performed some experiments on the recording of television pictures onto wax discs.* This he called 'Phonovision' and for a time caught the imagination of the prospective viewing public with this and other television-related inventions. Surprisingly few of these early recordings are still in existence.

It is hoped that this article will allow people to 'look back' to those early television pictures and will show that the old and the new technologies can be brought together by anyone having access to tape copies of the recordings and a personal computer.

The requirements for the computer are not strict. A minimum specification would include sufficient memory for a long sequence of frames, some sort of graphics capability allowing the pictures to be displayed with a few grey levels, an analogue-to-digital converter and a sampling clock for the converter and the computer. In my case, there is enough memory for 32 frames at less than 1Kbyte per frame and a converter capable of sampling at 15kHz to 8bits of accuracy (256 levels of voltage) under control of the computer and the sampling clock. For more detailed pictures either the sampling rate can be increased or the playback speed of the recording decreased.

Although the author had known about mechanical television for some time, it was only comparatively recently that examples were first heard on a BBC documentary record. Out of interest, I decided to display the sounds on this record as images, using a computer, as it was able to store the pictures as a sequence of samples. These pictures could be 'replayed' over and over again to check for movement, features and details. The replay was viewed on a graphics display, but an oscilloscope with control of X,Y,Z modulation by the computer would have been just as good.

It was clear from the start of these experiments that there were no synchronization pulses for identification of the start of lines and frames: the frames appeared to roll and drift in position due to playback speed variations. Synchronization of the early disc recordings was obtained by having the record platter rotation directly linked through a gearing arrangement to the scanning apparatus – figs. 4 and 5 show this arrangement clearly. The more common recordings of the mid 1930's were not linked in this way and relied on the record platter inertia to reduce picture 'hunting' or slippage.

If the original synchronous recordings had been available for these experiments a sampling clock for the computer could have been derived from the rotation of the record player, to ensure synchronization independent of playback speed. In their absence, I have evolved a method for re-aligning the sequences of pictures and inserting new synchronizing pulses, in an attempt to get nearer to re-creating the original scene quality.

30 line Baird standard

In a similar fashion to broadcast television today, the 30 line picture was created by scanning a spot of light of varying brightness in a particular pattern to form the display area. To re-create the scene as recorded, the spot had to follow this raster pattern exactly in synchronism with the video signal. If exact synchronization was not maintained, the picture would roll or slip in a similar fashion to an out-of-adjustment 'vertical hold' control on a modern tv receiver. Modern tv standards include provision for sync. pulses to 'tell' the receiver where the start of line and frame is: hence picture slippage is rarely a problem. A form of sync. on 30 line transmissions was obtained from a mixture of the inertia of the scanning disc and the actual scene content (as the television waveform was used to control the disc's rotational speed). Synchronizing the transmitter and receiver to mains frequency was only successful within the area served by a particular generator.

The scanning action on Baird 'Televisor' types of receiver was performed by a rapidly spinning disc which had a spiral pattern of holes spaced at equal angles around

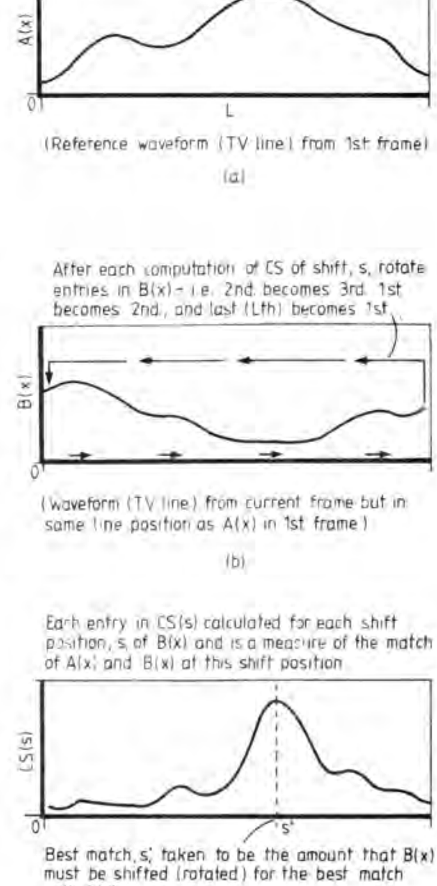

***Fig. 1.** Line matching. First two waveforms are reference and line to be matched by shifting. Samples of A multiplied by B samples produce 'score' at (c).*

*See references 1-6 for details of 'Phonovision'.

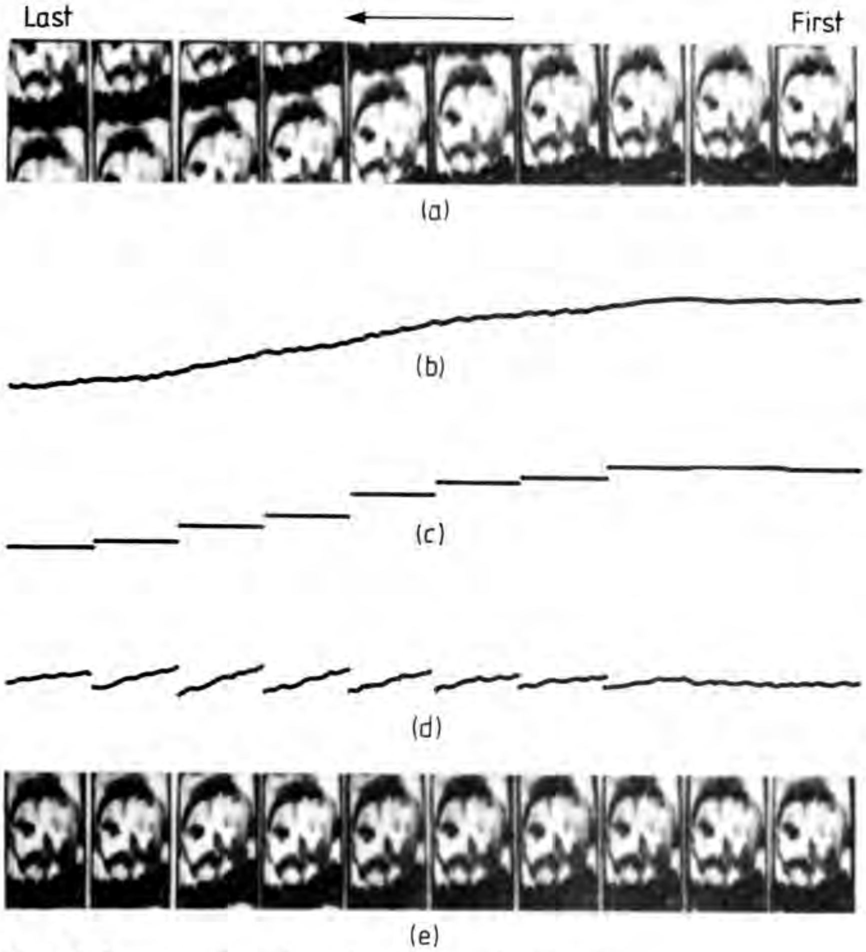

Fig. 2. Performance of one-dimensional line-matching system.

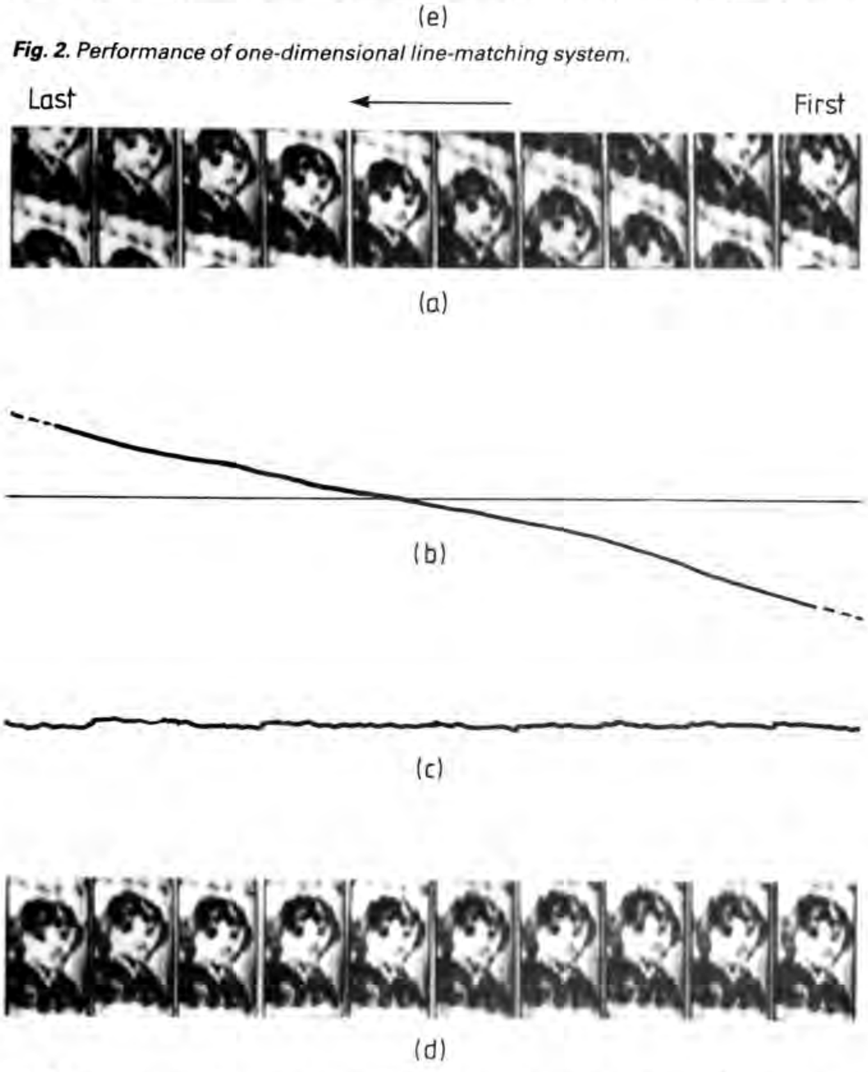

Fig. 3. Linear interpolation between adjacent frames instead of single-value shift of Fig. 2 (c). Picture tilt in each frame of (a) and removed at (d).

the disc, to give a small display area with a height-to-width ratio of 7:3. The frame or picture repetition rate was 12½Hz, each frame being made up of 30 lines. The spot of light that built up the raster scanned vertically from bottom to top to build up one line, each new line being placed slightly to the left of the previous line until a total of 30 had been scanned. The spot then returned to the position of the first line to start the next frame.

Correction of picture drift

In many applications of signal processing, the correlation or matching technique has grown in use throughout the years to become today a very powerful tool. Its main ability is, given two signals, to calculate a value whose magnitude indicates how similar the signals are. If one of the signals is delayed or shifted with respect to the other, repeated application of this matching technique can indicate how much one signal has to be shifted to match the other.

Variations of this technique were applied to short sampled extracts of recordings of early mechanical television pictures stored in computer. The aim was to find a method of accurately re-aligning a free-running sequence of frames for viewing and further processing.

Figure 2(a) shows a typical sequence of 10 frames digitized and stored in the computer memory: the first frame is on the right and all subsequent frames are to the left. The nature of the drift in frame position is quite evident. In the short space of time represented, the left and right edges of each frames have not drifted detectably, but, the images suffer from severe vertical drift in the position of line start and end (top and bottom), caused by wow and flutter in the recording medium of between 1 and 2%. The extremes of this variation would be equivalent to the difference between an image being perfectly level and one that is tilted by about 60°, corresponding to a change in the line start position from beginning to the end of a frame of about ⅔ line length. Figure 2(b) shows an estimate of the line-by-line positional error.

Also of importance is the frequency spectrum of this playback speed fluctuation. Figure 2(b) shows that fast fluctuations in speed are of much smaller amplitude than the slower frame-to-frame variations. The difficulty lies in obtaining correction methods able to cancel out *all* of these variations.

Method

Line matching. Figure 1 shows two waveforms, A and B. A is considered to be the reference and B is to be shifted to find the best match. As these waveforms represent two tv lines, the starts and ends of lines define limits. For this method to work, line B is assumed to be periodic, so that when shifted in one direction, the last sample in the line wraps around to become the first sample in the shifted line. Thus, the shifting appears to be a rotation. Waveform B is rotated a sample at a time. For each rotation, a matching score is calcu-

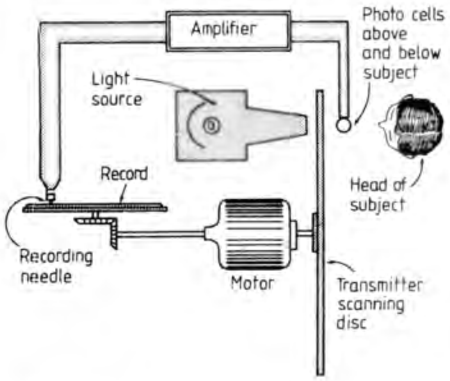

Fig. 4. Mechanical gearing of original apparatus provided steady sync.

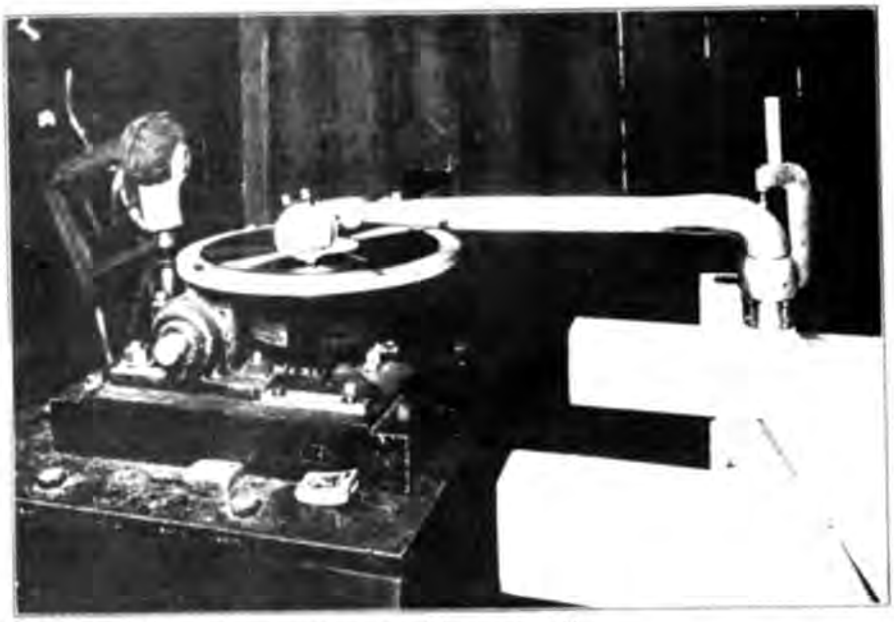

Fig. 5. System of Fig. 4 in use, recording the 'dummy head.'

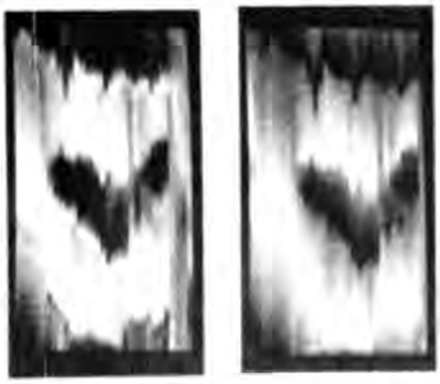

Fig. 6. Phase correction by all-pass filter.

Fig. 7. Digital filtering was carried out within the computer.

Fig. 8. Further examples of digital filtering to reduce effect of head-cutter resonance.

Fig. 9. Digital filter also used to reduce effect of low-frequency noise.

lated by multiplying each sample in the rotated line B by the corresponding sample in line A. The sum of these products is the matching score. The process is repeated by rotating line B and calculating another score for this new shift position, s. The equation below describes this sum of products.

$$CS(s)=\sum_{x=1}^{L} A(x).B(x-s)$$

where CS is the matching score, x is sample position from start of line, s is the current shift value in samples and L is the number of samples in the line.

The matching scores are stored in a list so that, at the end of one complete rotation of line B back to its starting position, the position in the list of the maximum score can be taken to be the number of samples by which B must be rotated to be 'lined-up' with A.

For each line in the current frame, the current line was matched against the corresponding line in the reference frame. The position of maximum score for the above equation was taken to be the value by which the current line had to be shifted to match best the line in the reference image. Initially the reference frame corresponded to the first frame in the sequence. After each line was corrected, the line in the reference frame was averaged with a fraction of the current line to take into account any scene change at the horizontal position of the line. The average was stored back in the reference frame.

The results of one-dimensional matching were somewhat mixed in success. For clear, stable and simple scenes, the results were excellent, leaving an extremely stable sequence of rectified frames with very low jitter in vertical position from one frame to the next. However, for fast-moving complex scenes, the performance was poor with unstable breaking up of the picture structure and consequent severe degradation in image quality.

Line-jitter removal. One of the recordings suffered from severe timebase jitter, causing large changes in the position of the start of the line on subsequent lines within a frame. In this case only it was considered worthwhile to use a different form of one-dimensional matching.

Instead of matching the current line in the frame being processed with the same line in the reference frame, the current line was matched against the previous line in the *same* frame. I have used this technique successfully with slow-scan television pictures received on the amateur bands[8]. Line-to-line jitter is removed in a fairly uncontrolled way at the expense of geometrical distortion of the picture. As it was considered important to maintain the picture geometry, this technique was not used on the other recordings.

The one recording which was processed using this technique suffered from static errors in the position of the start of each line in any frame. This was presumably caused by errors in the position of the holes in the scanning disc, causing some lines to start earlier or later than adjacent lines. Figure 11 shows a typical frame before and after correction of this fault.

Frame matching. The problem with the line matching technique was that it could only allow a small amount of lateral movement in the scene before instability degraded the picture quality. Because the matching algorithm was essentially one-dimensional in nature, it could not cope with any sideways movement. The algorithm had no 'knowledge' of any structure in adjacent lines. To attack this problem, a variation on the original method was devised and tested with excellent results.

This new method was based on the fact that the scene content (not position) varied little from frame to frame. There was no abrupt scene changes. Each frame can be thought of as a two-dimensional brightness distribution where each point (x,y) has an associated brightness value B(x,y). Using one frame as a reference, the idea was to 'slide' the frame to be corrected horizontally (in x) and vertically (in y) until a best match was found. The equation for calculating the matching score at any shift value (s,t) was derived from the one-dimensional equation given earlier and is given below

$$CS(s,t)=\sum_{x=1}^{L}\sum_{y=1}^{M}A(x,y).B(x-s,y-t)$$

where A(x,y) is the brightness in the reference frame at point (x,y), B() is the brightness in the current frame being processed at the shifted (x,y), L,M are the max. no. of samples and lines in x and y respectively, and CS (s,t) is the score for a possible match at shift value (s,t).

The 'sliding' of the frames is similar to the cyclic shift used when matching lines (Fig. 1) but is extended to two dimensions. A cyclic shift of the current frame was performed for each shifted position (s,t), and all possible shifted positions of one frame with respect to the other was used to create a list of scores for matching.

Using the position of maximum score, the current frame was cyclically shifted from its original position in both x and y directions to match it with the reference frame. Each point of the current and reference frames was then averaged with a user-selected weighted value. The averaged result was stored to become a new reference frame for the next frame in the sequence, which served to accommodate increasing differences between the successive and reference frames.

The positional errors that had to be corrected were only in the vertical direction, however. During the time taken for a short sequence, the horizontal drift could be ignored as it was 1/30th of the amplitude of the vertical drift in this case. Correction only in the vertical direction was achieved simply by removing the calculated lateral (x) shift *after* the full two-dimensional matching and shifting had been applied and the reference frame updated with the weighted average. The averaging of the reference frame with a fraction of the current frame had to be done by using the current frame shifted in *both* horizontal and vertical directions to track lateral movement in the scene.

The effect of using two-dimensional processing but only correcting vertically was to allow horizontal motion without any loss of stability in the processed sequence of frames. It is indeed most fortunate that the line scanning direction of the early mechanical television experiments was vertical as most scene motion in natural objects is typically only from side to side.

Figures 2(a) to 2(e) show the performance of the basic two-dimensional correction method described in the previous subsection. Figure 2(a) is the original digitized sequence showing significant vertical rolling of the frames in the sequence. Figure 2(b) shows an estimate of the actual positional error in the vertical direction on a line by line basis. Figure 2(c) displays the actual calculated result of how much each line in each frame had to be shifted. Each line in a particular frame was shifted by the same amount so that the correction function (Fig. 2(c)) shows only a stepped approximation to the actual positional error. Figure 2(d) is the residual error after applying the offset of Fig. 2(c) and Fig. 2(e) is the corresponding corrected sequence of frames.

By comparing Figures 2(b) and 2(c), a closer approximation to the actual positional drift could be made by performing a linear interpolation between the vertical shift values of adjacent frames. The method was quite simple to implement and provided considerable improvement to frames distorted by sampling at a slightly incorrect sampling rate.

Figures 3(a) to 3(d) show the result of processing a sequence digitized at slightly too high a sampling rate. The result here is equivalent to a 0.7% increase in the desired sampling rate. Figure 3(b) shows the linearly-interpolated error derived from the vertical shift calculated using the matching equation: the error wraps around as the shift value has cyclic symmetry. Applying this interpolated set of values (Fig. 3(b)) to the original sequence gave rise to the processed sequence as shown in Fig. 3(d). Figure 3(c) shows the residual error. The most significant result from this example is the removal of the overall tilt from each frame.

A natural extension of the original line matching idea resulted in a stable and robust method of accurately re-aligning a sequence of tv frames with no further information than the scene being transmitted. The restriction of only applying vertical correction resulted in the

Fig. 10. Good example of removal of hum by digital filter.

Fig. 11. Correction of variation of position of line starts within a frame.

Fig. 12. Set of pictures produced by system from noisy and distorted taped recordings. Pictures are still crude, but are enormous improvement on untreated versions.

preservation of horizontal motion with no loss in stability. Further correction for speed variations proved to be a powerful method for removing the 'tilted' effect on certain frames. This required little computational overhead.

Although not able to remove fluctuations faster than the frame rate, the frame matching technique has proved to be a useful tool in assisting the analysis of these early mechanical recordings.

Image quality

The quality of recording on wax discs in the 1920's was adequate for voice or music reproduction although it was far from hi-fidelity. When used for recording Baird's 30 line television signal, wax discs proved themselves to be quite a poor recording medium. The recording apparatus was not capable of recording the very high or the very low frequencies in the signal, and yet the shape of the recorded waveform – much less important for voice or music recording – had to be free of distortion for accurate reproduction of the scene being televised.

The limitations of base-band (unmodulated) recording on wax discs resulted in various types of distortion of the television waveform. The most common types observed on the discs of the period are: phase distortion – poor low frequency response, giving rise to phase shifts; low-frequency noise – eg, main hum aggravated by reduced signal level at low frequencies; high-frequency instability generated by head cutter resonance; noise caused by disc surface granularity; and residual timebase errors, giving rise to ragged edges to each frame.

Image filtering

One-dimensional filtering. Most types of distortion present on these recordings can be reduced by one-dimensional digital filtering techniques. This means that the television signal is treated as if it were an audio signal: the relationship between lines and frames is not used in the processing.

Phase distortion can be reduced by processing the signal through an all-pass phase shifter. It is the author's experience that using a simple electronic circuit to perform this function relieves the computer from time-consuming processing and gives instant feedback on the correction being applied to the signal. Figure 6 shows the result of phase correction.

Head-cutter resonance is predominant on one particular recording. Although external pre-filtering can reduce the effect of the resonance, digital filtering within the computer was found to be much more flexible and was able to reduce the resonance without adversely affecting the resolution of the picture. Figures 7 and 8 show the reduction possible by digital filtering.

Low-frequency noise was more effectively reduced again by using digital filtering. In one of the recordings, mains hum was present at a high level after attempting to recover low frequency information in the signal. The reduction of interfering signals on early Baird recordings is demonstrated well on Figs 9 and 10.

Although the waveform is by nature one-dimensional, the correlation between lines and frames can be used to suppress some types of irregularities in the signal. One particularly powerful technique involves processing the points in the scene exactly one frame apart, a process called temporal filtering.

Temporal filtering. One of the most effective processes on the frame- or line-matched sequence of frames was the temporal filter. Both surface noise and residual errors in the position of the lines were considerably reduced without affecting detail in the individual frames significantly.

The idea is that points having the same position along the same line on adjacent frames should be very close in their value of brightness. The filter creates a new point from the brightness values of the same point on three successive frames, chosen to be the middle value in brightness amongst the three. Using this median value allowed isolated errors to be corrected completely: line jitter, noise and to a lesser extent movement were suppressed without the blurring action of a spatial filter (i.e. one acting along or across lines). Reference 9 describes this technique applied to high resolution television.

Temporal filtering is both difficult and expensive to implement in high-resolution television because of the large amount of high speed memory required to store two (or more) frames. The memory and speed requirements for 30 line television however makes it very much simpler to demonstrate powerful image processing techniques such as this.

The author

After graduating from Glasgow University, the author spent several years with EMI Central Research Laboratories before joining Logica as an image-processing consultant. With a strong grounding in television, he has worked on digital video and audio systems, automatic inspection methods and, more recently, on the image-processing system for a fingerprint identification system. In his spare time, he explores aspects of television techniques.

Software

A program for acquiring data, displaying the result, re-aligning and processing a digitized sequence of 30 line television was written in machine code for acquisition and re-transmission: the processing routines could have been written in any language. A decisive factor in using machine code for the matching algorithms was the vast number of calculations required to match-up and re-align the frames. For line matching, this included about one million multiply operations for a sequence of 32 frames. Matching up 32 frames took 150 seconds in Z80 machine code with the processor running at 4MHz. Performing the multiply operation in hardware reduced the execution time to 65 seconds for a 32 frame sequence.

The implementation of the two-dimensional frame matching took considerably longer to execute than that of the one-dimensional line matching, since frame matching needed a greater number of multiply and accumulate operations (30 times) plus higher precision in the score value for matching. Considering that an image or frame in this case had 960 samples, point-by-point multiplication and score accumulation gave the staggering figure of just under one million of these operations per frame. The software implementation took 3 minutes per frame – 96 minutes in total for a complete 32 frame sequence. Performing just the multiply operation in hardware reduced this execution time to 80 seconds per frame – 40 minutes per sequence. A similar routine in Basic was estimated to take about 75 hours per sequence, while a compiled PASCAL routine on the same machine took 17 hours per sequence.

The one-dimensional and temporal filtering all required between 2 and 10 seconds to process a 32 frame sequence of 960bytes per frame. For high-resolution processing with 1920 bytes per frame, fewer frames could be held in memory. Because the amount of data was similar in both the low and high-resolution sequences, the processing time for each was similar. Fourier analysis of the signals to determine which filter to use took a few minutes for each frame.

Acknowledgements

Thanks go to Doug Pitt of the Narrow Band TV Association, Ronald Gibb of Strathclyde University AV unit, Ray Herbert, Ben Clapp, H. Spencer, Len Firmin MBS, John Ive and Mike Hallett of the IBA and Tim Voore for assistance in researches into mechanical television.

References

'Phonovision'

1. Dinsdale, A. 'Television' 2nd edn. Television Press Ltd. 1928, pp 142-145.
2. Moseley & Barton Chapple 'Television – Today and Tomorrow' 2nd edn. Pitman & Sons 1931, pp 127-130.
3. Newnes 'Television and Short-Wave Handbook' 1st edn., 1934, pp 47-50.
4. *Television and Short-Wave World* June 1935, p308 and p363.
5. Patent no. 289104, 16-APR-29, J. L. Baird.
6. Matthewson, D.K. *Television*, May 1983, pp 361-2.

For details of crrelation techniques

7. Gonzalez & Wintz 'Digital Image Processing' Addison-Wesley, 1977, pp 67-69 and pp 383-386.

Further application of correlation:

8. *Rad-Com.* Letter on computer-controlled slow-scan TV. October 1982, pp872.

Details of temporal filtering:

9. Huang, T.S. 'Image Sequence Analysis' Springer-Verlag, 1981, Chapter 4.

Short extracts from one of the recordings can be heard along with interesting conmmentary on the following:

10. 'We seem to have lost the Picture,' BBC documentary l.p. REB 239 from the series '40 Years of TV,' 1977.