Video Supervision and Documentation of Tear-Off Situations in Web-Inspection

Matthias F. Carlsohn[*a], Friedrich Aberle[b]

[a]Ingenieurberatung Dr. Carlsohn f. Computer Vision & Bildkommunikation,
D-28355 Bremen, Germany
[b] Sennheiser electronic GmbH & Co KG, D-30900 Wedemark, Germany

ABSTRACT

In web production a continuous work flow is very important, specially in paper-mills that produce high quality photo papers. Often rupturing of the paper-web cause falling-off for hours. Knowledge about the reasons for this kind of failures helps to tune the production parameters appropriately to avoid interrupts and enlarge the mtbf coherently.
From experience the operators know critical positions within the production line that they try to observe for deviations from the proper conditions.
Production speed, dimension of the production line and the positions from where the best view onto the web is given make continuous human visual inspection impossible.
Video cameras integrated within the production line substitute for human visual inspection. Continuously grabbed image sequences are buffered into a distributed short term memory for all video channels, simultaneously. This occurs synchronized upon trigger from light barriers checking the presence of the web at particular positions.
Positioning of up to 16 cameras is possible. Operators usually know those locations where tear-off can occur. It is unpredictable what time exactly has to be recorded to display successfully at multiple positions the events responsible for production interrupt. Therefore the marker for the trigger event is adjustable within a recording time window of 30s.
The solution is portable to other applications in web-inspection where for fast processes short clips of many coherent video streams have to be recorded due to particular trigger events.

Keywords: video surveillance, CCTV, multi-channel real time video recording, web-inspection, tear-off, video archiving, paper production, distributed systems

1. INTRODUCTION

Recordings of long video sequences are still the domain of video tapes because of price and availability of media.[1,2.] In cases where from 24 hours recording the interesting or critical events will late some few seconds only within a couple of days no sophisticated method of systematically browsing the material for the important instances exists. Exhaustive search would be needed. Consequently, large storage archives of video tapes contain often close to 100% useless information. This is too expensive concerning storage amount, investigation time for browsing and relevant information yield.

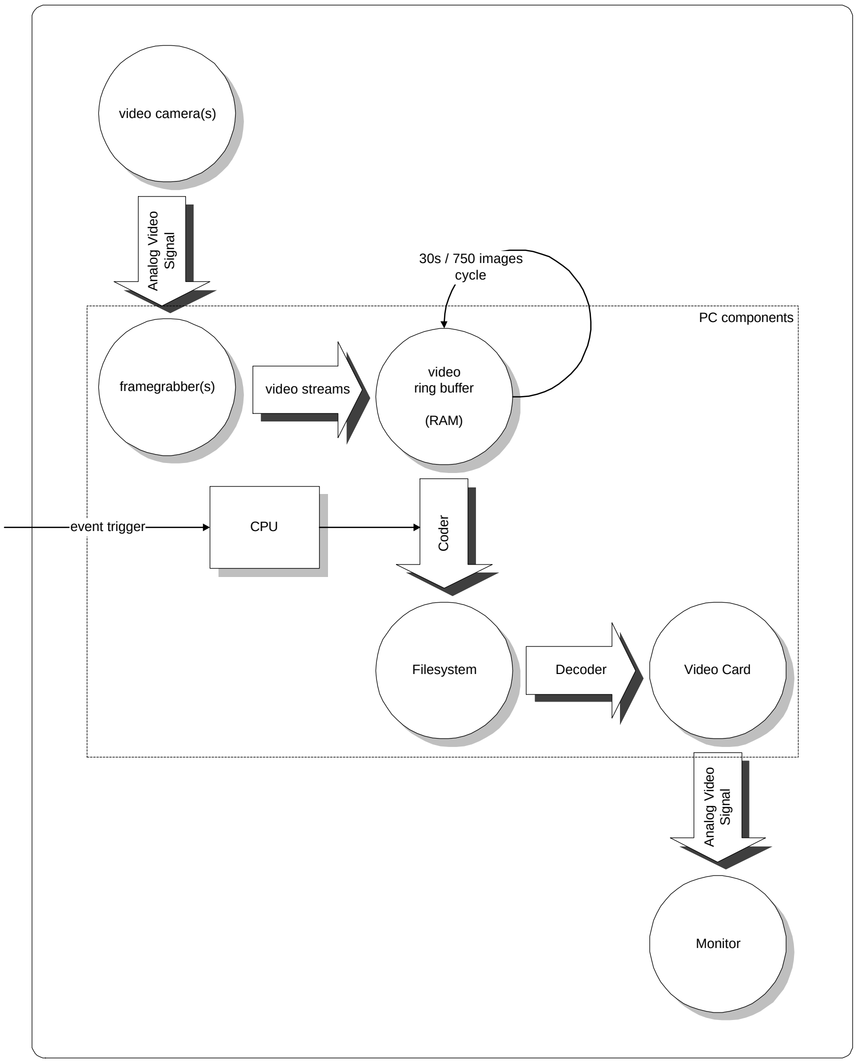
Reduction of the recorded data to an extract of exactly that portion of video information that contains the important time intervals for a particular application is necessary.
Determination and selection of semantic importance of information involving appropriate sensors are needed instead of mathematical methods for compression in the sense of classical information science. This measure avoids a huge amount of unimportant information waste.

[*] Correspondence: Email:Matthias.Carlsohn@t-online.de; Telephone:++49 421 2052055, Fax: ++49 2052056

Therefore mass of information has to be buffered into a short time memory for later access for permanent archiving. Selection of what should be finally stored has to be automated with respect to the application. Of course additional information, e.g. time and date stamps are required as well as classical image data compression for long term storage.

Concerning price and availability the solution should be as close as possible to video tape recording.

2. SYSTEM CONCEPT



Actually, PCs are available everywhere for reasonable price. Performance in data throughput and memory capability for main memory rams and disk space are continuously increasing. Dedicated hardware, for example frame grabbers become more and more standard and inexpensive, simultaneously.

These insights lead to the idea to design a solution that makes use from standard PC components as far as possible. Only the application is added by software plus some simple intercommunication device to set up a distributed system that can be scaled to the number of video sources required (figure 1).

The paper discusses the general system architecture, system operation and data management in a case study for a particular web-production application.

Figure 1: Video Supervision System (VSS) general architecture of a single PC configuration

2.1 System requirements and architecture

The PC contains a processor of at least 450 MHz and a main memory of 500 MB to 1.500 MB depending on the number of video sources and desired recording time. Ethernet LAN-cards ensure general intercommunication among the distributed network. Industrial standard frame grabbers are responsible for acquisition of video data streams where up to 4 can be plugged into a single PC without restrictions when monochrome information is exploited. For monochrome cameras that can be externally synchronized also RGB grabber can be involved each of them fed from 3 monochrome video sources.

The application software consists of one or more recorder(s) and a player. The recorder is responsible for control of frame grabbers, administration of main memory and permanent archiving of video streams onto disk upon command either initiated automatically by a sensor or manually by an operator.

Recorders occur multiple times within a VSS system. They are distributed over the number of PC stations involved in the entire system. The player is able to select and display the stored video streams from different PC stations, all members of the VSS. The only interface among recorder(s) and player is the file system of the operating system.

2.2 Case Study: Tear-Off Situations in a Paper Mill

Recently, web-inspection became another standard application in automatically visual inspection for quality control.[3.] If wallpaper prints, plastic foils, textile or wooden web material is checked, the goal is almost the assurance of a homogeneous material output. No artifacts should reduce the quality of the product like inclusions, bubbles, color deviations, holes, etc..

Beside these demands given by the quality requirements of the product there are additional needs given by the production process. In web production a continuous work flow is very important. In a paper-mill for high quality photo papers a web of 4m width and 400m length is transported by a speed of approximately 500 m/s along a production line of nearly 100 meter length. The web itself is much longer because of the many diversions by return sheaves. At the beginning of the production line the material consists of approximately 99% water and 1% paper pulp. At the end the relation is nearly vice versa. All the way in between one has the different mixtures of the two main components winded as a web like a large conveyer belt around rolling-mills.

Interrupts in this production scenario by rupturing of the paper-web cause falling-off from half an hour up to eight hours. Therefore it is very important to know the reasons that caused the failures to tune the production parameters appropriately that avoid interrupts as good as possible and enlarge the mtbf coherently.

From experience the operators know some critical positions within the production line that they try to observe for deviations from the proper conditions.

Because of the production speed, the large dimension of the production line and the positions from where the best view onto the web is given human visual inspection is obsolescent.

Surveillance by video cameras along the production line substitutes for human visual inspection (figure 2). Obviously, it would be useless to get 24 hours the day images and nothing happens.

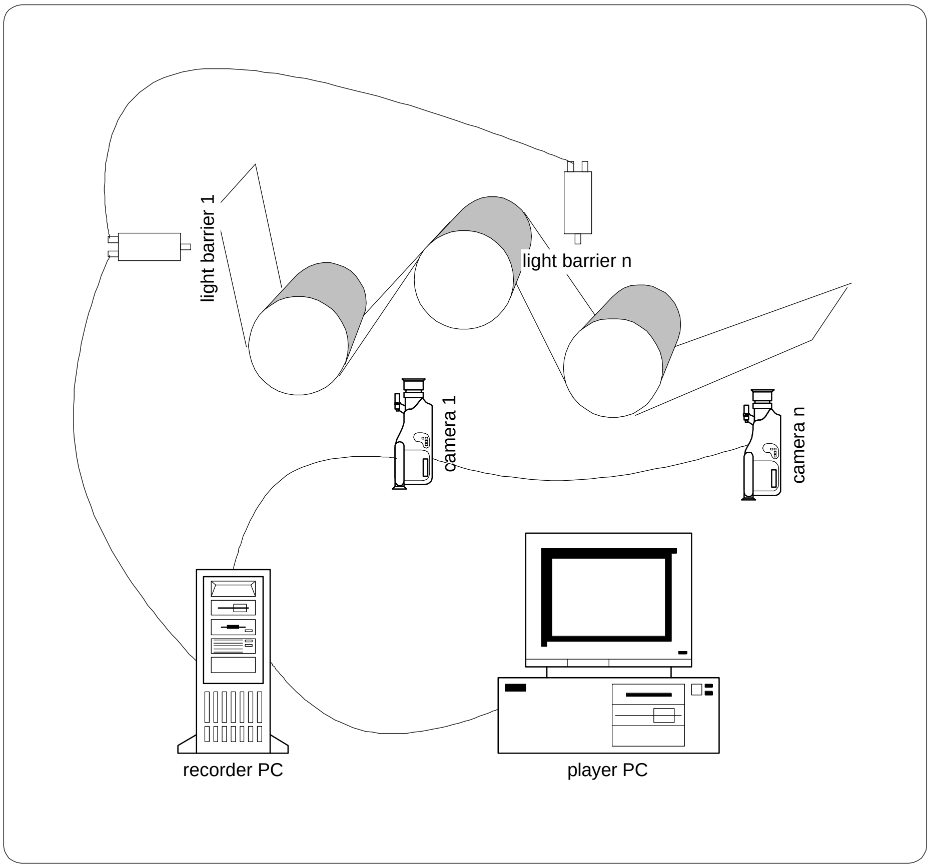


Figure 2: Web-inspection and tear-off control within a paper-mill

For that reason at particular positions light barriers check continuously if the web is present. When the paper is missing suddenly, a trigger signal starts a procedure for permanent archiving of the continuously grabbed image sequences for all video channels, simultaneously (ref. figure 3).

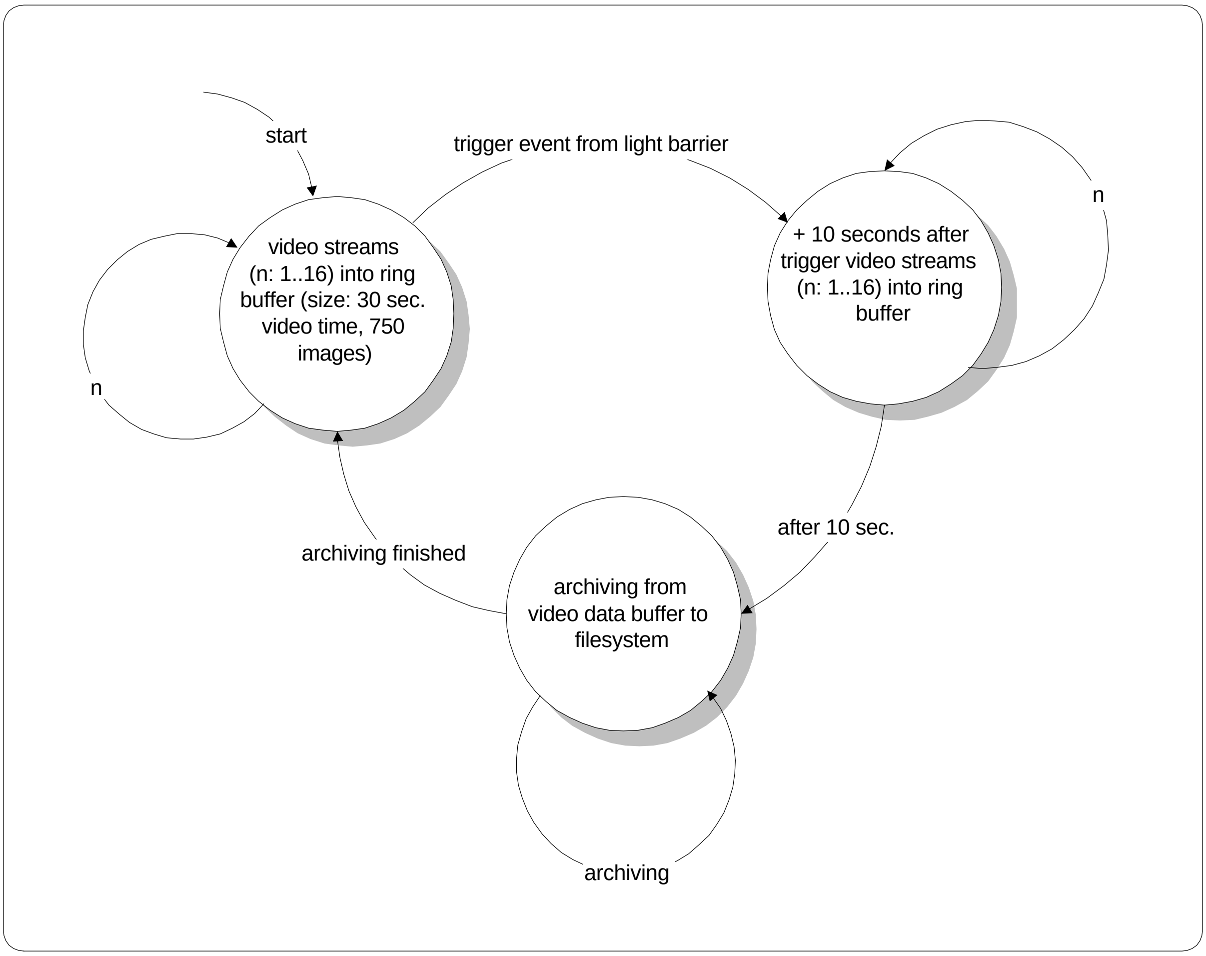


Figure 3: State diagram of the image sequence buffering and archiving cycle

After initial start all frame grabbers within the distributed system in figure 4 begin to acquire images that are continuously stored into a FIFO buffer showed in figure 5.

Up to four PC stations placed along the production line allow to grab up to 16 independent video channels at the same time. Images have a resolution of 640x480 pixels. These are coded by 8 bit for gray value representation. For a video clip of 30s real time one need approximately 230 MB ram n-times for n video channels distributed over the main memory of a couple of PCs (see figure 4). The buffer in figure 5 contains at each instant a constant number of recently recorded images. For each video channel the trigger event can be combined individually with one particular image of every sequence, e.g. each channel can have its own origin from which locally visual history and future can be recorded and displayed. This is a very flexible concept. In the real case of the paper-mill application the assignment of trigger event to the "origin" image is equal for all channels. This particular image is the reference or origin of the time window and is labeled by "0", consequently (figure 5). With respect to the trigger instant and its connected origin of the image sequence "future" images are grabbed because, e.g. 10s after the trigger event FIFO buffering stops and archiving takes place. This is shown in figures 3 and 5. Now, the data are transferred from the short term buffers (FIFO) distributed over the main memory of the PCs involved (cf. figure 4) into local files headed by additional information. These are the trigger image, date and time, number of images, recorder status (master/slave), sequence of fields (even /odd), time between frames (40 ms).

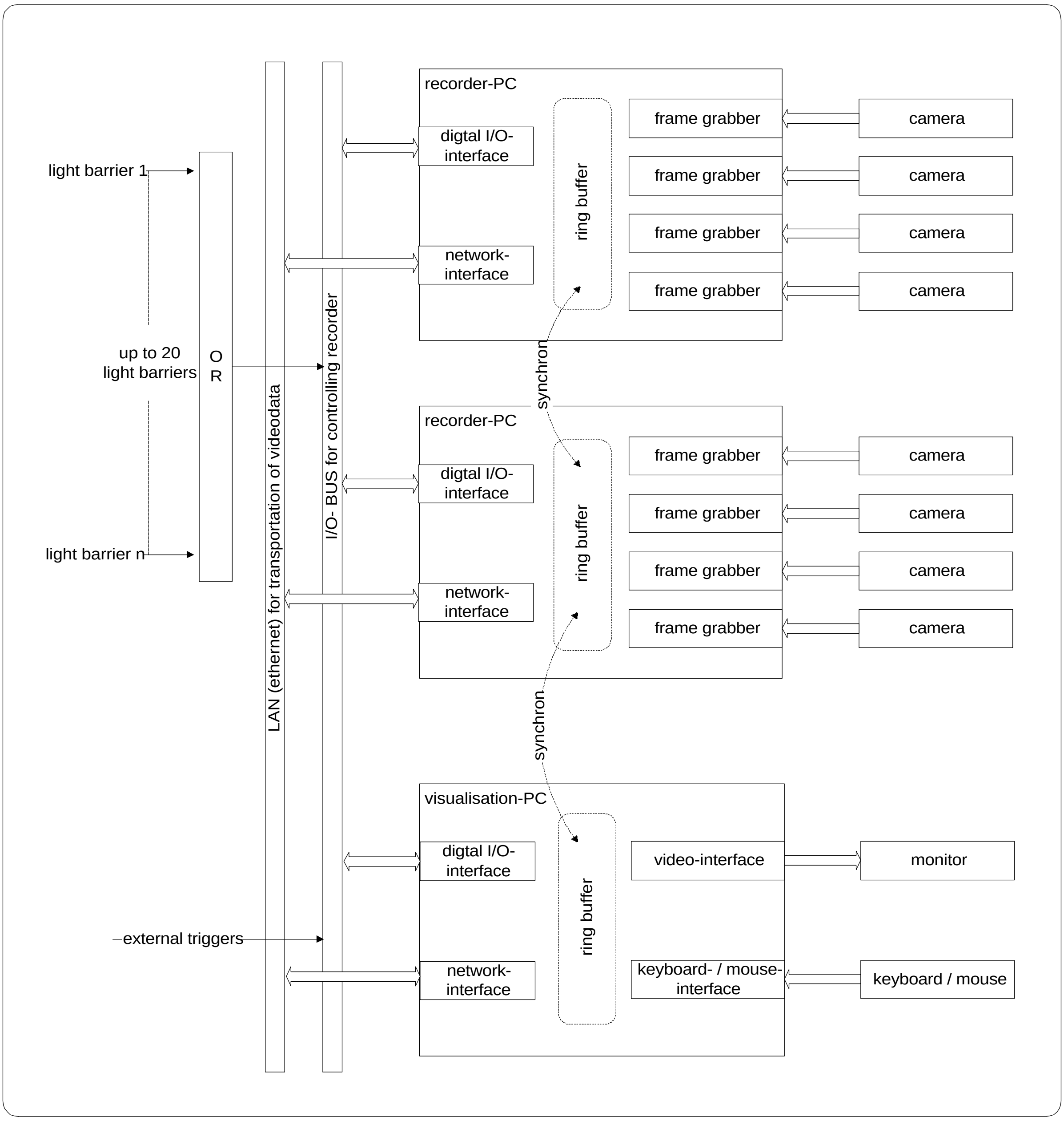


Figure 4: Distributed system architecture of VSS involving multiple PC stations

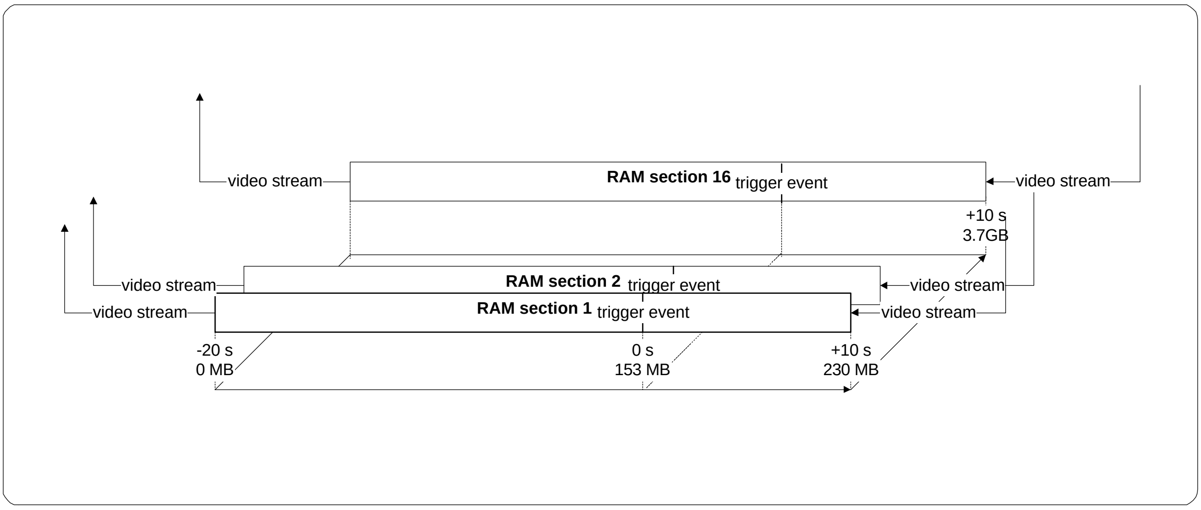


Figure 5: Distributed video sequence short term buffer (FIFO principle)

Files are stored into pre-defined local directories configured for all recorders individually. The recorder is able to recognize even and odd fields for proper interlaced and also for non-interlaced representation of video. The player recombines the locally distributed files to a compound of video streams. Fidelity concerning alignment of even and odd fields of all particular trigger images (origins) is maintained.

For archiving purpose the coherent files include the video sequence in three resolutions (full, quarter, 1/16) by sub-sampling image data by a factor ½ and ¼ in both directions, simultaneously.

One of the PC stations within the VSS has to be defined as master recorder. All the others are slaves. The master transmits the code for the actual file names represented as numbers to all slaves for the next recording event. Additionally, software (initiated by the operator) and hardware triggers (caused by the sensors) are forwarded via the I/O links of the line printer interfaces (cf. figure 4).
File names begin either with a leading "tmp" that means that they can be overwritten or with a leading "_" what indicates that a certain number of files are generated before the oldest one is replaced by a new one due to available disk space. Networking among PCs allow to overcome bottlenecks of single PCs concerning their resources. A local area network (LAN) is built up by ethernet, mainly for file collection transfers for the player. Additionally, control functions are submitted making use of the parallel port interfaces. A dedicated line printer-adapter interfaces this control structure (figure 4).

Synchronization is performed via an interface using the parallel port to ensure that the recorded data are taken due to the same trigger. The solution is portable to other applications in web-inspection where for fast processes short clips of many video sources have to be recorded due to particular trigger events.

3. ARCHIVING

Permanent archiving needs some additional measures to ensure an easy and reliable access on the image information, specially within a distributed system.

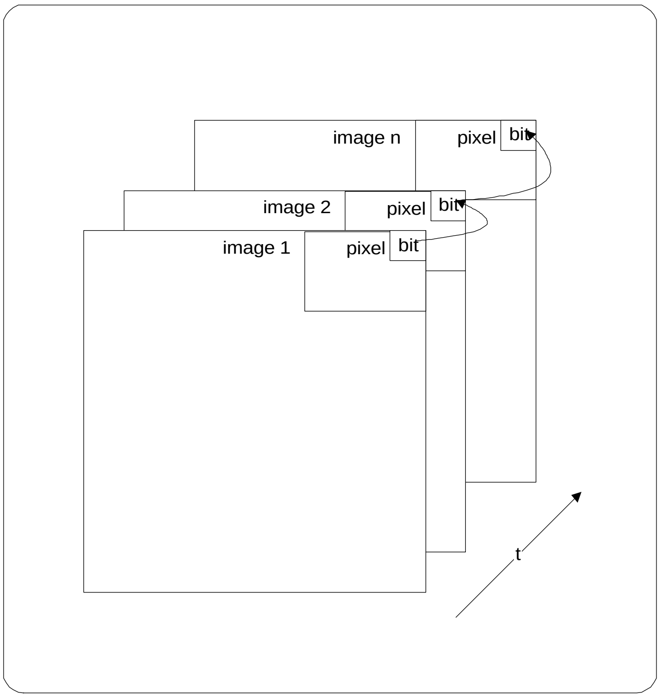
3.1 Storage of video sequence files on disk

Synchronization among all members of the VSS in figure 4 is performed via the line printer I/O-interface where the master recorder defines the enumeration for all slaves, that is the basis for the design convention of the file names. When the player is showing the video files taken from the entire system, a common file name is the unique identifier for coherent

scenes taken from different video sources independently from the location where they are stored (disk, directory paths, etc.).
Depending on the number of channels that are displayed at the same time, e.g. 1, 4 or 16, the appropriate image resolution is used. Of course, also live images can be shown in this manner but not on video speed. Up-date rate is very slow (approx. 2s). This is no limitation of the main function because the processes under inspection are so fast, that human operators are unable to recognize them in real-time not to speak about their reaction time for any control manipulations.

3.2 Modified run length encoding of difference images

Additionally, VSS provides an option for image data compression for long term archiving by a bit-wise run length encoding of difference values as depicted in figure 6.[4.] The local changes in gray values are coded in bit planes, that are the subsequent differences in luminance from the first image in the sequence to its successors. The first image is stored entirely as original. Then pixel by pixel the gray value differences between the second image and the preceding original are computed. This procedure is repeated for all following pictures of the sequence.

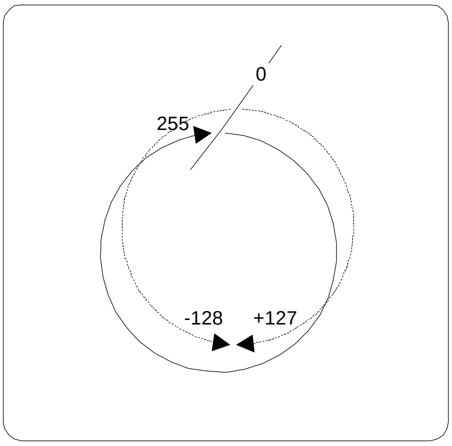


Prior to coding some pre-processing takes place because of the presence of noise in order to avoid inefficient run length encoding. For that purpose, median filtering of all images removes single noise artifacts.[5.] Otherwise singular noise pixels would cause flipping for many bits of the difference values from image to image. Noise pixels will be replaced by the median that is very similar to the locally dominant gray values. This cause more stable conditions in the temporal bit structure of the PCM coded gray value differences at each particular pixel address. Consequently, long runs of ones or zeros at particular bit positions occur for a couple of subsequent images.

Finally, the least significant bit (LSB) in an 8 bit PCM representation encodes to a certain amount the signal noise and can be neglected. This reduces gray rendition of the difference values to an amount that does not affect the fidelity of images very much and can be omitted for run length encoding. This is done because arbitrarily flipping bits cannot efficiently be coded. This would increase the number of data instead of their reduction.

Figure 6: Image data compression by temporal bit-wise run length coding along the time axis of the video sequence

Calculating the absolute gray value differences of subsequent images result integers ranging from 0 to 255. Using the code model of a number cycle, difference values can be represented by a coherent scale of unsigned integers. Alternatively, another scale of signed integers ranging from -127 to +128 express the same information as printed in figure 7. The absolute gray value difference can be described by the shortest circle arc between two gray values. Making use of the second scale at the circle one get numbers less or equal 128. Sign information can be derived from the orientation (clockwise + or counterclockwise -) of the shortest path between pairs of numbers in figure 7.



The bit planes of these differences are run length encoded and stored into files. Expected gray value differences at a particular pixel address in subsequent images are small due to the high amount of temporal and spatial autocorrelation of images when changes in the scene are rather smooth.
Statistically, difference values are often very small, i.e. close to zero. Zeros can be encoded very efficiently. But a change of a difference value from 0 to -1 in 2's complement representation would cause flipping of all bit positions. For that reason coding of binary numbers is converted due to table 1.
Depending on the gray value gradient with respect to time the differences can be positive or negative. This leads to another problem, namely that the sign bit is flipping even when the gray difference is small. Fortunately, a temporal trend over a couple of subsequent images yields sufficient long runs of positive or negative signs for efficient encoding of the most significant (MSB) bit plane. On average run length encoding achieves a compression ratio of 3:1.

Figure 7: Coding of the absolute values of the amount of gray value differences into signed integers

decimal numbers	binary numbers before conversion	binary numbers after conversion
+127 (127)	01111111	01111111
...	...	...
+1 (1)	00000001	00000001
0 (0)	00000000	00000000
-1 (255)	11111111	10000001
...	...	...
-127 (129)	10000001	11111111
-128 (128)	10000000	10000000

Table 1: Conversion of coding of signed integers

Runs are coded as number of packets of zero bytes per 8 subsequent images. Bytes containing ones are stored uncoded. The details of the entire image data compression algorithm is summarized in table 2.

```
1.    noise suppression by median filtering
2.    image 1 is stored as original
3.    for image 2 to image 750: compute difference to the preceding image
4.    coding of negative numbers from 2's complement into representation of table 1
5.    for MSB until LSB-1
      {
            for pixel 1 until pixel 307200
            {
                  for image 1 until image 750
                  {
                        write bit plane (in byte packets) of pixel address run length encoded into file
                  }
            }
      }
```

Table 2: Algorithm of data compression by run length encoding of bit planes of temporal subsequent gray value differences

4. RESULTS AND OUTLOOK

Since one year the first systems are working now. They are integrated into the production lines of different paper-mills. Apart from slight adaptations in the user interface for daily dealings the systems work reliably all the time and substituted the former video tape recording system completely. Some customer demands were added during this time to tune the system as close as possible to the daily requirements and needs in functionality of the production operators who investigate the video archives after failure situations.

Investigation of video sequences is done using the graphical user interface of the player shown in figure 9. Additionally, to the choice among different channels, resolution and speed modi a scroll bar also shown in figure 8 assists the operator in his analysis. Off-line playing of recorded video streams, forward and backward, in single image steps and in video sequences enables the operator to zoom in space and time the documented details responsible for the tear-off situation in web-production.

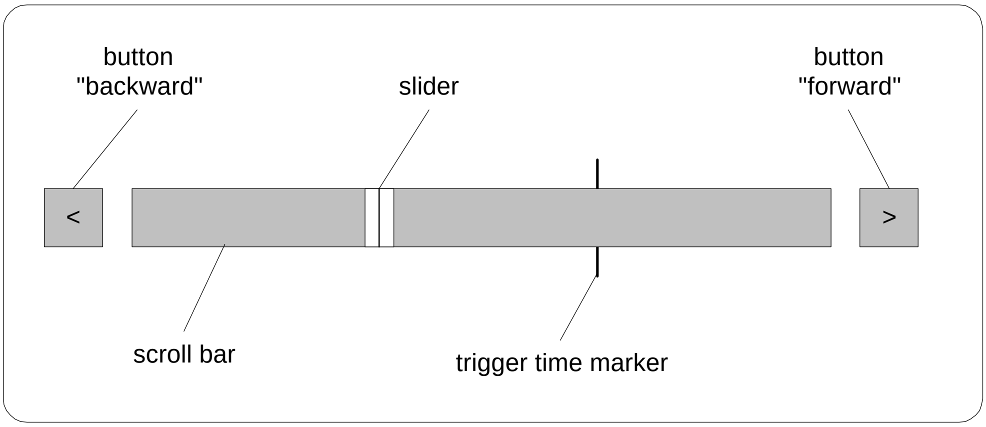


Figure 8: Scroll bar for off-line playing of recorded video streams, forward and backward, in single steps and image sequences

The main result of this work is that analog recording onto video tapes could completely replaced by digital recording using this special cctv solution of the video supervision system VSS.

Figure 9 shows a screen shot of a typical tear-off situation within a paper-mill where rupturing of the paper-web triggered the recording. The scene shows an example of the player presenting 4 channels of coherent video streams taken from different camera positions. The left hand windows show that the web disappeared already while the right hand windows show still pieces of paper and the ruptured web. This is due to the size of the production line where cameras can be far away from each other.

Depending on the sensors involved not only rupturing of the web can be the trigger for the recording process because this is a typical post mortem analysis situation. In future applications more sophisticated sensors arrangements could give early alerts indicating, for example an unstable system status. Slow motion analysis of coherent video sequences taken from different positions within the production line could help the operator during production to tune parameters for optimization of the entire process and the quality of the final product. Finally, this could avoid or at least reduce system crashes.

Of course additional sensors are thinkable that indicate not only real accidents like rupturing of the paper web but giving alerts also during other deviations from standard operation conditions. This would give the operators the possibility of interaction with the process upon video information taken during production instead of a post mortem analysis after failures.

Figure 9: VSS player showing 4 video channels in a tear-off situation within a paper production line

REFERENCES

1. A. Wege, *Video-Überwachungstechnik*, Hüthig GmbH, Heidelberg, 2000,
2. E. Götz-Meyn, W. Neumann, *Grundlagen der Video- und Videoaufzeichnungstechnik*, pp.12, Hüthig GmbH, Heidelberg, 1998,
3. WW-Elektronik, *Web-Eye 200*, Manual (in German) Version 1.5
4. M. Rabbani, P.W. Jones, *Digital Image Compression Techniques*, pp.51-54, SPIE, Vol. TT7, Bellingham, 1991
5. A.K. Jain, Fundamentals of Digital Image Processing,pp. 246-249, Prentice Hall, Inc. Englewood Cliffs, 1989