Nationaal Lucht- en Ruimtevaartlaboratorium
National Aerospace Laboratory NLR

NLR-TP-2003-407

Off-line synchronisation of measurements based on a common pseudo-random binary signal

P.J. Hoogeboom

This report is based on an article in Behavior Research Methods, Instruments & Computers (BRMIC), August 2003 by Phychonomic Society.

Customer: National Aerospace Laboratory NLR
Working Plan number: V.1.C.4
Owner: National Aerospace Laboratory NLR
Division: Flight
Distribution: Unlimited
Classification title: Unclassified
August 2003

Approved by author:	Approved by project manager:	Approved by project managing department:
24/9	24/9	

Summary

Post-processing (or off-line) synchronisation methods rely on some overlap of information between two or more simultaneous measurements. The methods vary in the amount of explicit data needed to record the overlap: on one hand the full recording of an (absolute) time source for each measurement or, on the other hand, the reliance on information overlap of different parameters in the measurements. Whenever it is impossible to record the time from a central clock in all simultaneous measurements, it is often possible to record a simple common signal. Using a binary Pseudo Random Noise (PRN) code to modulate the common signal enhances extraction of relative time information. The simplicity of this binary code makes it suitable for recording on various media (digitally or within audio and video streams) without the requirement for specialised and/or complex converters. The PRN technique has been applied successfully by NLR in several projects executed in a range of environments over the last decade and has shown to provide time-difference information with high precision.

Contents

(16 pages in total)

1 Introduction

The recording of data on a single recorder is often used as a method to preserve the time-relations between the various measured signals. The advent of modern computers allows storing of data in separate data files using information from a common clock, giving more freedom in data storage whilst maintaining the required time-relations. For ambulatory experiments however, it is often necessary to record data with separate devices without the availability of overlapping clock information. In those cases, it is often possible to record a simple common signal at all recording places that can be used later for synchronisation. The signal used can be divers (e.g. electrical, sound or light). For example, when recording video and performance data, one can use a flashing (infrared) Light Emitting Diode (LED). The flashing sequence can be recorded electronically with the performance data, whilst the (ambient) light changes can later be recovered from the video information. This paper provides a solution for the commonly encountered problem of how to determine time differences between simultaneous measurements using (pieces of) overlapping information.
An example of a 'generic' set up of a measurement system with multiple recorders (and thus the data synchronisation problem) is given in figure 1. Other examples of measurement systems with similar synchronisation problems are reported in for instance (Harlaar, Redmeijer, Tump, Peter & Hautus, 2000, Geuze & Hunnius, 2002).

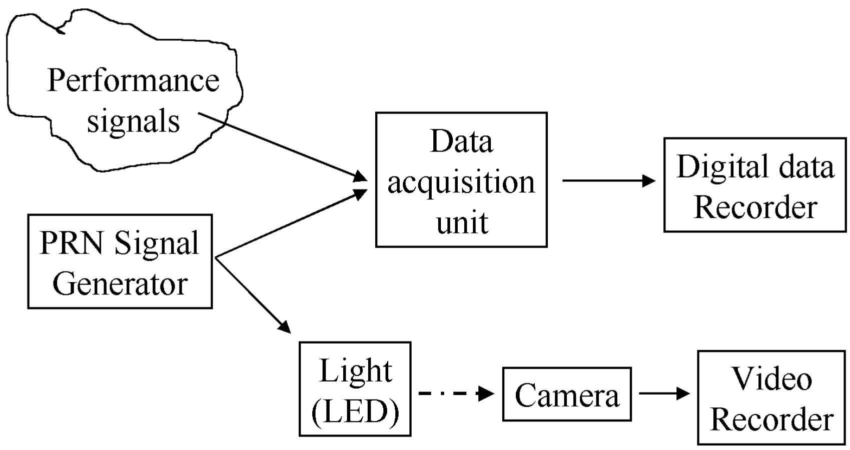

Figure 1: Example measurement set-up with a data synchronisation problem

The synchronisation problem is illustrated in figure 2, which shows two separate measurements 1 and 2, each containing multiple recorded channels (respectively parameters 1.m and 1.n).

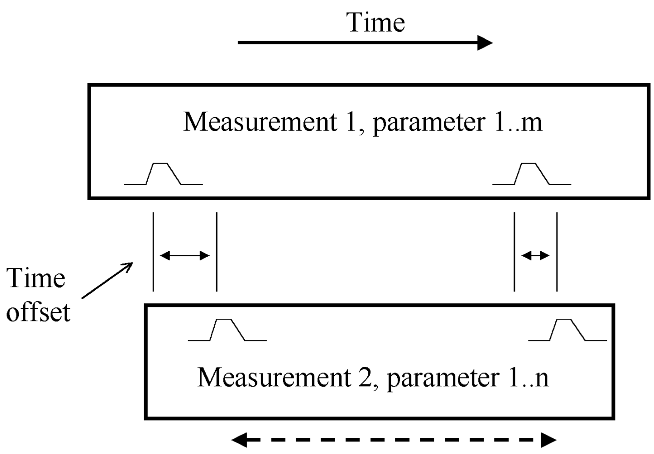

Figure 2: Two simultaneous measurements, each with a start and stop pulse. For the synchronisation, the location of the start and stop pulses need to be matched by shifting measurement 2 to the left (time-offset) and by stretching it (time-gain).

For the synchronisation the following needs to be achieved:

- Correction of the start time of each measurement. Variations are caused by the impossibility to start all recordings at exactly the same time. So the measurements need to be shifted in time in order to match the location of the recorded start pulse.
- Compensation of differences in recording speeds. This is shown in the figure as a different interval between the start and stop markers. Several causes can be mentioned for such a difference. The simplest one is a difference in recorder clock speed (based on practical experience, we can state that clock-speed differences are normally less than approx. 0.2%; Stein & Vig, 1992; Vig, 2001).
- Resolving 'Measurement gaps' or 'frame-drops': they occur when data records are missed in the recording process.

2 Synchronisation methods

If it is assumed that no measurement gaps are present, the simplest solution is to shift the second measurement in time to match the start of the first measurement, followed by stretching or compressing the second measurement to also match the end of the first measurement. In other words: ***'SHIFT and STRETCH to Match'***, or mathematically: apply a time-offset and a time-gain. The actual time of a given value of a parameter from measurement Y is then given by:

$$T_{py} = T_{measurement\ Y} + T_{offset\ Y} + (T_{parameter} + T_{p\_offset}) * Time\_gain_{Y}$$

With:

T_{py}	Time of parameter p of measurement Y
$T_{measurement\ Y}$	Start time of measurement Y
$T_{offset\ Y}$	Offset in the start time of measurement Y (relative to the reference measurement)
$T_{parameter}$	Time of the parameter value relative to the start of the measurement
T_{p_offset}	Optional offset specific for parameter P (relative to the other parameters in the same measurement). In most cases this value is assumed to be equal to zero and, therefore, will be further ignored in this paper.
$Time\_gain_Y$	Time gain for measurement Y

In order to synchronise multiple measurements off-line, some overlapping common information needs to be present in each measurement. The simplest one is to record a unique event like a start-marker. This start-marker is subsequently used to shift the measurements in time. Unfortunately, it only provides a time-offset and no time-gain. As a consequence, longer intervals of measurements will reduce the accuracy progressively. A first extension to the start-marker method is the addition of an end-marker. Using two markers, well separated in time, allows for a determination of the time-gain. This 'two marker' method has advantages in case one of the markers was not recorded, but is limited by the accuracy of the measurement of the marker information. For example, the signal conditioning (like amplification and filtering) and sampling frequency determine the obtainable reconstruction and timing accuracy of the marker signals. If the required time-resolution is more demanding, the approach of using just one or two markers may be insufficient and time has to be recorded in a more continuous way.

2.1 Continuous time recording

The need for synchronisation requires that (absolute) time can be reconstructed in all simultaneous measurements. If data of different modalities are to be recorded (e.g. performance data and video), some translation of the time-information to the other modalities can be required. For video-media, use can be made of the Vertical Interval Time Code (VITC) method, which allows recording of digital information in hidden or invisible video lines. Those hidden lines are not converted to computer based video formats (unless a work-around is provided for), making the VITC method often less suitable. An alternative is to record time information in the visible part of the video signal, since this survives the conversion to digital video formats. However, the difficulty is that one is to retrieve the recorded time information in a computer readable format. Another drawback is the limited video resolution.

The audio channel can also be used for storing time-related information. The bandwidth of the audio-channel is limited, resulting in restrictions for the stored amount of information. Also the time-relation between video and audio is only guaranteed for some hundreds of milliseconds since the design driver for the conversion soft-and hardware was 'lip-synchronic' replay. So, although feasible, techniques utilising the audio-channel are highly dependent on the used hardware set-up and quality for the digitisation of the analogue information. The remaining of this paper focuses on a practical method, which does not require much space within the video image and that can also be ported to the audio domain or to other digital and analogue recording systems.

2.2 Use of information overlap

Measurements can be synchronised by maximising the similarity of common information at each point in time. To this end, three different interactive synchronisation methods are tested in the NLR HEART data analysis package (Hoogeboom, 2000):

1. Trial and error: the user is given the option to numerically enter the time-offset and time-gain values of a measurement, after which the data is updated to reflect the changes (e.g. time-traces of the signals to be matched).
2. Graphical slewing of measurements: this method allows the user to select a characteristic point in one of the measurements, followed by a similar point in the other measurement. Subsequently, the time difference between the points is used to update the time-offset.
3. Graphical matching of four points: two reference points are compared with two others in the measurement to be synchronised. This method allows the determination of the time-offset and the time-gain.

All three options are powerful and very useful, but the signal synchronisation is essentially performed manually. Fortunately, automated synchronisation can be achieved by a matching process, corresponding in mathematical terms to maximising the correlation of the two signals. In this case, the ‘standard’ correlation value needs to be extended to a correlation function C(dt), which provides the measure of similarity between two signals with a time-offset dt:

$$C(dt) = \frac{\sum x(t) * y(t-dt)}{\sqrt{\sum x(t)^2 * \sum y(t-dt)^2}}$$

Normally, the function is computed over a certain window length (-N..N), as is implied by the sigma function. Also it is assumed that the mean of x(t) and y(t) are equal to zero.

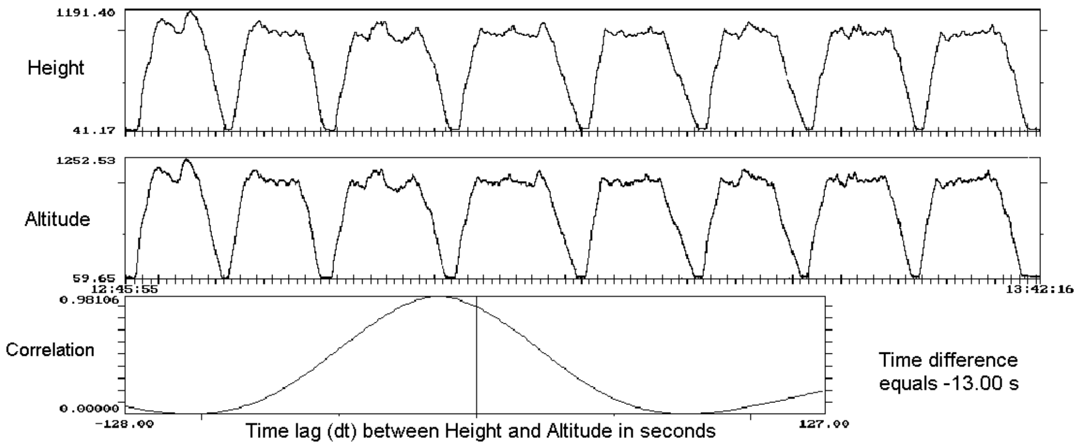

Figure 3: Example correlation function (bottom trace) showing the correlation of two signals (upper two traces) relative to the time offset between those two. The correlation function peaks left of the "no-time offset" location in the centre, indicating the upper trace is delayed by 13 s relative to the middle trace

In figure 3 two time traces are plotted: the first (top) one is the height above a reference system as derived from a Global Positioning System (GPS). The second channel shows the altitude of the aircraft, as derived from barometric air pressure. Both signals have similar patterns, with the barometric altitude signal slightly leading the GPS height signal. The third trace shows the correlation function C(dt) of the two signals. This function peaks left of the centre line (which equals to zero time difference between the signals, or dt=0 s). The time difference between two signals is determined to be –13 s, meaning that the altitude is 13 s in advance to of the GPS height channel. The correlation function peak itself is not very sharp, and limits the accuracy of the time-difference estimation.

2.3 Pseudo Random Noise (PRN)

To enhance the resolution of the correlation function, a special digital signal can be used. The requirements for its use can be summarised as having a high correlation for signals without time-offsets and a low correlation for signals with some time difference. Also the 'code repetition-distance' should be sufficient to avoid inappropriate locks. Substantial research has been performed in the area of PRN codes (Brown & Hwang, 1992; Dixon, 1994; Gold, 1967; Gold, 1968; Pursley & Roefs, 1979, Stein & Vig, 1992). Some navigation systems like the GPS (Anonymous 1997; Lin & Tsui, 2000) even rely solely on similar types of codes. Also, similar techniques are widely used in cellular phone systems.

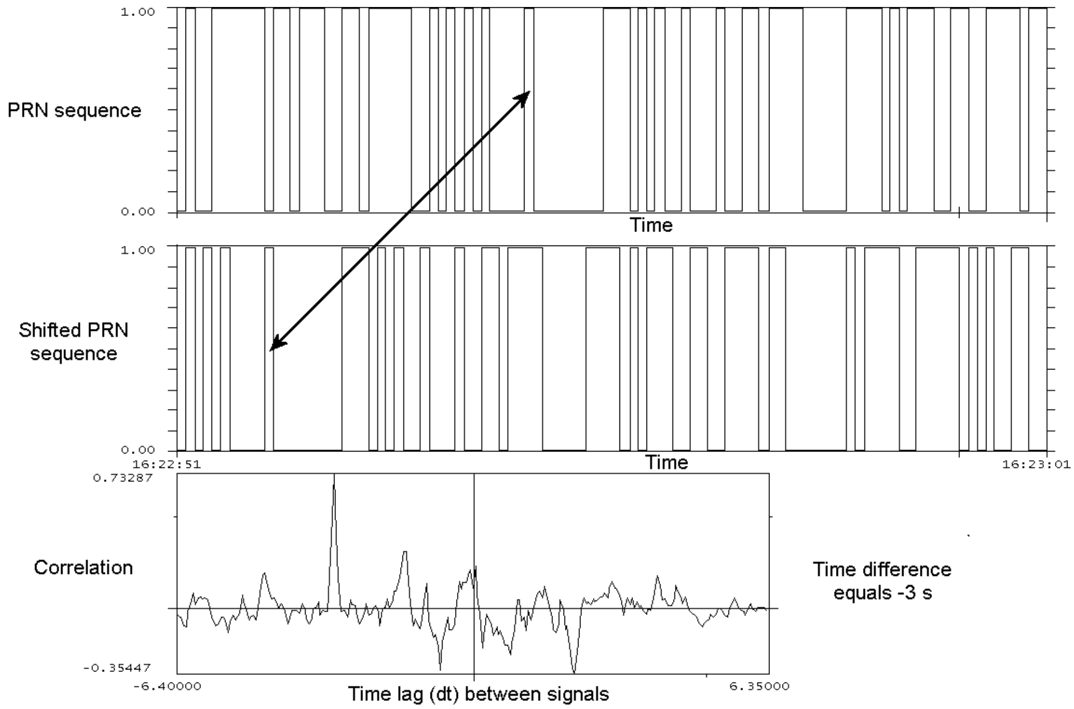

Figure 4: Correlation function of a pseudo-random noise (PRN) signal for an interval of 10 s and a relative time-difference of 3 s, with the middle trace leading the upper trace.

Figure 4 shows a typical example of two different measurements containing overlapping PRN information 'PRN sequence' and 'Shifted PRN sequence'. The code generation frequency is 10 Hz. The correlation function C(dt) of the two signals is indicated in the third trace, and peaks about 3 seconds before the 'no-time-offset' location (vertical line in the middle of the window). This is confirmed by the two time-histories, from which it is easy to see that the shifted PRN sequence is earlier than the original PRN sequence.

The peak in the correlation function is now sharp and also relatively high when considering a time-difference of 3 s combined with a data-window length of only 10 s. Note: increasing the data-window length enlarges the peak/noise ratio, but at the cost of prolonged computation times.

The practicality of the PRN based time-difference detection was tested with several data sets in the context of the Visual Lab project (Noldus, 1999). The accuracy of the time-difference determination was found to be dependent on the used PRN code, the applied data window length, the minimum time-step (dt) used in the correlation function, the method to locate the peak in the correlation function, and the time-gain difference between the measurements. The time required for the computation of the correlation function depends on a similar set of parameters. A higher accuracy requires longer computation times. Therefore for most applications a trade-off exists between the required accuracy and the allowed computation times.

Surprisingly, it was found that major distortions to the PRN signal are allowed: due to the correlation process local differences (e.g. small measurement errors) are averaged out and have almost no influence on the accuracy of the determined time difference. This property of the correlation function makes the PRN technology relatively robust compared to using a single and dual marker technique.

3 Detection of measurement gaps using the PRN code

So far, only the linear situation has been looked at -- meaning the derivation of the time-offset and time-gain -- ignoring the possible presence of missing records. To test the behaviour of the PRN for such cases, a test-file was created consisting of a continuous PRN sequence and a (deliberately) interrupted PRN sequence.

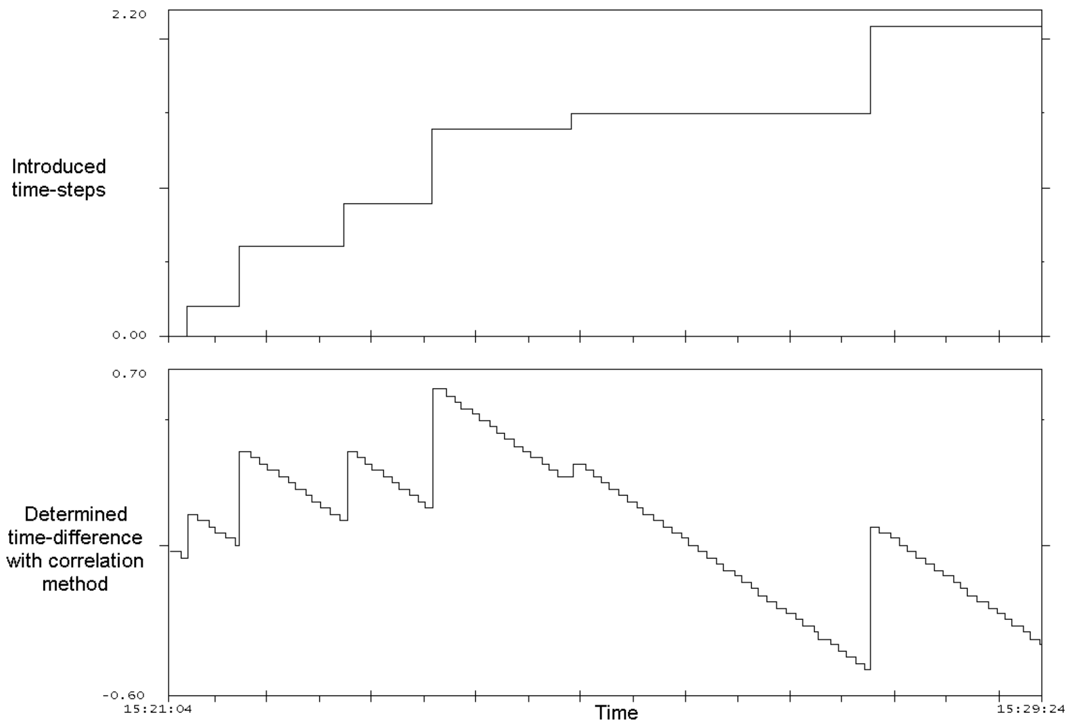

Figure 5: Measurement gap detection: the first trace indicates the location and magnitude of artificially introduced record losses between two PRN recordings. The second trace indicates the determined time-difference between the two recordings using the correlation-based approach. Both the occurrence and the magnitude of the measurement gap are detected. The slope in the second trace indicates a difference in time-gain between the two recordings.

The top time-trace ‘offset’ in figure 5 presents the time-offset of the second PRN sequence relative to the first. Each ‘jump’ indicates that a certain number of records were discarded in the second sequence. The magnitude of the jump indicates the amount of time skipped. The interrupted PRN sequence was stretched by applying a gain difference of 0.5 %. Both data sequences were compared using the correlation technique with a sliding window of 100 clock intervals, stepping at a pace of 10 clock intervals. (Note: the clock interval, or chip, is the inverse of the code generation frequency; with a code generation frequency of 10 Hz, 10 intervals correspond to a window step size of 1 s). The result of the automatically determined time difference is given by the second trace, which 'jumps' at the expected locations. The slope of the second curve is caused by the artificially introduced time-gain difference. Stretching the second run compensated the lack of data records, therefore the jumps were expected to go in the opposite direction (compared to the slope caused by the time-gain difference). As such it can be concluded that the PRN technique can a) detect measurement gaps and b) provide quantitative information on both the location in time and on the magnitude of the gap.
The relatively clear detection of the location of the “time-offset jump” can be explained by the mechanism of the correlation algorithm in combination with the applied peak detection algorithm. Close to a time jump, the correlation function C(t) will show two peaks. One peak belongs to the data segment just before the time jump, whilst the other peak belongs to the data segment after the jump. The magnitude of the peaks depend on the relative size of the data segments, with the largest peak belonging to the largest data segment within the correlation window. At the moment both peaks are of equal magnitude, the size of both data segments is equal and hence the time-jump is in the middle of the correlation window. Any change in this situation will force the peak detection algorithm to select the highest peak, which causes a jump in the detected time-difference. As such, the accuracy of the jump-location depends on the accuracy of the peak detection algorithm and on the used time-increment in the correlation function.

4 Practical recording of PRN in video signals

The practicality of the PRN technique can be illustrated by our experiences gathered in amongst others the DIVA (Design of human factors Interfaces and their Validation in Aeronautics) project (Nibbelke, Ferro & Hoogeboom, 2001) which was co-sponsored by the European Commission. Within this project pilot/crew performance and behaviour were measured in a flight simulator. The recorded signals included physiology, data from the simulator as well as video observations. In this specific case, security reasons prohibited a direct connection of the recording equipment (consisting of several computers) to the simulator network. Additionally, to minimise the risk of electrical shocks for the crew, it was (and is) not allowed to connect the physiological recording equipment directly to the simulator computers. Therefore a simple interface was used through a specially developed ‘ break-out’ box, providing the necessary insulation. The simulation equipment transmitted through this box the digital PRN code to the recording equipment. A green LED allowed recording of the PRN information on one of the four videos taken from the cockpit. In the video-stream, the LED image had a size of approx. 15 pixels (horizontally). After combination of the four video streams (resolution reduction by a factor 2) and conversion to MPEG 1 (another resolution reduction by a factor of 2) only a couple of pixels remain in the final digital image. An example screenshot of the resulting video is given in figure 6.

Figure 6: Example video screenshot: the arrow indicates the LED location

The digitised video was subsequently used to extract the PRN information by allowing the data analyst to select an area on the screen. The height and width of the selection-square were approximately 3 pixels. The video was replayed in a window with a width of 352 pixels. The total intensity of the square was calculated based on the RGB values of the pixels within the selection-square (repeated for each video frame). If the intensity of the square was equal or above a user selectable threshold, then the status of the LED was assumed to be ‘ON’, otherwise it was set to ‘OFF’. The status of the LED was subsequently stored in a data file for further processing. During this experiment, the following lessons were learnt:

- The background of the LED should be dark enough to be able to discriminate between the two states.
- The use of a stationary scene camera worked well. During each simulator run the location of the LED in the image was stationary. However between different runs, the location could change a little, forcing the data analyst to re-adjust the selected detection area.

- Sometimes data was lost due to a person walking between the camera and the LED. Depending on the selected threshold and the colour of the clothes, the LED status was fixed either to the 'ON' or the 'OFF' mode.
- With the small selected size of the detection-square, the observed intensity difference was about 80 (at a scale between 0 and 255).

As a whole, it was surprisingly simple to retrieve the code information.
The correlation function of the video-derived PRN signal and the separately measured electrical signal indicated that the video was delayed at the start of the measurement by approximately 2.31 s. At the end of the measurement, the video was leading with 1.4 s. The total time-difference was 3.7 s over a measurement length of 2244 s. In other words, a time-gain difference of 0.165%.
An alternative approach, utilising the signal changes to estimate the clock-interval from observed data, revealed that the actual code generation frequency was 1.79 Hz. The time-gain difference between the two measurements was estimated to be 0.18%, which is close to the value derived above. Note: this alternative method is less precise and is not able to estimate the time difference between the measurements. The main benefit of this alternative method is that it can be used to bring the time-gain differences within the acceptable range for the correlation technique.
To detect possible unwanted frame-drops (measurement gaps), the time-difference development between the video and physiological signal was determined with the result as shown in figure 7.

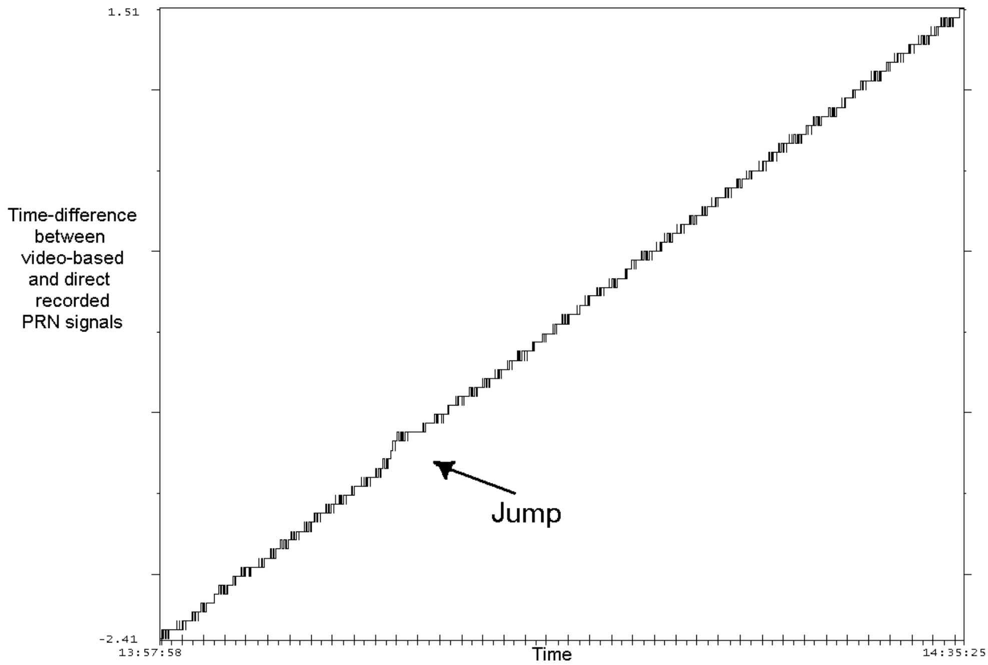

Figure 7: Time difference development for an example video. A moving window is used for the determination of the time difference. The vertical step-size is caused by the (selected) resolution of the time-determination algorithm. The deviation from the linear line ("Jump") is caused by a drastic change in illumination of the environment affecting the LED on/off detection.

The used data window was 100 clock intervals, which corresponds to 55.9 s (code generation frequency = 1.79 Hz). The correlation function step resolution was 0.1 clock interval (or 0.1/1.79 = 0.0559 s). The calculation window was moved in steps of 1 second over the complete measurement. The expected linear relation is present. The small jump, occurring at one-third of the measurement, coincides with a change in the experimental procedure: the first experimental run ended and the cabin lights were turned on to allow the pilots to complete some questionnaires. Some PRN data points, as derived from the video information, were therefore found to be incorrect. Especially the on-off times showed deviations (the on-times were too large).

The vertical step size of the line is caused by the temporal resolution of the correlation step (set at 0.1 clock interval) in combination with some remaining noise. Therefore the determined peak location alternated between the two resolution values surrounding the actual time-difference. A smaller correlation step-size or an alternative peak estimation method would reduce the vertical step size at the cost of increasing the computation times. Also note that the estimate of the time-gain and time-offset between the measurements can be improved by taking the best fit of a line through all the data points.

5 Conclusions

This paper presents an approach to determine time-offsets between different data sets with high accuracy. The method is based on the use of a correlation function in combination with a binary Pseudo Random Noise (PRN) signal. Related signals are also used in equipment like satellite navigation systems (e.g. GPS) and telephone communication services (GSM). The time-difference determination method has been successfully applied by NLR in research projects in various circumstances. The technique is a likely candidate to be implemented in future Noldus products (Noldus, Trienes, Hendriksen, Jansen & Jansen, 2000).
Even though the technique is powerful, it was found that the robustness of the correlation function for larger differences in recording speed (time-gain difference larger than approx. 1%) needs improvement. Therefore, in those cases alternative time-gain difference estimations have to be used to increase the reliability of the proposed synchronisation technique. Fortunately, the PRN signal provides this information in the form of observable signal level changes. Experience indicates that the observed time gain difference is normally less than approx. 0.3%, and so the sensitivity to the time-gain differences of the method is adequate for practical applications.
So far it was not necessary to use the details of the PRN code: the signal waveform is totally deterministic and can be generated afterwards. Correlating the measured signal with the expected code sequence is an excellent method for reducing measurement errors and re-sampling noise. In that reuse the technique discussed and illustrated has sufficient growth potential for future challenges in synchronisation of valuable, but hard-to-record data.

6 References

1. Anonymous (1997). Navstar GPS Space Segment/ Navigation User Interfaces, ICD-GPS-200, Revision C, Initial Release. *GPS NAVSTAR JPO,* http://GPS.losangeles.af.mil
2. Brown, R., & Hwang, P. (1992). *Introduction to Random Signals and Applied Kalman filtering* John Wiley and Sons*,* New York
3. Dixon, R.C. (1994). *Spread spectrum systems with commercial applications, third edition,* John Wiley and Sons, New York
4. Geuze,R.H., & Hunnius, S. (2002). Measurement and analysis of eye movement and heart rate as markers of visual attention in babies. *Measuring Behavior 2002, 4th International Conference on Methods and Techniques in Behavioral Research* (Amsterdam, The Netherlands, 27-30 August 2002), 82-83, http://www.noldus.com/events/mb2002
5. Gold, R. (1967). Optimal binary sequences for spread-spectrum multiplexing, *IEEE Transactions on Information Theory, vol.IT-13*, 619-621

6. Gold, R. (1968). Maximal recursive sequences with 3-valued recursive cross-correlation functions. *IEEE transactions on Information Theory, vol.IT-14*, 154-156
7. Harlaar, J., Redmeijer,R.A., Tump, P., Peters, & R., Hautus, E. (2000). The SYBAR system: integrated recording and display of video, EMG and force plate data. *Behavior Research Methods, Instruments & Computers, 32*, 11-16
8. Hoogeboom, P.J. (2000). Human Factors Evaluations, data Analysis and Reduction Techniques (HEART), *Proc. 2nd IEEE Benelux Signal Processing Symposium (SPS-2000), Hilvarenbeek, The Netherlands*, Brussels, IEEE, www.ieee.be
9. Lin, D.M., & Tsui, J.B.Y (2000). Comparison of acquisition methods for software GPS receiver. *ION GPS 2000 conference, September 19-22,* Fairfax, The Institute of Navigation, 2385-2390
10. Nibbelke, R., Ferro D., Hoogeboom P.J. (2001): Design and Evaluation with the Human in Mind. *Air&SpaceEurope, Proceedings of the Aeronautics Days 2001, Hamburg,* Amsterdam, Elsevier.
11. Noldus, L.P.J.J. (1999). Visual Lab: Integrated Measurement, Analysis and Visualization of Behavior and Physiological Signals. *Senter project BTS-99013,* April 1999- March 2002, http://www.noldus.com/events/mb2000/program/sig_visuallab.html
12. Noldus, L.P.J.J., Trienes, R.J.H., Hendriksen, A.H.M., Jansen, H. & Jansen R.G. (2000). The Observer Video-Pro: new software for the collection, management, and presentation of time-structured data from videotapes and digital media files. *Behavior Research Methods, Instruments and Computers, 32,* 197-206.
13. Pursley, M.B. & Roefs, H.F.A. (1979). Numerical evaluation of correlation parameters for optimal phases of binary shift registers. *IEEE Transactions on Communications*, *vol. Com-27,* 1597-1604
14. Stein, S.R. & Vig, J.R. (1992). Communications frequency standards. In Froehlich, F.E. and Kent, A. (Eds.), *The Froelich/Kent Encyclopedia of Telecommunications*, Vol. 3, Marcel Dekker Inc., New York, 445-500.
15. Vig, J.R. (2001). Introduction to Quartz Frequency standards. http://www.qct.nl/theory/toc.htm