

# Tracking of Hidden Image Structures Employing Gold Sequences

C. Schütz\*, K. Kroschel\*\*.

\*Fraunhofer Institut für Informations- und Datenverarbeitung IITB, 76131 Karlsruhe  
Germany (Tel: xx49-721-6091-225; e-mail: schuetz@iitb.fraunhofer.de).

\*\* Tel: xx49-721-6091-587; e-mail: kroschel@iitb-extern.fraunhofer.de

---

**Abstract:** Tracking of man-made objects offers a great flexibility in designing adequate markers for tracking purposes. Invisibility for human observers and simultaneous recognisability of these markers for the tracking system causes more effort in marker design. An approach to design and track markers using Gold codes is shown.

*Keywords:* Image processing, tracking, Gold code, steganography, watermarking

---

## 1. INTRODUCTION

Optical tracking of objects is performed using an image acquisition system based on common hardware (camera, internal or external A/D-converter, PC) and specialised software. The software algorithms need to be adjusted to the visual appearance of the objects. Real-world objects generally satisfy at least one of commonly known requirements for tracking. Colour, large gradients of brightness orthogonal onto edges, or a fixed known shape of an object serve as criteria for usual tracking algorithms. Tracking of objects surrounded by a large amount of clutter is a more challenging task. Fusion of multiple features is an approach to that task, marking of the objects may further alleviate the tracking complexity.

The Situation Table with Fovea Tablett® (Peinsipp-Byma et al., 2007) is used as a multi-user platform in applications like disaster management and environmental planning where the scene of interest is presented by maps or by air-borne images. The Situation Table consists of a work station, a horizontal screen which acts as a work-table displaying the situation overview, a vertical screen which acts as a board for additional information, and tablet PCs as so called and patented Fovea-Tabletts® (abbr.: FT) (Patent 2007), which offer detailed information. Fig. 1 shows the table and two FTs. In order to allow high precision positioning of each tablet PC, the FTs are marked underneath by an identification marker with an integrated circular bar code. A camera below the work-table with a transparent tabletop records an image of the work-table including all FTs with their respective identification marks. While the bar code of the identification mark identifies the FT, the design of the identification mark enables an exact measurement of the FT's position and rotation angle. In Fig. 2 the so called Multi-Cursor-MarkerXtrackT marker (MC-MXT marker) and an image taken by the camera from underneath the work-table are shown. Details of the tracking process and the markers may be found in Rehfeld (2001). In order to avoid interference of

the recorded image with the projected image onto the work-table, the measuring camera mounted under the table operates at near infrared ( $0.78 \mu\text{m} - 1.5 \mu\text{m}$  wavelength).



Fig. 1. A Situation Table with two Fovea Tablett®

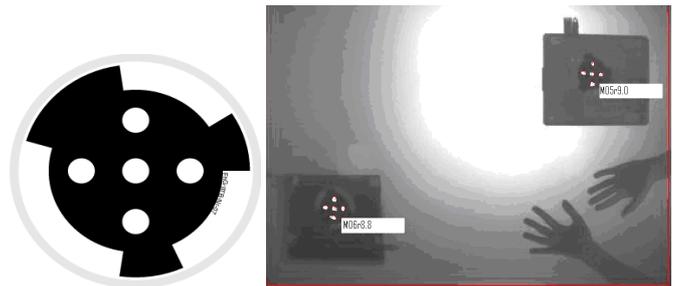


Fig. 2. Left: an MC-MXT marker. Right: tracking the tablet PCs from underneath the table

In a future version of the table, the large horizontal projection screen will be replaced by a luminous LCD, so tracking from underneath will no longer be possible. The left part of Fig. 3 shows a prototype of the Situation Table with cameras tracking the FTs from above.

Tracking from above requires additional markers noticeable to the computer vision system, because the “natural” structure

of the FTs occasionally may not be distinguishable from the image presented on the table. Mainly, additional markers are required because the shape of the FTs themselves does not vary enough to differentiate them from each other.



Fig. 3. Left: horizontal large display surface (gray) and two little displays (black) with high resolution lying upon the large display, vertical display for additional information (black), cameras 1.6 meters above the tabletop. Right: detailed view of cameras and illumination

For tracking purposes, these markers have to be suitably textured, but the resulting saliency of the markers collides with an invisibility favoured by the human observer. Combining visibility and invisibility causes the usage of watermarking. The signal of the information dedicated to the observer has a very large intensity compared to the embedded watermarking. In other words: the local signal-to-noise ratio is very low and the hidden marker has to withstand this low SNR. The MC-MXT markers used so far are based on a maximum contrast within the marker. Hiding these markers in an image will immediately impede any matching and the tracking will fail. Thus replacements for MC-MXT markers have to be developed.

## 2. TRACKING BY CORRELATION

For the localisation and tracking of objects in a noisy environment, correlation techniques are often used and have proven to be robust (Kroschel et al., 2005). In the problem discussed in this paper the marker is the object to be localised and tracked, and the noise is the optical information carried by the map or air-borne image.

From the point of view of correlation techniques the most appropriate marker is the one with a correlation function with a sharp peak, i.e. a peak with a large magnitude with respect to its environment and a minimal extension. A first approach could be to use the MC-MXT marker also for the new Situation Table given in Fig. 3.

The correlation function of the MC-MXT marker, shown in Fig. 4, proves that it is not suitable for a noisy environment. Due to the low ratio of the main lobe peak and the side lobe peaks of the correlation function and the low gradient of the main peak, noise will prevent a robust tracking.

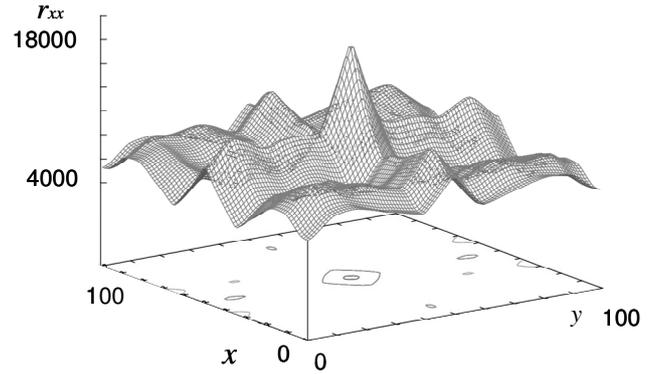


Fig. 4. Auto-correlation function of the MC-MXT marker

Tracking by means of correlation requires codes providing an optimal ratio of the main peak and other local peaks of the correlation function. R. H. Barker (1953) found optimal codes in this sense of length

$$N \in \{2,3,4,5,7,11,13\}. \quad (1)$$

The relation of the maximum peak and the other peaks of the correlation function of Barker codes is  $\frac{1}{N}$  with  $N$  given in (1). Up to now sequences larger than  $N=13$  have not been found. Short sequences do not withstand large levels of noise, thus sequences with properties similar to those of the Barker sequences, but with a larger number of elements have to be analysed.

In deep space probe communications, Gold codes are widely used to encode digitally transmitted data. R. Gold (1967, 1968) proved the existence of a series of maximum-length sequences with correlation properties similar to those of the Barker codes. The spectra of these codes are extremely broad-band which is comparable to the noise degrading the transmitted data in a noisy environment. This technique is therefore called spread spectrum technique and Gold codes are used for data transmission over channels with a low signal-to-noise ratio (SNR). Fig. 5 shows two shift registers of length  $m=5$  and the resulting Gold sequence with  $n=2^m-1=31$  elements.

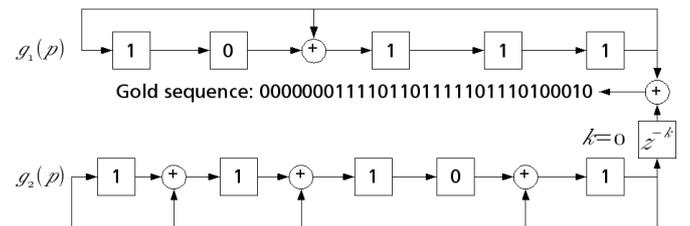


Fig. 5. Gold sequence generation: Generator polynomials  $g_1$  and  $g_2$ , delay  $z^k$ , in this example generation without delay, i.e.  $k=0$ , all additions modulo 2.

The Gold sequence generated by the polynomials  $g_1$  and  $g_2$  yields the signal  $s$  shown in Fig. 6. Adding white noise of identical power results in the signal  $s_d$  shown in Fig. 7. Due to the low sampling rate introduced by a camera system, only low-frequency components of the noisy signal remain in  $s_d$ . Thus  $s_d$  is a step function with a step width equivalent to the

resolution in time. Given an appropriate SNR, the Gold sequence is visually undetectable.

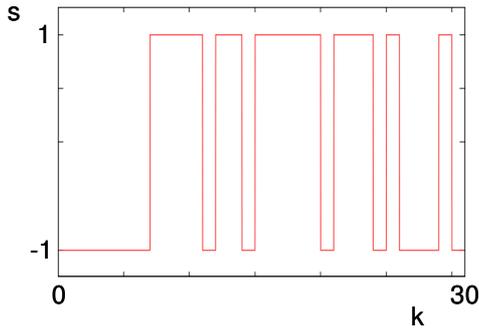


Fig. 6. Binary signal, equal to the Gold sequence shown in Fig. 5

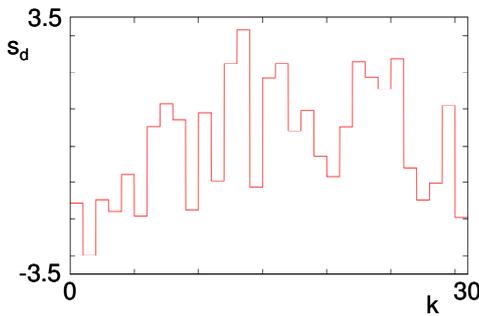


Fig 7. Binary signal with added white noise of identical power after signal digitisation with low sampling rate, the Gold sequence is visually not detectable any more

The auto-correlation function of the Gold code of length  $n=31$  is shown in Fig. 8. The main peak is sharp and the ratio of the main peak with respect to the other peaks is much higher than in the auto-correlation function of the MC-MXT marker given in Fig. 4. Therefore, this type of marker is more suitable for tracking purposes, because the determination of the sequence position will be more precise than using the MC-MXT marker.

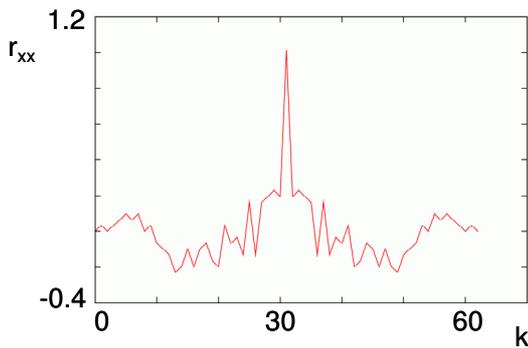


Fig. 8. Auto-correlation of the Gold sequence shown in Fig. 6

Due to the short length of the chosen Gold sequence, the additional noise has a significant effect on the cross-correlation of the noise sequence and the noise-free sequence, shown in Fig. 9. Despite the fact that the ratio of the main peak with respect to the other peaks is now reduced in comparison to the auto-correlation function given in Fig. 8,

the main peak is still appropriate to localise and to track the signal pattern. However, a further increasing noise level will impede the retrieval of the Gold sequence.

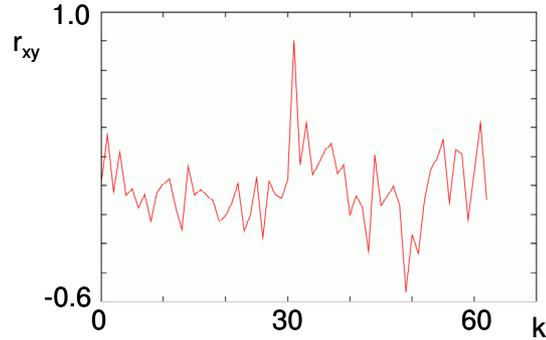


Fig. 9. Cross-correlation plot of the Gold sequence and the sequence with additional white Gaussian noise of SNR = 0 dB shown in Fig. 7

### 3. 2D-MARKERS BASED ON GOLD CODES

While one-dimensional code sequences are suitable for 1D localisation, tracking of the FTs requires a 2D localisation. Arbitrary positions and orientations of all FTs need to be considered. Any embedded linear code sequence of the Gold type will either fail in correlation due to an angle between the original and the embedded code, or a very large set of rotated and dilated marker patterns has to be correlated to the embedded marker, which requires a very long computing time.

A rotational invariant two-dimensional implementation of the Gold code sequence of length  $n=31$  generated by  $g_1$  and  $g_2$  is shown in Fig. 10.

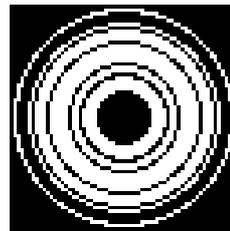


Fig. 10. Gold sequence of length  $n=31$  as a concentric ring structure.

2D-correlating an image  $G$  with an additively embedded marker  $H$  causes a large computational effort. In space domain the 2D-cross-correlation  $r_{G,H}$  of  $G$  and  $H$ , each with  $U$  rows and  $V$  columns, is calculated by

$$r_{G,H}(u,v) = \frac{1}{UV} \sum_{u'=0}^{U-1} \sum_{v'=0}^{V-1} G_{u',v'} H_{u'+u,v'+v} \quad (2)$$

Every single value of the correlation matrix requires  $O(UV)$  arithmetic operations, thus the complete correlation matrix requires  $O(U^2V^2)$  operations. Usage of the representation in the frequency domain is much more efficient. A 2D-fast-Fourier-transform requires

$O(UV \cdot \log_2(UV))$  complex multiplications and the same number of additions. Conjugate complex multiplication in the frequency domain requires  $O(UV)$  computations, and the inverse Fourier transform is of the same computational complexity as the forward Fourier transform. The overall computational effort is  $3 \cdot O(UV \cdot \log_2(UV))$  and drastically reduced in comparison to the effort needed in space domain for large values of  $U$  and  $V$ . Typical values in the application presented here are  $U \geq 1000$ ,  $V \geq 1000$ . Efficient algorithms are discussed in Jähne (1997).

Due to the short code length, the marker has to be embedded with a graylevel of 56 out of 255 values to generate a correlation peak large enough to be correctly detected. This graylevel does not provide the desired invisibility, and the marker may be detected in the upper left corner of Fig. 11 and may be detected easily in the centre of Fig. 12. The  $SNR \approx 0.31$  ( $SNR \approx -5.1dB$ ) has to be much lower to keep the marker unobtrusive.



Fig. 11. Image with embedded Gold sequence 31



Fig. 12. Enlarged view onto the upper left corner of the image shown in Fig. 11, the concentric rings are noticeable in the centre of the enlarged view

Gaining invisibility while preserving the tracking accuracy may be achieved using a larger sequence length. A 2D-Gold code sequence of length  $n=255$ , shown in Fig. 13, withstands higher noise levels, thus the intensity of the embedded marker may be reduced.

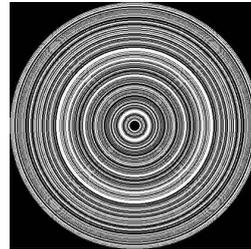


Fig. 13. Gold sequence of length  $n = 255$  as a concentric ring structure.

The auto-correlation of this sequence, shown in Fig. 14, yields a very sharp peak compared to the auto-correlation of the short sequence, shown in Fig. 8. In Fig. 15 the cross-correlation function of the Gold sequence and the sum of the sequence and white noise of identical power is shown. Even in this case at  $SNR = 0dB$  the correlation peak is unique and sharp.

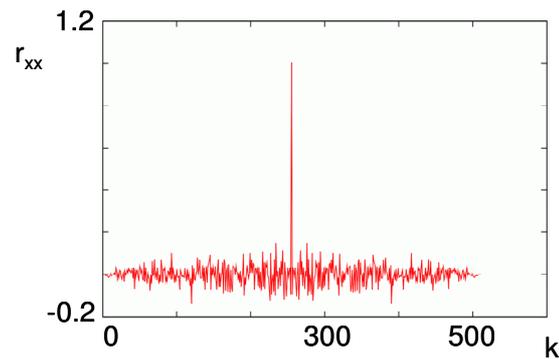


Fig. 14. Auto-correlation of the Gold sequence  $n = 255$

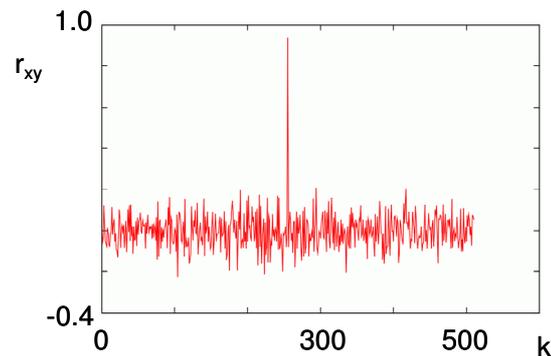


Fig. 15. Cross-correlation plot of the Gold sequence and the sequence with additional white Gaussian noise,  $SNR = 0dB$

#### 4. LOCALISATION WITH GOLD CODE MARKERS

Limitations in real-world application are imposed by the camera resolution needed to capture the total Gold sequence which is due to the concentric circle 2D shape which has the

dimension  $2n$  pixels in both height and width. The example shown in Fig. 13 contains a Gold code sequence of length  $n=255$ . The complete sequence has to be radial, so the marker incorporates  $2n=510$  rows and columns. Due to the above mentioned rotational invariance, only the position of concentric circles may be determined, while determining their orientation forces the embedding of two distinct Gold sequences into the image on each FT. An example is shown in Fig. 16.

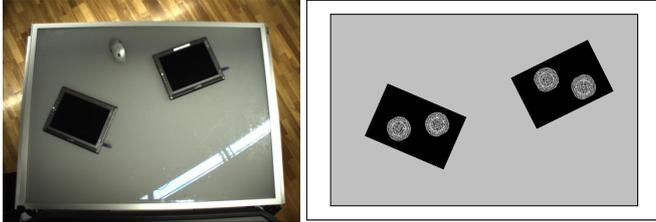


Fig. 16. Left: top camera view onto the horizontal display and the two high-resolution-displays. Right: view showing Gold sequences with high contrast

The sequence shown in Fig. 13 has been embedded into the image shown in the left part of Fig. 17 with the brightness of white stripes at the graylevel of 255 out of 255 values. In Fig. 18 the same Gold sequence embedded at the identical position with graylevel of 9 out of 255 values is shown. Human observers will not be able to detect the sequence at graylevel 9 in static images. Previous to the correlation, both the image and the marker were normalised,

$$G_{uv} \in [01], H_{uv} \in [01] \quad \forall u \in [0 \ U-1], v \in [0 \ V-1], \quad (3)$$

and the dc component was removed as well.



Fig. 17. Image with embedded Gold sequence, brightness of 255 out of 255 values

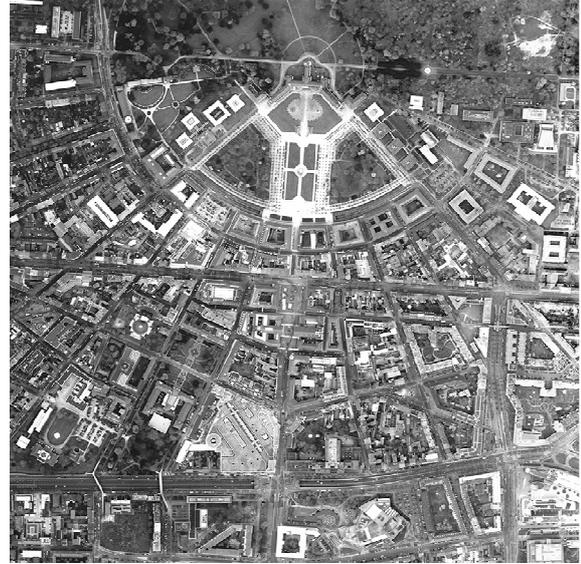


Fig. 18. Identical image as in Fig. 17 with the identical Gold sequence at the identical position, brightness of 9 out of 255 values

The 2D-cross-correlation of the marker and the image shown in Fig. 17 yields an outstanding correlation peak at an  $SNR \approx 3.93$  ( $SNR \approx 6.94dB$ ), the correlation is shown in Fig. 19. The intensity of the marker is not suitable for watermarking an image without attracting attention, as the marker is obviously perceptible in Fig. 17.

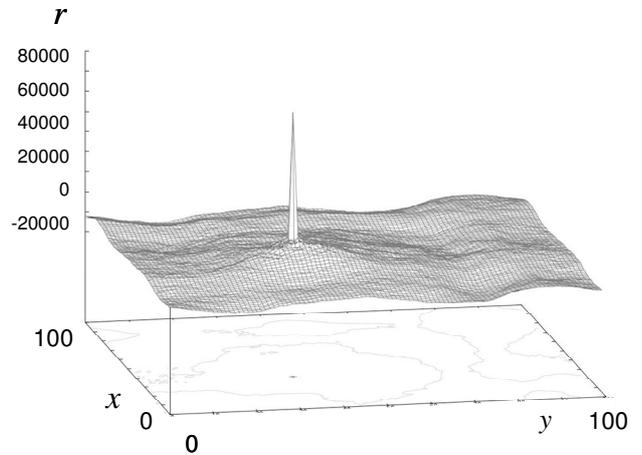


Fig. 19. 2D-cross-correlation of the marker shown in Fig. 13 and the image shown in Fig. 17, subsampling of the correlation plane by the factor 10 in columns  $x$  and rows  $y$

The 2D-cross-correlation of the marker shown in Fig. 13 and of the image of Fig. 18 results in a correlation peak at the position where the marker was inserted into the original image. Though the  $SNR \approx 0.005$  ( $SNR \approx -23.1dB$ ) is very low, the correlation peak indicates the correct position of the marker. Fig. 20 shows a plot of the total correlation plane subsampled by the factor 10. Gold codes have an excellent robustness against additional noise, i.e. against the image which is seen by the human observer.

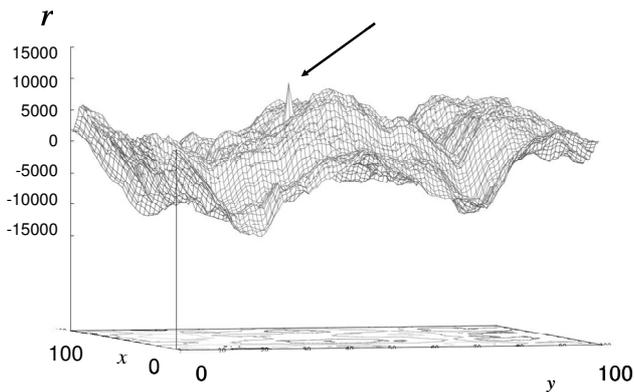


Fig. 20. 2D-cross-correlation of images shown in Fig. 13 and Fig. 18, subsampling of the correlation plane by the factor 10 in columns  $x$  and rows  $y$ , the arrow indicates the correlation peak

If the main peak is not much larger than other peaks of the cross-correlation function, the needle-shape of the peak will be useful as a criterion for position detection. Therefore, the peak may be extracted using a non-linear filtering technique, e.g. median filtering with a filter kernel size of  $3 \times 3$ , as shown in Fig. 21.

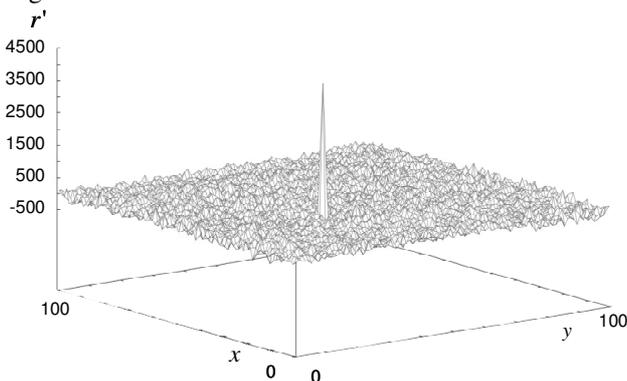


Fig. 21. Non-linear filtered 2D-cross-correlation function presented in Fig. 20, subsampling of the plane by the factor 10 in columns  $x$  and rows  $y$

The result of this post-processing, plotted in Fig. 21, yields a 2D-representation which is almost as good as the one gained for a high SNR, given in Fig. 19.

In future work an alternative approach for attaining the invisibility of the markers will have to be explored: Time-variant markers will be displayed with high intensity, but within a very short time interval. Due to the slowness of the human visual system a short-time marker will not be recognised. Blinking markers, which are displayed alternately unmodified and inverted, might be even less conspicuous, as long as the mean value is equivalent to the original image.

Using time-varying markers will increase hardware requirements. Either the cameras tracking the scene from above will have to provide an acquisition rate at least twice as high as the frame rate of the FTs, or the cameras and the FTs will have to be synchronised.

Therefore particular attention was paid to the marker design in order to avoid the usage of time-dependent markers.

## 5. CONCLUSION

The introduction of markers based on Gold codes for localisation and tracking of objects in an image environment has proven to be flexible and robust. On the one hand, the marker is not visible to the human observer due to low signal amplitudes, on the other hand the vision system is able to detect the marker with high reliability. This reliability is guaranteed by an appropriate selection of code length and depends on the required SNR.

## REFERENCES

- Barker, R. H. (1953). "Group synchronizing of binary digital systems". In Jackson, W. (ed.), *Communication Theory*. Butterworths, London.
- Gold, R. (1967). "Optimal Binary Sequences for Spread Spectrum Multiplexing". *IEEE Trans. Inform. Theory*, vol. IT-13, pp. 619-621.
- Gold, R. (1968). "Maximal Recursive Sequences with 3-Valued Recursive Cross Correlation Functions". *IEEE Trans. Inform. Theory*, vol. IT-14, pp. 154-156.
- Jähne, B. (1997). *Digitale Bildverarbeitung*, ch. 2.5. Springer Verlag Berlin Heidelberg, Heidelberg
- Kroschel, K., Schlosser, M., Grimm, M. (2005). "Tracking a Moving Person Based on a Microphone Array and a Stereo Camera". *Proceedings zur 31. Deutsche Jahrestagung für Akustik, DAGA '05*, München, pp. 771-772
- Patent DE 10 2004 046 151.1-32 (2007). "Vorrichtung zur visuellen Darstellung graphischer Detailinformationen".
- Peinsipp-Byma, E., Eck, R., Rehfeld, N., Geisler, J. (2007). "Situation Analysis at a Digital Situation Table with Fovea-Tablet®". In Erbacher, R. et al. (eds.), *Visualization and Data Analysis 2007*, Proceedings of SPIE Volume: 6495.
- Rehfeld N. (2001). "Codierte Marker als Mess- und Interaktionskomponente für das Kindermuseum ZOOM in Wien". *Fraunhofer IITB Jahresbericht*, 2001, pp. 32-33.