# Modulation Concepts for Visible Light Communication Using Video Displays

Ruediger Kays, Senior Member, IEEE
Communication Technology Institute
TU Dortmund University
Dortmund, Germany
ruediger.kays@tu-dortmund.de

***Abstract*— Visible light communication (VLC) is an attractive alternative to radio frequency communication for applications requiring a wireless short-range link. Different solutions have been developed, usually applying LEDs as transmitters and photodiode receivers. This paper refers to a new system, which uses video displays as massive parallel modulators and cameras as receiver. Data transmission is designed such that it can run on top of a video reproduction without annoying viewers. Different modulation schemes for such a transmission setup can be applied. Their parameters have to be chosen as a compromise between visibility of the data modulation and decoding performance in terms of data rate and bit error probability. In this paper, a system model is introduced that supports the design of the key components of the system.**

***Index Terms*—free space optical transmission, visible light communication, screen-to-camera communication, video device based optical transmission.**

## I. Introduction

Using visible light for data transmission is a concept with interesting features. It does not occupy scarce radio spectrum and can be used without any license. Furthermore, VLC provides a visible, well-defined coverage area and does not give a reason for users to be afraid of electromagnetic radiation. Typical transmitters and receivers are cheap. An interesting characteristics of some VLC concepts is the option to re-use existing devices like LED lamps as transmitters. Many publications describe VLC systems, and key aspects for different applications have been standardized in IEEE 802.15.7-2011 [1]. Good overviews on technical aspects of using general-purpose lighting for transmission can be found in [2] and the references therein. Concepts for transmission links supporting high data rates are described e.g. in [3].

The application of cameras instead of photodiodes as receiver has been proposed in several publications (see e.g. [4]), especially for applications in infrastructure-to-car-communication. The idea of screen-to-camera-communication also has been considered in the past. System proposals such as PixNet [5], CoBra [6] and LightSync [7] use LCD screens as transmitters and mobile phones as receiver. The authors of these papers propose several concepts to overcome limitations of existing smart phone hardware.

This paper deals with a different approach for visible light communication. Like in the systems mentioned before, a combination of a video display and a camera is used as a massive parallel visible light transmission channel. However, in the approach described in this paper, the basic concept is different. Taking advantage from the ever-increasing performance of video displays and cameras, the intention is to provide a data transmission running simultaneously to a video representation on the same screen. Modulation amplitudes are quite small and can be added to the pixel amplitudes of the video content. While the additional data modulation should be nearly invisible for human viewers, a high-resolution camera picking up the scene will be able to demodulate the data. Fig. 1 shows the principle arrangement. This concept has been presented for the first time in [8].

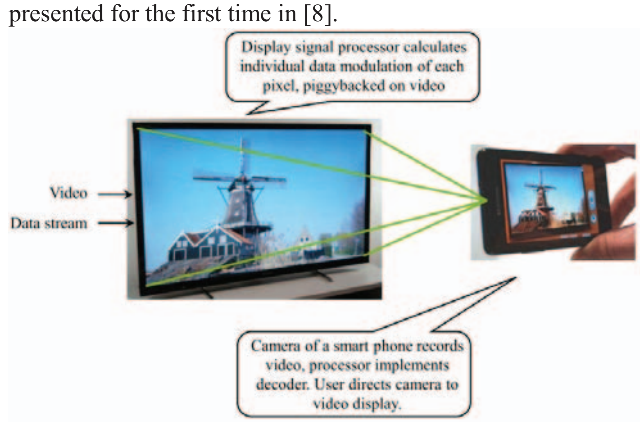


Fig. 1. Basic transmission concept

Many application scenarios of this technology can be designed and appear to be very attractive. While the basic principles remain constant, the technical details have to be adjusted to the applied hardware setup.

This paper describes results on the design of the modulation schemes and related parameters. In chapter II, options for the modulation are described in more detail. This is the basis for the system model, which is explained in chapter III. Results on picture quality evaluation are given in the fourth chapter. Chapter V concludes the paper and gives an outlook on future work.

## II. Modulation Concept

The video display provides a large number of pixels, each consisting of a specific arrangement of subpixels for RGB reproduction. Generally, rendering of moving pictures means that each frame of a video sequence is represented by a matrix of subpixel values. In the proposed system, the video display will carry a modulation for data transmission on top. In order to allow the separation of the overlaid data sequence from the video content, a simple differential modulation is proposed. Some part of the video information is repeated, whereas data is modulated as a symmetric Manchester code and added to the video signal. Suitable modulation schemes can be characterized by their spatial-temporal structure and the modulation amplitude. One option is the repetition of video information in temporal domain (frame repetition). In this case, a spatial data pattern is added to each pair of frames containing the same video content. Another option is a spatial repetition of video content (e.g. side by side horizontally) with differential data added per pixel pair. In this case, a joint repetition and modulation of the complete color triple can be considered as well as a repetition and modulation of the chrominance only. A modulation solely of the chrominance while keeping luminance unmodulated significantly reduces the visibility of the pattern. However, it is demanding with respect to camera performance and signal processing.

In the following, the focus will be put on the temporal differential modulation as a promising candidate for good performance. Fig. 2 shows a block diagram of the modulation system.

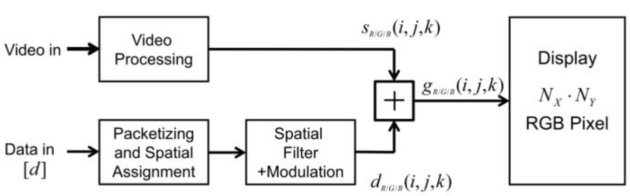


Fig. 2. Block diagram of the signal processing in temporal differential modulation

We assume a display providing $N_y$ lines, each containing $N_x$ pixels consisting of subpixels for RGB reproduction. The video input signal is processed to provide a display input *s(i,j,k)* providing the frame repetition required for the selected modulation scheme:

$$
\begin{aligned}
&s_{R/G/B}(i,j,k+1) = s_{R/G/B}(i,j,k) \\
&\quad for\ \ 0 \le i < N_X, 0 \le j < N_Y, k = 2m,\ m \in \mathbb{Z}
\end{aligned}
\tag{1}
$$

The indices $i$ and $j$ specify the horizontal and vertical pixel position on the screen, whereas $k$ is the number of the reproduced frame. Index m is the counter for the input frames of a video sequence. The amplitude range of the video signal should be limited to allow for the addition of small data amplitudes without clipping. If this is not possible, measures like forward error correction can be introduced to prevent data transmission from being corrupted.

A stream of data that has to be transmitted first has to be split up to slices of length $L$. Each slice contains the amount of data that can be transmitted in one pair of frames, counted by $m$. It is assigned to the two dimensional transmission array. A straightforward approach is a direct mapping of data bits to pixel triples line by line. In this case, slices of $N_x$ bits from each transmission block are taken and assigned successively to the pixels of each line in one frame pair:

$$
\begin{aligned}
&d(l) \rightarrow d(i,j,m) \qquad 0 \le l < L, L = N_X \cdot N_Y \\
&i = l \bmod N_x \\
&j = \lfloor l / N_x \rfloor
\end{aligned}
\tag{2}
$$

An important parameter is the modulation amplitude $A$ for data transmission. On principle, the amplitude can be selected independently for the three display primaries to optimize the system performance. In our proposal for temporal differential modulation of the luminance, we select equal amplitudes for R, G and B:

$$A_R = A_G = A_B = A \tag{3}$$

The differential modulation method assigns a sequence of {-A,A} to *d=0* and {A,-A} to *d=1*, respectively. Modulated data symbols and processed video amplitudes are added to provide the display input *g(i,j,k)*:

$$
\begin{aligned}
g_{R/G/B}(i,j,k) &= s_{R/G/B}(i,j,m) + A \cdot (2 \cdot d(i,j,m) - 1) \\
g_{R/G/B}(i,j,k+1) &= s_{R/G/B}(i,j,m) - A \cdot (2 \cdot d(i,j,m) - 1)
\end{aligned}
\tag{4}
$$

An example of a modulated frame sequence is shown in Fig. 3. Adding the modulated data (here with *A=4*) to the video input yields the display amplitudes in the right column.

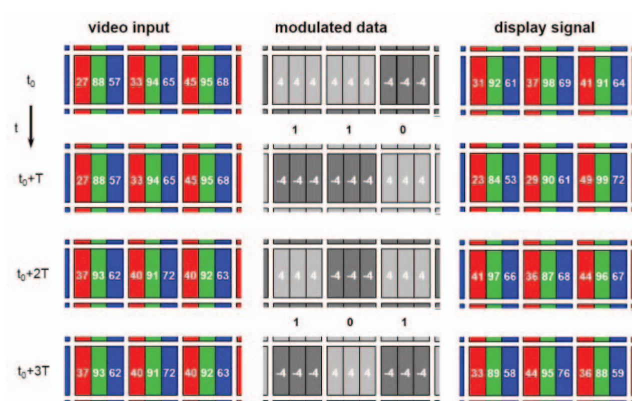


Fig. 3. Example of display amplitudes in temporal differential modulation

Direct assignment of one bit to one single screen pixel provides a high data rate. Assuming a display frame rate of 100 Hz and a full HD display, the resulting data rate without

978-1-4799-8748-1/15/$31.00©2015IEEE

overhead for synchronization and error correction can be as high as

$$1920 \cdot 1080 \cdot 50b/s = 103{,}68Mb/s. \qquad (5)$$

Certainly, this kind of modulation requires a high camera resolution. Due to the sampling theorem, even under ideal conditions the camera has to provide more than 8.3 Megapixels for luminance recording. In order to relax the camera requirements, modulation can be restricted to a part of the screen. Another approach is a spatial resampling and filtering of the overlaid data, as introduced by the signal processing blocks in Fig. 2.

A straightforward concept is the assignment of each data bit to a block of $B_X$ x $B_Y$ display pixels. In this case, equation (2) for data assignment is replaced by:

$$\begin{aligned} &d(l) \rightarrow d(i,j,k) \qquad 0 \le l < L \\ &L = \lfloor N_X / B_X \rfloor \cdot \lfloor N_Y / B_Y \rfloor \\ &i = (l \cdot B_X) \bmod N_x + r_X, \quad r_X = 0...(B_X - 1) \\ &j = \lfloor l / N_X \rfloor \cdot B_Y + r_Y, \; r_Y = 0...(B_Y - 1) \end{aligned} \qquad (6)$$

As there is no principle reason to select different block sizes for both dimensions, the evaluation in section IV of this paper have been done for square blocks:

$$B_x = B_y = B. \qquad (7)$$

If a slice of data for transmission does contain less than $L$ bits, bit stuffing should be performed.

## III. System Model

Up to this point, the key elements of the transmission scheme have been described in a pragmatic way. First experiments using a PC as display signal generator, a standard computer LCD screen as transmitter display and high quality cameras have shown that the concept is feasible if filter design and synchronization procedures are well designed.

In order to support this design process and evaluate the influence of all components and parameters, a system model has been developed. On principle, the system is a massive parallel transmission of temporal data sequences that might be impaired by noise, spatial crosstalk of data, interference from overlaid video reproduction as well as temporal inter-symbol-interference.

Concerning temporal ISI, a temporal synchronization and suitable selection of shutter time can help to avoid any problems. Measurements of recent displays have shown that typical response times are sufficient to avoid problems of temporal ISI, as long as no further display internal processing is involved. This means that the calculated display input *g(i,j,k)* has to be fed to the display panel directly. Fig. 4 gives an example of a measurement of a LCD screen. In this case, a frame rate of 60 Hz was applied. The differential modulation of the video sequence by the data can be easily detected.

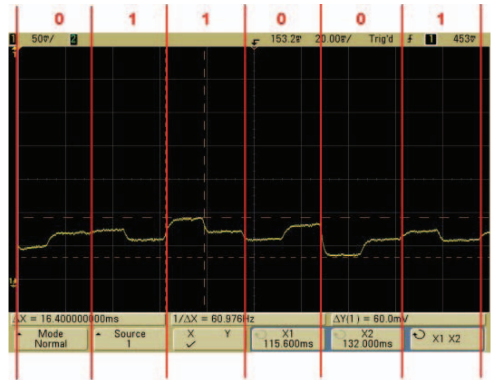


Fig. 4. Measured display luminance as a function of time when performing temporal differential modulation on a computer LCD monitor (frame rate 60 Hz). Input data of measured pixels are marked in red.

Assuming a synchronous recording using a shutter speed of 10ms should avoid any temporal inter-symbol interference. Certainly, ISI can be minimized by selecting shorter exposure times. However, this comes at the cost of reduced signal to noise ratio. A compromise has to be found for the involved display technology.

The spatial multiplex in this massive parallel transmission system is much more complex to model, as described in the following. While in the preceding sections, pixels have been considered as discrete elements and addressed by indices *i* and *j*, respectively, continuous spatial coordinates *x* and *y* are now introduced in the description. For simplification, the same letters are used for the coordinate system of the transmitter as well as that of the receiver. In reality, the camera lens performs a transformation between both coordinate systems. As this paper focusses on principle aspects of this complex transmission system, the transformation process is neglected in the formal description. Fig. 5 describes the transmitter side for transmission of one frame. By Fourier transformation, the system can also be analyzed in the frequency domain.

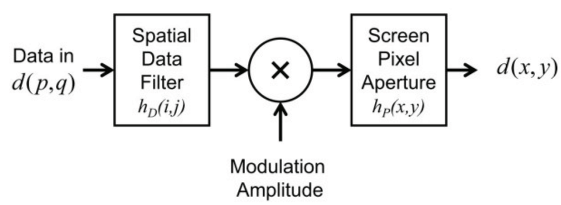


Fig. 5. Spatial system model of a display as data transmitter

Input data is assigned to the two-dimensional plane of the frame. As the sampling pattern of the data can be different to the pixel pattern (like in the data transmission in blocks of pixels as described above), indices p and q are introduced here. The spatial data filter, characterized by its impulse response $h_D(i,j)$, performs resampling to the display pixel pattern and spectral filtering. The resulting data amplitudes are multiplied by the modulation amplitude, which is selected as a compromise of visibility of the data pattern and system robustness. Due to the nonlinearity of typical displays, which is defined by the gradation function, there will be a certain amount of crosstalk coming from the video signal, influencing the modulation amplitude. The resulting amplitudes of the discrete pixels are transferred to a spatially continuous luminance density by the subpixels for R, G and B. This process is described by the impulse response $h_P(x,y)$, which is given by the pixel geometry of the involved display. As modulation of the different primaries is done, and the pixel positions for R, G, and B usually differ due to the spatial color multiplex in flat screens, the system model has to represent the exact geometry of the display.

The system model for the receiver side, as given in Fig. 6, is more complex. For the sake of simplicity, the same symbols are used for indices and coordinates, although in practice transmitter and receiver will apply different geometry and sampling patterns.

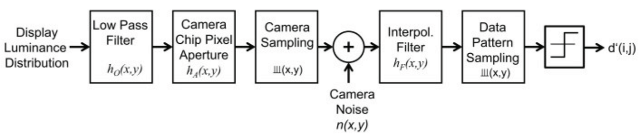


Fig. 6. Spatial system model of a camera as data receiver

The luminance distribution of the screen is picked up by the lens and projected onto the camera chip. The camera lens acts as a spatial low pass filter which is described by its spatial impulse response $h_O(x,y)$. As the conversion between luminance and the electrical signal is almost linear, this process does not influence the abstract model. Therefore, no functional block is introduced here. The aperture of the pixel elements in the camera chip, characterized by the aperture function $h_A(x,y)$, reflects the geometry of the camera. As we consider single chip consumer cameras, the mosaic pattern also has to be considered when modelling the characteristics. The camera chip performs a two dimensional spatial sampling of the signal, represented by the following functional block in the figure, that may cause aliasing. Camera noise is added to the signal of each pixel individually. As the generated noise, mainly generated by dark current and shot noise, depends on the total exposure of the pixel, the overlaid video signal influences the noise power of each pixel.

In the next processing step, the camera signal will be processed by an interpolation filter specified by the impulse response $h_F(x,y)$. This filter is a key element in the processing chain. It has to be developed such that the following two-dimensional sampling delivers optimum signal amplitudes for the decider.

Considering this model, it is obvious that fine data patterns require higher camera resolution, more precise filtering and lower noise, compared to coarser spatial data patterns. Certainly, visibility of the overlaid modulation in the video presentation increases with higher video amplitudes or coarser data patterns. In order to evaluate a set of suitable operational parameters, subjective tests of perceived picture quality have been carried out.

## IV. Picture Quality Evaluation

The perceived quality of video scenes with overlaid data transmission has been evaluated in subjective tests. Several sets of parameters for the modulation amplitude in combination with different block sizes for the data structure have been presented to subjects. Three block sizes were selected, namely data blocks of 1x1, 2x2 and 3x3 pixels. A temporal differential modulation of the luminance was carried out, using the same modulation amplitude A for R, G and B. A was varied between 2 and 7, based on an eight-bit representation of pixel amplitudes at the display input. The results of this test series are shown in Fig. 7. The viewing distance was three times picture height. In combination with the display frame rate of only 50 Hz, the results can be regarded as a lower bound for perceived picture quality, because visibility of artifacts decreases with greater viewing distance and increased frame rates. Furthermore, it can be expected that the simple data repetition scheme for block generation provides a picture quality which is worse compared to performance when using an optimized low pass interpolation filter.

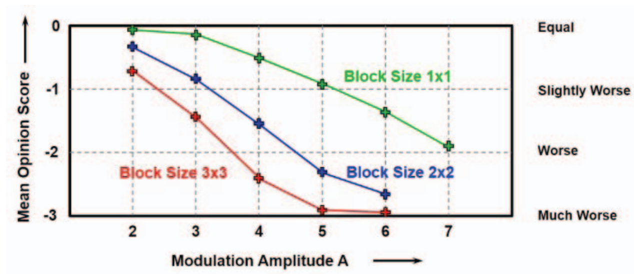


Fig. 7. Subjective evaluation of perceived video quality for different modulation parameters in temporal differential modulation

Obviously, even under these critical evaluation conditions modulation amplitudes of *A=4* appear to be acceptable when applying a direct mapping of data to pixels, i.e. *1x1* blocks. If larger blocks are used, the acceptable modulation amplitudes are lower. In case of *2x2* blocks, *A=3* appears to be the upper limit, while for *3x3* blocks, modulation amplitude should be limited to *A=2*.

## V. Conclusion and Future Work

Visible light communication using video displays and cameras is a promising approach for a variety of applications where data transmission is provided parallel to a video reproduction. Data rates of 100 Mb/s appear to be feasible for technologies available today. Subjective tests have shown that modulation parameters can be selected such that a reliable transmission is possible while perceived video quality is degraded only slightly.

The design of a suitable demodulation process still needs further research. Temporal and spatial synchronization as well as high quality interpolation filters to compensate geometric distortions caused by misalignment of the camera are key aspects of future work. The adaption of the modulation and decoding schemes to different application scenarios - from personal communication with small distances between transmitter and receiver to large multicast scenarios in public areas - is a demanding task. Future work should also provide input to standardization initiatives.

## REFERENCES


[1] IEEE Std. 802.15.7-2011, “Short-Range Wireless Optical Communication Using Visible Light,”, Sept. 2011

[2] J. Gancarz, H. Elgala, and Th. D.C. Little, “Impact of Lighting Requirements on VLC Systems,” IEEE Communications Magazine, Vol. 51, No. 12, pp. 34-41, Dec. 2013

[3] L. Grobe et al., "High Speed Visible Light Communication Systems," IEEE Communications Magazine, Vol. 51, No. 12, pp.60-66, Dec. 2013

[4] C. Danakis, M. Afgani, G. Povey, I. Underwood, and H. Haas, “Using a CMOS Camera Sensor for Visible Light Communication,” Proc. 3rd IEEE Workshop on Optical Wireless Communications, pp. 1244-1248, Dec. 2012

[5] S.D. Perli, N. Ahmed, and D. Katabi. "PixNet: Interference-free wireless links using lcd-camera pairs," 13th International Conference on Mobile Computing and Networking (MobiCom 2010), Chicago, USA, pp 137-148, Sept. 2010

[6] T. Hao, R. Zhou, and G.Xing. "Cobra: Color Barcode Streaming for Smartphone Systems," 10th International Conference on Mobile Systems, Applications and Services (MobiSys 2012), Low Wood Bay, Lake District, UK, pp. 85-98, June 2012

[7] W. Hu, H. Gu, and Q. Pu, "Lightsync: Unsynchronized Visual Communication over Screen-Camera Links," 16th International Conference on Mobile Computing and Networking (MobiCom 2013), Miami, USA, pp 15-26, Sept. 2013

[8] R. Kays, "Visible Light Communication Using TV Displays and Video Cameras," Proc. IEEE International Conference on Consumer Electronics, Las Vegas, USA, pp. 582-583, Jan. 2015.