Northumbria
University
NEWCASTLE

Smartphone Camera Based Visible Light Communication

R. Boubezari[1], H. Le Minh[1], Z. Ghassemlooy[1] and A. Bouridane[2]
[1]Optical Communications Research Group, [2]Computer and Electronic Security Systems, Faculty of Engineering and Environment, Northumbria University, Newcastle upon Tyne, NE1 8ST, UK
[rayana.boubezari,hoa.le-minh,z.ghassemlooy,ahmed.bouridane]@northumbria.ac.uk

Abstract— The paper proposes a novel camera-based receiver for visible light communications for a short range mobile-to-mobile communications link. The receiver captures data from the screen of a transmitting smartphone and uses the speeded up robust features algorithm to effectively detect it. The receiver performs a projective transformation to accurately eliminate perspective distortions caused by the displacement of the devices. The paper also introduces a quantization process in order to suppress the inter-symbol interference resulting from the dynamic nature of the environment. A range of experiments are carried out in order to evaluate the system performance when the position parameters are varied. We show that the proposed system is capable of achieving a very high success rate of 98% in recovering the transmitted images under test conditions.

Keywords—Image processing; Projective transformation; Smartphone; SURF; VLC

I. Introduction

RADIO frequency (RF) technologies are extensively used in wireless based communications systems. RF applications vary from short range indoor communications such as near field communications (NFC), Wi-Fi or Bluetooth to mobile cellular networks [1] up to medium-to-long range links such as millimeter wave backhauling [2] and satellite-to-Earth communications [3], respectively. The explosion in mobile devices usage over the last decade has led to the ever increasing usage of the limited and expensive radio spectrum. For instance, wireless and mobile devices will exceed traffic from wired devices by 2016 (54%) [4]. In Europe, the total frequency spectrum for public mobile communications is now over 500 MHz [5], with the additional spectrum becoming available in the future at higher-frequency carriers. The spectrum has become rather a limited resource when considering the rapid growth in the number of interconnected mobile users and the emergence of Internet of things (IoT) paradigm. RF based technologies are facing a substantial demand for the capacity that simply cannot be met with current spectral allocation [6], therefore new solutions must be found.

One of the key candidates to solve the so-called capacity crunch in certain applications is the visible light communications (VLC), which potentially offers 10,000 times more bandwidth capacity than the RF based technologies [7]. The VLC technology is mostly based around the intensity modulation of white light emitting diodes (LEDs) (in-organic [8] and organic [9]), which can be switched on and off at a very high rate [10], thus enabling data communications, illuminations, sensing as well as indoor localization [11]. LEDs are widely used in everyday infrastructures including homes, offices, street and traffic lights and smartphones [12]. In display devices such as smartphones or computers, the individual elements of the pixel arrays can be independently modulated [9] and captured with a camera (i.e., optical camera communications) in order to recover the transmitted information [13].

Current smartphones have such capabilities in detecting information. For example, in [14] a CMOS camera sensor was used to capture the data from visible LED light source. A similar approach was proposed in [15] where a message was encoded using Red, Green and Blue (RGB) LEDs in order to improve the data rate. In [16], an experimental screen-to-camera VLC system was reported, where a 256-LED transmitter and a commercial webcam were used as the transmitter (Tx) and receiver (Rx), respectively. In order to detect LED transmission on the captured frames, four LEDs were used as a reference. To locate the reference LEDs, an edge detection scheme based on Hough transform was used in [17]. In [18], the capacity of smartphone camera (SC) communications considering the inherent distortions and varying the distance between the Tx and the Rx was investigated. However, in this work the interference between pixels was not considered in the analysis. In [19], a camera based VLC link was demonstrated within an aircraft cabin, where successive 2D visual codes were displayed on the in-flight entertainment screen and captured by SC.

In this paper we investigate, for the first time, the performance of the SC link under different mobility conditions. This is a timely and important investigation as it takes advantage of the widely used screens, available in public places and personal devices. The system is portable as it can be readily implemented on smartphones, without any hardware modifications or additions. Any type of digital data can be converted into transferable images consisting of thousands of bits per image, which can then be displayed on the screen of a smartphone and captured by the camera of a second smartphone to recover the information within a single frame. Similar to a dynamic 2D barcode, it offers the possibility of transferring a significant amount of data by simply displaying it on the screen of a smartphone screen. The potential of such a technology is even higher if RGB images are used for transmission, since a single coloured cell is represented by three components instead of one.

One of the main issues in smartphone VLC (SVLC) is user mobility, which can introduce geometric distortions in the transmitted images. In this work, the user space was sampled at a number of spots chosen as representative of all possible positions of the mobile devices. The communication channel and system performance were then experimentally evaluated. Additionally, SVLC is an independent end-to-end communication system as it only relies on the screen and the camera on the hardware side, and signal and image processing

techniques on the software side, to transmit and receive data without the need for any other communication technologies.

The original contributions presented in this paper are (i) the use of the smartphone screen to display the digital information encoded in series of images and (ii) the use of the speeded up robust features (SURF) algorithm to detect information on laden images at the Rx. Additionally, in order to overcome perspective distortions resulting from the mobility of the Tx and the Rx, a projective transformation is applied. The user's movement can also cause considerable inter-symbol interference, which is mitigated using a novel quantization scheme based on the intensity of the pixels in received images.

In this work, an experimental investigation is carried out showing high transmission success rates as a function of the transmission distance and orientation angles of the smartphone.

The rest of the paper is organized as follows. Section II is the theory section is which explains the different computer vision techniques utilized in this system. The smartphone VLC system description is given in Section III. Experimental setup and results are outlined in Section IV followed by the conclusion and future work discussion in Section V.

II. THEORY

A. Computer vision – SURF (speeded up robust features)

The object detection is an important application in computer vision, which gives machines the ability to observe the environment, retrieve and analyze data, and make decisions accordingly. Object detection is a challenging topic especially in dynamic environments where a number of parameters need to be considered when selecting an algorithm. The most important parameter to be considered is the targeted object variations on different images since images can be taken under various conditions including illumination, scale, viewpoint and background clutter. For instance, recognizing an object at night can be more challenging than during the day. Other aspects can be challenging as well such as the viewpoint, scale, targeted object occlusion and background clutter [20]. Different types of object detection algorithms exist as outlined in [21]. One of the most robust classes uses the concept of keypoints, which describes specific properties of an image such as important changes in intensity and orientation of pixels, T junctions, corners and others. The keypoints include a set of specific properties of the image such as the intensity and orientation. In this scheme, SURF algorithm is chosen since it is a local feature extractor and descriptor, invariant to scale and rotation of the tracked object that is being used in real time applications [22]. Keypoints detected using SURF are defined using a set of unique 64-element descriptors. Though descriptors are unique, two keypoints from different images, which point to the same area in the same object, have close descriptors. The similarity of the keypoints for the received and the original images determines the accuracy of the targeted object localization.

B. Binarization – Otsu method

Binarization is the process of converting a grayscale image into a black and white image using thresholding [23]. Two different thresholding methods applied separately are investigated in this work. A hard threshold selection will be explained in the system description section, and a soft thresholding method using Otsu method, which relies on the image histogram to compute an optimal threshold value [24] will also be discussed. Otsu method [25] is based on a statistical separation of image pixels in two classes in order to binarize the image. The threshold is calculated using moments of the first two orders, which are average ave and standard deviation std [25]. In addition, in order to obtain a result independent from the dimensions, the histogram H is normalized as:

$$H_i = \frac{n_i}{N}\ , \tag{1}$$

where n_i represents the number of pixels at the illumination level i in the image and N is the total number of pixels on the image. The average and standard deviation are defined by:

$$ave(k) = \sum_{i=1}^{k} i \times H_i, \tag{2}$$

$$std(k) = \sum_{i=1}^{k} H_i\,. \tag{3}$$

For each value of $k = 1, \dots, 255$, s^2 is calculated:

$$s^2(k) = std(k).\big(1 - std(k)\big).\big(ave(255).std(k) - ave(k)\big)^2. \tag{4}$$

The value of k that maximises the function s is considered as the optimal threshold value to binarize the image. This is because it maximizes the separability of the resulting classes in the gray levels [25]. The threshold k can be determined from:

$$s^2(k) = \max\big(s^2(k)\big). \tag{5}$$

C. Projective transformation

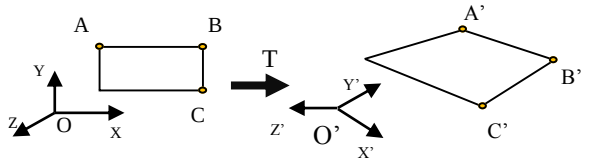


Fig. 1. Rectangle distorted with a projective transformation

In Fig. 1, "A" is a point in the O_{XYZ} frame of reference and A' is its projective transformation in the $O_{X'Y'Z'}$ frame of reference, where $A = TA'$. The points A and A' homogeneous coordinates are (A_X, A_Y, A_Z) and $(A'_{X'}, A'_{Y'}, A'_{Z'})$, respectively. The transformation matrix $\mathbf{T} = [t_{ij}]$, which is an invertible 3×3 matrix and the relationship between the coordinates of both points is given by:

$$\begin{bmatrix} A_x \\ A_y \\ A_z \end{bmatrix} = \begin{bmatrix} t_{11} & t_{12} & t_{13} \\ t_{21} & t_{22} & t_{23} \\ t_{31} & t_{32} & t_{33} \end{bmatrix} \times \begin{bmatrix} A_x' \\ A_y' \\ A_z' \end{bmatrix}. \tag{6}$$

In order to determine all elements of **T** matrix, the three points A, B and C and their respective projective transformation A', B' and C' are required.

III. SYSTEM DESCRIPTION

Fig. 2 shows a schematic diagram of a SVLC, which is composed of a pair of transmitting and receiving smartphones. The transmitting phone uses the screen for transmission and

the camera of the receiving phone is used for capturing the emitted signal from the Tx. The emitted light beam from the Tx is an encoded image where the black and white cells represent the information. The camera at the Rx captures an image provided it is within its receiving field of view (FoV).

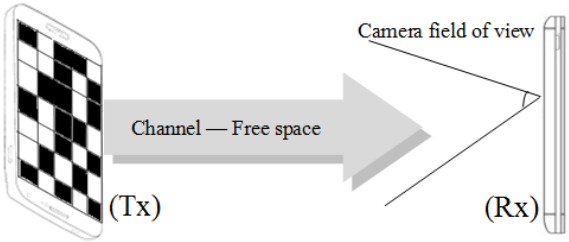


Fig. 2. A typical smartphone VLC system

1. Transmitter

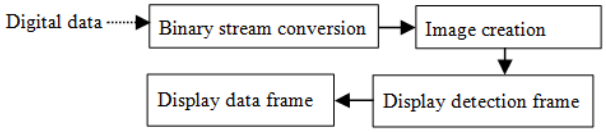


Fig. 3. SVLC Tx block diagram

The Tx's block diagram is depicted in Fig. 3. Initially the input information is converted into the corresponding binary stream, which is then used to create an image composed of a number of cells, where each cell contains a number of pixels. Pixels within the same cell are assigned the same intensity. The mage creation step is equivalent to signal modulation in communication. Three ways to modulate the signal are considered:

a. Black/white cells: The colour of each cell is either black or white and represents a single data bit. The output of the modulation is a matrix of M × N cells, which are assigned an intensity level of either 0 or 255.
b. Grayscale cells: For the grayscale option, every consecutive 8-bit binary pattern is represented by a decimal value between 0 and 255. The output is also a $M \times N$ matrix.
c. RGB cells: In the colour format, pixels are represented by three different colour components (i.e., RGB); with the components of each pixel assigned an intensity level between 0 and 255. Therefore, each cell representing 24-bit input data stream. The output is three-dimensional matrix of M × N × 3.

The dimensions of the generated matrices depend on the device screen size. In the case of a considerably long data stream, the data will be divided into smaller streams prior to conversion into a series of images. While the RGB cells method offers higher data transmission capacity, a simpler to implement black/white cell method was adopted, since it was considered to be more appropriate for investigation of the feasibility of SVLC under dynamic conditions. In order to ensure accurate detection at the Rx, a detection frame (i.e., a pilot signal) known by both the Tx and the Rx is first displayed, followed by a series of data images, see Fig. 4. The detection frame is periodically transmitted with the transmission interval depending on the degree of mobility of the Tx and the Rx.

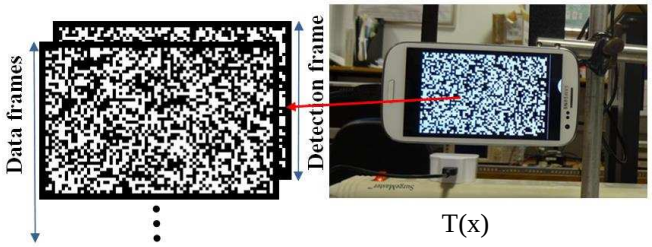


Fig. 4. Detection and data frames at the Tx

2. Receiver

The Rx's flowchart is represented in Fig. 5. The signal acquisition consists of capturing both the detection and the data frames, successively. The SURF algorithm is then applied on the original detection frame and the one captured by the Rx. Keypoints (kps) are then detected on both of them and their descriptors are calculated.

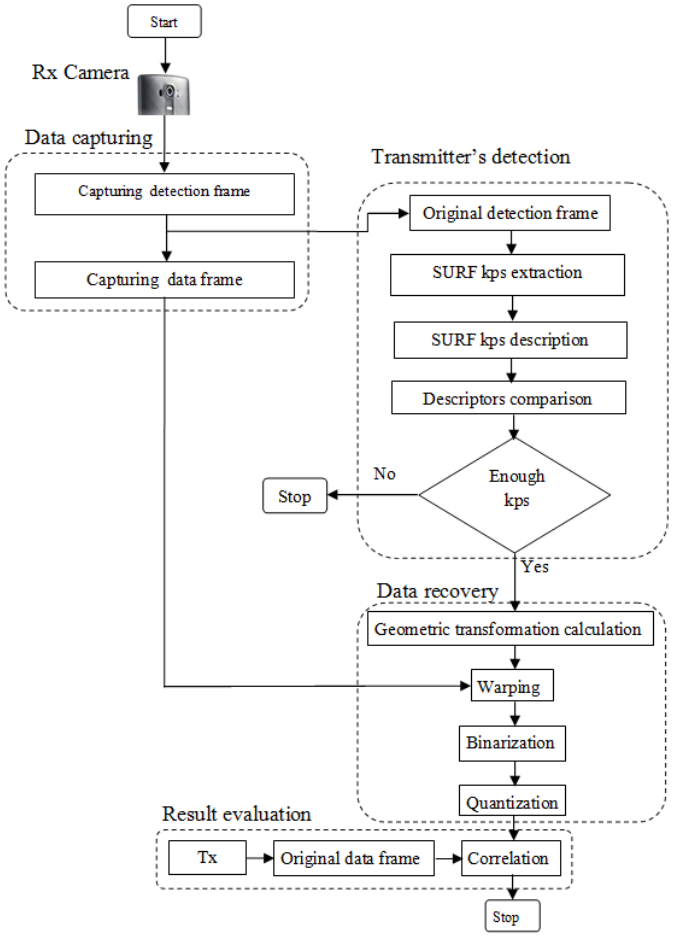


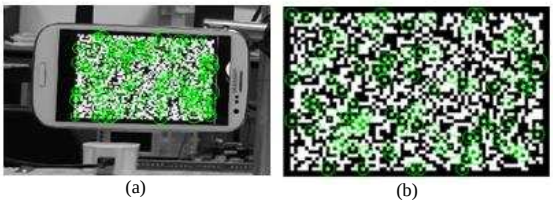
Fig. 5. SVLC Receiver's flowchart

These are then compared using the Euclidian distance and the matched keypoints are used for the projective transformation calculation. The following steps consist of the data frame warping, binarization and quantization. Finally, the received data is compared to the originally transmitted frame

by determining the correlation between them. The following steps consist of the data frame warping, binarization and quantization. Finally, the received data is compared to the originally transmitted frame by determining the correlation between them.

a) Tx detection, path loss and background noise

Similar to other communication systems, the transmission path loss and the noise due to the ambient light and other sources will affect the link performance. Therefore, a robust detection technique at the Rx is required and the SURF algorithm is adopted. Since SURF is invariant to the scale, detecting the Tx using this algorithm allows longer range communications. The accuracy and uniqueness of the descriptor calculated for each detected keypoint ensure the Tx's detection in any environment, regardless of the background objects. Fig. 6 depict how keypoints are detected on both the original and received detection frames.



(a) (b)

Fig. 6. SURF keypoints determination on the detection frame: (a) at the receiver, and (b) on the original image

b) Warping and perspective distortions

Perspective distortions occur when capturing the frame while the users are moving, which results in distortions of the received signal, see Fig. 7.

(a) Transmitted image (b) Received image

Fig. 7. (a) Transmitted image, and (b) perspective distortions on the received image

The selected keypoints from the pre-learnt detection frame and the captured frames are used to compute the projective transformation matrix, which is then used to correct the distorted image to its original format, see Fig. 7(a).

c) Binarization and intersymbol interference (ISI)

Depending on the number and size of cells in the transmitted image, the ambient lux level and the distance between the Tx and the Rx, the lines separating the recorded cells can fade away, thus resulting in ISI leading to poor cell recovery. Therefore, a binarization of all the pixels in the image is required prior to the cell quantization in order to reduce the ISI effect on the received signal. The emitted data is represented using the black and white cells in the images. However, standard colour cameras capture RGB images, where each pixel is represented by three colour-intensity components. The grayscale conversion step returns one component by pixel but the illumination level will vary between 0 and 255. Since only values 0 and 255 (white and black) were originally transmitted, then values in between need to be processed to restore the only two above states with minimum errors. Initially, a hard threshold binarization method was implemented. The details of this method are given below. The threshold T_b represents the theoretic mean value of the cells intensity, which is $T_b = 127$.

Given I the received image of size M × N:

$$\text{If } I(i,j) > T_b \text{ then } I(i,j) = 255 \quad \text{else } I(i,j) = 0, \tag{7}$$

where $0 < i \leq M$ and $0 < j \leq N$.

However, hard thresholding selection does not take into account the different lighting conditions, thus T_b might not be ideal under a different ambient light. In fact, the binarization threshold and the link performance are strongly dependent on the ambient light. In order to make the system independent and efficient under various lighting conditions, the reliability of the automatic thresholding scheme based on Otsu method is tested and compared to the hard thresholding as described above. The binarized image is used for cell quantization and reconstruction prior to recovery of the bit stream.

d) Quantization

Assuming that the total number of cells is known in advance, the Rx will scan through the received frame left-to-right, up and down in order to reconstruct the cells. Each cell is checked individually. Within each cell, if 50% or more of the pixels is black, then the remaining pixels will be converted into black. On the other hand, the whole cell will be converted into white if white pixels represent 50% of the total pixels within the cell or more. This process is illustrated in Fig. 8, where a cell is highlighted before and after quantization.

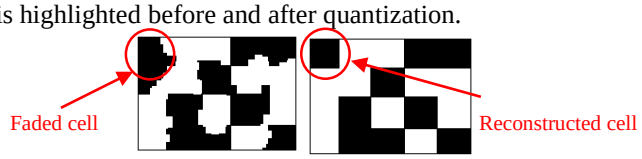


Fig. 8. Before and after quantization

IV. Comparison with Previous Work

In this section, SVLC is compared to similar systems, described in Visual MIMO [18], Pixnet [26] and COBRA [27]. Some key properties are listed in Table 1.

All the systems compared in this section use standard cameras. Although the distance at which Pixnet and Visual MIMO operate is 10 times greater than the one in SVLC, both screens used for transmission are of a bigger size. In COBRA, smartphones are used for transmission and reception. Coloured blocks are used where one pixel is actually represented by three components and hence, three channels and the optimum number of cells are higher than the maximum experimented with SVLC. However, the system operates at shorter distance, which reduces the mobility.

Table 1. Comparison of SVLC with existing systems

	Pixnet	**Cobra**	**Visual MIMO**	**SVLC**
Transmitter	32 inch LCD screen	5inch smartphone screen	21.5inc LCD screen	4.8 inch smartphone screen
Receiver	Off-the shelf camera	smartphone camera	Tablet camera	smartphone camera

Cells	Black/ white	Coloured	Black/white	Black/white
Distance	Up to 14m	5inch	5m	Up to 1.2 m
Number of cells	81x81	800x400	15x15	Up to 170x120

V. Experimental Set Up

Fig. 9(a) illustrates the tilted position set-up whereas Fig. 9(b) shows the rotation angle and the distance . The Tx is tilted with a certain angle ranging from 0° to 75° where 0° is the initial position (i.e., both smartphones are in parallel planes). Firstly, random binary streams were generated using Matlab, and then converted into black and white images prior to loading to the smartphones. The cells in the data frames are equally distributed in colour between black and white since the binary 0 and 1 are equally generated during the image formation process. The transmitting smartphone was used to display the data matrix as a black/white image whilst the camera of the receiving device captured the image for further processing using Matlab.

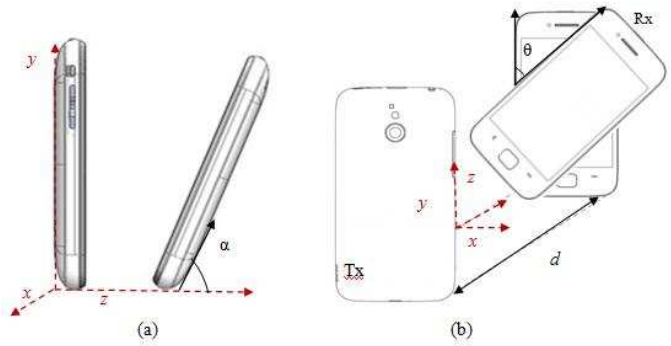


Fig. 9. Mobile VLC: (a) distance and tilting angle, and (b) the rotation angle

Measurements were carried out by varying the transmission distance d, rotation angle θ , and tilting angle α. In the setup (see Fig. 9), the Tx was located at a fixed position, whereas the Rx was moving along the z-axis. The built-in camera of the Rx captured the displayed data on the Tx. A captured image background may include other irrelevant objects located around the Tx. Both the Tx and Rx were exposed to the normal office lighting conditions using fluorescent lamps (i.e., the ambient background noise). At the Rx, Samsung Galaxy S3 smartphone back and front cameras were used to capture the image from the Tx mobile. Additionally, the detachable Sony QX-10 camera was used, which is a high quality and very high resolution camera, as a reference camera to compare with the performance of Galaxy S3 camera. No optical image stabilizer (OIS) was used in either S3 camera and the snapshots from both cameras were taken under the same test conditions for comparison purposes. Table 2 shows the key parameters adopted for the experimental setup. The captured image from smartphone camera is further processed using the above algorithms developed in Matlab to recover the information data stream, which is compared with the originally transmitted signal.

Table 2. Experiment parameters

Parameter	**Value**
Number of cells ($M \times N$)	3750 (75 × 50)
Number of pixels per cell	10 × 10
Number of cell colour states	2 (black and white)
Tx smartphone screen resolution (Samsung S3)	1280 × 720 (4.8 inches HD super AMOLED)
Rx smartphone front camera (Samsung S3) effective pixel	1.9 MP
Rx smartphone back camera (Samsung S3) effective pixel	8 MP
Reference camera (Sony QX-10) effective pixels	18 MP
Link length d	10 to 110 cm
Parallel rotation angle θ	0° to 360°
Tilting angle α	0° to 75°
Measured ambient level (lux) (Normal office lighting)	570 lux (Note: ISO office standard lighting level is ~400 lux)

The system performance was evaluated in terms of the success rate (i.e., 1 – bit error rate) by means of correlation of the transmitted and the received data. Five images were taken for every single position and the average of the bit success rate of these five samples was calculated. Note that, no error correction was introduced in the system.

VI. Results and Discussion

Fig. 10 shows a comparison of the data recovery success rate against the transmission range for three different cameras. The Samsung back camera reaches 98% data recovery over the short range (10 – 50 cm). Over longer transmission range, the bit success rate drops to ~60% until the detection is no longer possible beyond 70 cm. The Samsung front camera achieves a data recovery greater than 85% within the range of 10 to 25 cm. The performance is drastically decreased since the number of keypoints is not sufficient for the Tx's detection when the distance is over 30cm. In contrast, the reference camera Sony QX-10 has a much higher resolution (18 Megapixel) and a better quality optical lens. In this case the communications range can be extended up to 110 cm. At this distance the system performance is highly dependent on the camera quality (resolution and lens), even though the bit success rate is somewhat decreased. The bit success rate at shorter distances (10 – 60 cm) are very satisfactory throughout, and at these transmission distances similar results can be obtained using the back camera of the smartphone, both being in the range of 85% and 100%.

In order to binarize the data correctly regardless of the background light and the Tx's screen illumination, an automatic solution for the optimal threshold calculation was developed and implemented. Here the Otsu method was adopted for optimising the system performance [15]. To evaluate the reliability of this method, the performance of Otsu method, the soft decision thresholding method (ST) was compared to the hard decision thresholding method (HT). The

captured images are washed out when the Tx was exposed to the ambient light at the level of 570 lux which is in the standard office lighting condition. As a result, the black colour was brighter and the process with HT tended to favour it over the "white colour". Consequently, the ST value should be closer to 255 and greater than the hard threshold T_b to allow more accurate separation between cells of different colours with improved recovery. Overall, the Otsu method showed a better performance compared to the HT scheme. The second geometrical test was carried out by varying the angle of rotation of the Tx where both smartphones are in parallel planes in order to evaluate the data recovery performance at the Rx, see Fig. 11.

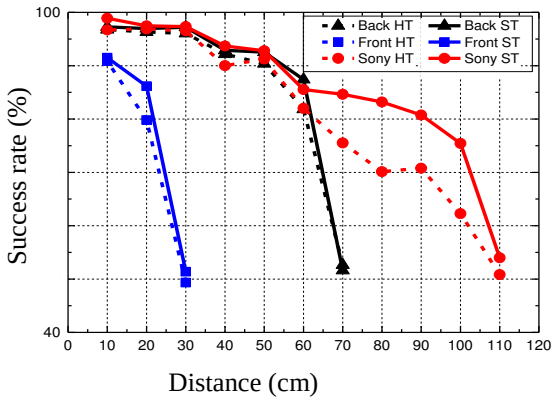


Fig. 10. Success rate of data recovery against the link distance d for hard thresholding (HT) and soft thresholding (ST) schemes

For all of the three cameras the distance d was kept at 15 cm since all cameras displayed high performances at this distance, thus enabling the investigation of the effect of the rotation angle independently. Both Sony and Samsung back cameras displayed similar performances where both success rates are near 100% over the transmission range of 0° to 360°.

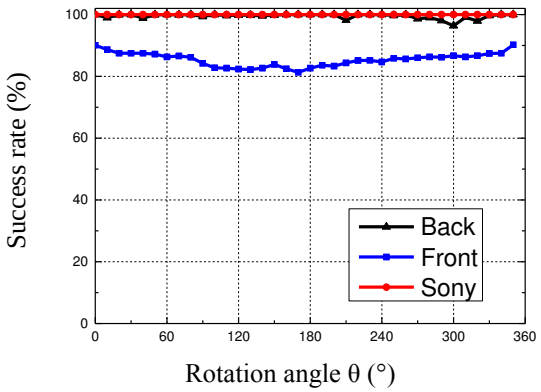


Fig. 11. The successful rate of data recovery as a function of the rotation angle θ

The Samsung front camera showed fairly good performance with slightly decreased success rate when the angle increased towards 180° and reaching 90% success rate when the angle is closer to 360°. The system performance against the tilted positions of the Tx was further investigated (see Fig. 12). Both Sony camera and Samsung back camera achieved improved results within the transmission range of 0 and 50°. The performance does drop slightly at 60°. The Tx is not detected beyond the angle of 60° as the number of keypoints is not sufficient. The Samsung front camera is less robust but achieves acceptable results only within the range of 0° to 20°. These results again indicate the importance of camera quality on the system performance. The results also show that the use of Otsu method is more advantageous than the HT method. It is to be noted that new generation of Smartphones will offer higher resolution cameras and screens, thus offering higher potential when adopted in SVLC.

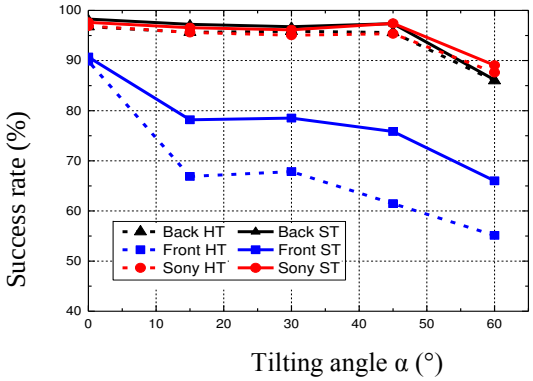


Fig. 12. Success rate of data recovery against the tilting angle for hard thresholding (HT) and soft thresholding (ST) schemes

Next we carried out experiments by changing the number of cells displayed on the Tx's screen in order to extensively evaluate the performance of the Rx capability under different threshold schemes. The number of cells was varied from 3,000 to 100,000 over a 15 cm link distance. For Sony camera and Samsung Back camera, the performance decreases slightly when increasing the number of transmitted cells until it drops after 90,000 cells. However, the Samsung front camera is more affected whilst increasing the number of cells, and the recovery rate drops below 70% when transmitting over 30,000 cells.

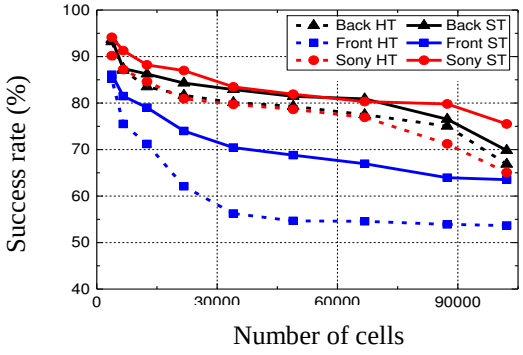


Fig. 13. Bit success rate while increasing the number of cells transmitted for hard thresholding (HT) and soft thresholding (ST) schemes

Note that 3750 black/white cells were transmitted per frame (i.e., 3750 bits). Given that a data rate $R = (FPS-1) \times M \times N \times \log_2(K)$, where FPS is the camera rate and one frame per second would be devoted for the detection frame. K is the

number of colours used in the data encoding. A data rate of 112.5kbit/s could be achieved with real time processing. However, R might be slower due to the need for synchronisation at the Rx. Using 10,000 cells and maintaining an 80% transmission rate, a three times higher R could be obtained with RGB images, thus offering ~300 kbit/s. It can be anticipated that R would be much higher with the new high quality camera phones. In addition, grayscale or RGB methods could be adopted instead of the black/white scheme.

VII. Conclusion And Future Works

This paper demonstrated a mobile VLC system by means of investigating transmission and detection of encoded images. SURF, projective transformation and quantization were appropriately used to realise objective detection. Tests were carried out under various position-displacement conditions for smartphone VLC. The proposed system offered the suppression of path loss, inter-symbol interference and perspective distortions. The achieved results outlined a reasonable performance for standard smartphones and provided a benchmark for short range data communications as evaluated against the Rx camera type.

In future the experimental work will be further extended to different mobile platforms. Different testing conditions (outdoor or sunlight), hand-shaking and fast movements will be investigated for a more extensive evaluation of SVLC.

References

[1] A. H. Sakr and E. Hossain, "Analysis of K-tier uplink cellular networks with ambient RF energy harvesting," Ieee Journal on Selected Areas in Communications, vol. 33, pp. 2226-2238, Oct 2015.

[2] C. Dehos, J. L. Gonzalez, A. De Domenico, D. Ktenas, and L. Dussopt, "Millimeter-wave access and backhauling: the solution to the exponential data traffic increase in 5G mobile communications systems?," IEEE Communications Magazine, vol. 52, pp. 88-95, 2014.

[3] T. Hirano, K. Miyazaki, M. Hatamoto, R. Yasumitsu, K. Hama, and F. Watanabe, "RF-MEMS for reconfigurablesatellite antenna," in Antennas and Propagation, 2007. EuCAP 2007. The Second European Conference on, 2007, pp. 1-5.

[4] Cisco, "The zettabyte era—Trends and analysis," ed, 2015.

[5] D.Lewin, P.Marks, and S.Nicoletti, "Valuing the use of spectrum in the EU," June 2013.

[6] W. Yuanquan and C. Nan, "A high-speed bi-directional visible light communication system based on RGB-LED," China Communications, vol. 11, pp. 40-44, 2014.

[7] D. A. Basnayaka and H. Haas, "Hybrid RF and VLC Systems: Improving user data rate performance of VLC systems," in Vehicular Technology Conference (VTC Spring), 2015 IEEE 81st, 2015, pp. 1-5.

[8] M. Hoa Le, D. O'Brien, G. Faulkner, Z. Lubin, L. Kyungwoo, J. Daekwang, et al., "100-Mb/s NRZ visible light communications using a postequalized white LED," IEEE Photonics Technology Letters, vol. 21, pp. 1063-1065, 2009.

[9] P. A. Haigh, F. Bausi, H. Le Minh, I. Papakonstantinou, W. O. Popoola, A. Burton, et al., "Wavelength-multiplexed polymer LEDs: Towards 55 mb/s organic visible light communications," Ieee Journal on Selected Areas in Communications, vol. 33, pp. 1819-1828, Sep 2015.

[10] D. O'Brien, L. Zeng, L.-M. Hoa, G. Faulkner, J. W. Walewski, and S. Randel, "Visible light communications: Challenges and possibilities," in Personal, Indoor and Mobile Radio Communications, 2008. PIMRC 2008. IEEE 19th International Symposium on, 2008, pp. 1-5.

[11] H. Le Minh, D. O'Brien, G. Faulkner, L. Zeng, K. Lee, D. Jung, et al., "High-speed visible light communications using multiple-resonant equalization," IEEE Photonics Technology Letters, vol. 20, pp. 1243-1245, 2008.

[12] Y. Wang, N. Chi, Y. Wang, L. Tao, and J. Shi, "Network architecture of a high-speed visible light communication local area network," IEEE Photonics Technology Letters, vol. 27, pp. 197-200, 2015.

[13] L. Pengfei, Z. Min, Z. Ghassemlooy, M. Hoa Le, T. Hsin-Mu, T. Xuan, et al., "Experimental demonstration of RGB LED-based optical camera communications," IEEE Photonics Journal, vol. 7, pp. 1-12, 2015.

[14] C. Danakis, M. Afgani, G. Povey, I. Underwood, and H. Haas, "Using a CMOS camera sensor for visible light communication," in IEEEGlobecom Workshops (GC Wkshps) 2012, pp. 1244-1248.

[15] World.casio.com, "Casio unveils prototype of visible light communication system using smartphones at CES - 2012 - News - CASIO," ed, 2012.

[16] N. Trang, L. Nam Tuan, and J. Yeong Min, "Practical design of screen-to-camera based optical camera communication," in International Conference on Information Networking (ICOIN), 2015, pp. 369-374.

[17] F. Dai, N. Zhang, and J. Xue, "Primal sketch of images based on empirical mode decomposition and Hough transform," in Industrial Electronics and Applications, 2008. ICIEA 2008. 3rd IEEE Conference on, 2008, pp. 2521-2524.

[18] A. Ashok, S. Jain, M. Gruteser, N. Mandayam, Y. Wenjia, and K. Dana, "Capacity of pervasive camera based communication under perspective distortions," in IEEE International Conference on Pervasive Computing and Communications (PerCom) 2014, pp. 112-120.

[19] T. Fath, F. Schubert, and H. Haas, "Wireless data transmission using visual codes," Photonics Research, vol. 2, pp. 150-160, Oct 2014.

[20] S. V. Kothiya and K. B. Mistree, "A review on real time object tracking in video sequences," in Electrical, Electronics, Signals, Communication and Optimization (EESCO), 2015 International Conference on, 2015, pp. 1-4.

[21] C. Schmid and R. Mohr, "Local grayvalue invariants for image retrieval," Ieee Transactions on Pattern Analysis and Machine Intelligence, vol. 19, pp. 530-535, May 1997.

[22] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, "ORB: An efficient alternative to SIFT or SURF," in IEEE International Conference on Computer Vision (ICCV), 2011, pp. 2564-2571.

[23] B. Su, S. Lu, and C. L. Tan, "Robust document image binarization technique for degraded document images," IEEE Trans Image Process, vol. 22, pp. 1408-17, Apr 2013.

[24] L. Dongju and Y. Jian, "Otsu method and K-means," in Ninth International Conference on Hybrid Intelligent Systems 2009, pp. 344-349.

[25] N. Otsu, "A threshold selection method from gray-level histograms," IEEE Transactions on Systems Man and Cybernetics, vol. 9, pp. 62-66, 1979.

[26] S. D. Perli, N. Ahmed, and D. Katabi, "PixNet: interference-free wireless links using LCD-camera pairs," presented at the Proceedings of the sixteenth annual international conference on Mobile computing and networking, Chicago, Illinois, USA, 2010.

[27] T. Hao, R. Zhou, and G. Xing, "COBRA: color barcode streaming for smartphone systems," presented at the Proceedings of the 10th international conference on Mobile systems, applications, and services, Low Wood Bay, Lake District, UK, 2012.