PixNet: Interference-Free Wireless Links Using LCD-Camera pairs

Samuel David Perli
CSAIL, MIT
samuelp@mit.edu

Nabeel Ahmed
CSAIL, MIT
nabeel@csail.mit.edu

Dina Katabi
CSAIL, MIT
dina@csail.mit.edu

ABSTRACT

Given the abundance of cameras and LCDs in today's environment, there exists an untapped opportunity for using these devices for communication. Specifically, cameras can tune to nearby LCDs and use them for network access. The key feature of these LCD-camera links is that they are highly directional and hence enable a form of interference-free wireless communication. This makes them an attractive technology for dense, high contention scenarios. The main challenge however, to enable such LCD-camera links is to maximize coverage, that is to deliver multiple Mb/s over multi-meter distances, independent of the view angle. To do so, these links need to address unique types of channel distortions, such as perspective distortion and blur.

This paper explores this novel communication medium and presents PixNet, a system for transmitting information over LCD-camera links. PixNet generalizes the popular OFDM transmission algorithms to address the unique characteristics of the LCD-camera link which include perspective distortion, blur, and sensitivity to ambient light. We have built a prototype of PixNet using off-the-shelf LCDs and cameras. An extensive evaluation shows that a single PixNet link delivers data rates of up to 12 Mb/s at a distance of 10 meters, and works with view angles as wide as 120°.

Categories and Subject Descriptors

C.2.1 [**Network Architecture and Design**]: Wireless communication

General Terms

Algorithms, Design, Experimentation, Measurement, Performance

Keywords

Optical Links, Camera, OFDM, Perspective Distortion

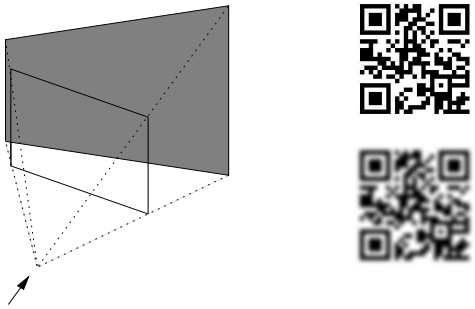
1. INTRODUCTION

Cameras and LCDs are abundant in today's environment, both in stand-alone form and embedded in laptops, smart phones, and PDAs. This abundance creates an untapped opportunity for using these devices for wireless communication. In particular, LCDs can encode data into visual frames, allowing camera-equipped devices to download this information. These LCD-camera links are highly directional due to the short wavelengths in the visible light spectrum. Hence, they can provide a form of interference-free wireless communication, that is a multitude of LCD-camera links can operate simultaneously in a dense area without interfering with each other. Consider for example a trade show, where companies demonstrate their latest products. Thousands of participants tour the booths, exchange business cards and brochures, and come back loaded with handouts and booklets. LCD-camera links enable any pair of participants to use their smart phones to wirelessly exchange brochures and even videos of their products, without having to worry about interference from the hundreds of concurrent communications going on simultaneously in the trade show. More generally, LCD-camera links can potentially evolve into a new wireless technology that is useful in dense high-contention scenarios, similar to how Bluetooth targets low-power scenarios, and whitespaces target long-range communication.

While they offer new opportunities, LCD-camera links bring about new challenges. Specifically, an LCD-camera link exhibits three main types of distortions:

- *Perspective distortion.* Since they operate in the visible light spectrum, LCD-camera links require line of sight. This requires the designer to pay special attention to where the LCDs are placed in order to maximize coverage. If an LCD and a camera can communicate only when they are perfectly aligned, the resulting system is less likely to be broadly applicable to different usage scenarios. In contrast, if an LCD and a camera can communicate in the presence of view angles, similar to how a human sees a screen even when he looks at it from an angle, coverage is significantly extended. The challenge is that the image of a rectangular screen becomes a trapezoid when viewed from an angle, as shown in Fig. 1(a). As a result, some pixels on the LCD screen expand at the camera, while others shrink. The more coverage one would like from an LCD-camera link, the greater perspective distortion these links must be able to tolerate.
- *Blur.* Any handshaking or movement while capturing an image or a lack of focus can introduce blur in the image, which causes the pixels to blend together, as in Fig. 1(b). An LCD-camera communication system

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
MobiCom'10, September 20–24, 2010, Chicago, Illinois, USA.
Copyright 2010 ACM 978-1-4503-0181-7/10/09 ...$10.00.



(a) Perspective Distortion (b) Blur

Figure 1: Example distortions of the LCD-Camera channel.

must be able to deal with such blending and still successfully recover the transmitted bits.

- *Ambient Light.* Ambient light is a source of noise for LCD-camera links because it changes the luminance of the received pixels. This can cause errors in the information encoded in the pixels, resulting in information loss at the receiver.

Thus, the LCD-camera channel needs a new transmission scheme that can handle the above distortions, that are significantly different from the distortions seen in RF wireless channels.

Past work in the area of computer graphics has looked at these problems in the context of 2D barcodes, e.g., QR code [2] or Data matrix [3]. These codes are printed on walls or objects. Users with a camera phone can take a picture of these barcodes, decode them, and obtain a description of the attached object or surrounding space [37, 38]. Barcodes however have relatively low information density and must be read at close proximity [18, 30]. In contrast, we aim to develop an LCD-camera link that supports high data rates at multi-meter distances and wide view angles.

This paper presents PixNet, a system for transmitting information over LCD-camera links. In contrast to all past work on 2D barcodes, which encode information directly in the visual domain, PixNet encodes information in the *frequency domain*. Such a design is inspired by the popular OFDM transmission scheme, widely used in modern RF technologies. However, unlike existing RF-based OFDM schemes that encode data in time frequencies, PixNet encodes data in two-dimensional spatial frequencies. More importantly, PixNet generalizes OFDM receiver algorithms to deal with the unique distortions of the LCD-camera link. Using PixNet we show that such a generalized frequency-based design provides a unified framework to deal with the distortions in the LCD-camera channel.

PixNet has the following three components:

(a) Perspective Correction Algorithm: A picture taken by a digital camera is a sampled version of the captured object. Perspective distortion occurs when the sampling frequency is irregular. For example, a rectangular screen becomes a trapezoid if the columns on the right are sampled at a lower frequency (i.e., with more pixels) than those on the left (Fig. 1(a)). Since PixNet operates in the frequency domain, it naturally addresses irregularity in the sampling frequencies. Specifically, PixNet generalizes the OFDM algorithm for correcting the sampling frequency offset (SFO) between sender and receiver to allow it to work with irregular sampling offsets. Once the receiver knows the sampling frequency offset in each part of the image, it re-samples the image at the right frequencies to correctly recover the bits encoded in the frame.

(b) Blur-Adaptive Coding: Approaches that encode bits directly in the visual domain, like 2D barcodes, fail quickly in the presence of blur because the bits blend together. In contrast, since PixNet encodes information in the frequency domain, it is more resilient to blur. Blur, in the frequency domain, translates into attenuation in the high frequencies while the low frequencies remain intact. Therefore, PixNet naturally identifies the frequencies affected by blur and prevents the error from spreading into other bits. PixNet treats different frequencies differently: frequencies that are badly attenuated due to blur are not used for transmitting information; frequencies mildly affected by blur are used for transmitting information but protected with a high redundancy error correction code; frequencies not affected by blur are used for transmission and protected with a low redundancy error correcting code.

(c) Ambient Light Filter: Approaches that encode information directly in the visual domain have to perform a special preprocessing step referred to as light balancing [22]. In contrast, PixNet operates in the frequency domain. Since ambient light changes the overall luminance, it only affects the zero frequency (commonly referred to as 'DC') that captures the average luminance of pixels in the image. Thus, PixNet can filter out the impact of ambient light simply by ignoring the DC frequency.

We have built a software prototype of PixNet and evaluated it using commodity LCDs and cameras. Empirical results using Dell 30 inch screens with consumer cameras such as Casio EX-F1 and Nikon D3X reveal the following findings:

- Using PixNet a single LCD-camera link can deliver data rates of up to 12 Mb/s, at a distance of 10 meters.
- PixNet's links support wide view angles. Specifically, PixNet delivers 8 Mb/s at view angles as wide as 120°.
- We also compare PixNet to a baseline system that uses the popular Quick Response (QR) 2D barcode and stacks as many of them as can fit in an LCD frame. Our results show that in comparison with QR codes, PixNet delivers up to $2x - 9x$ higher throughput depending on the distance and can also tolerate $3x$ wider view angles.

1.1 Contributions

- To the best of our knowledge, PixNet is the first system where LCD-camera links are shown in a working deployment to deliver high throughput data rates over multi-meter distances and wide view angles.
- PixNet presents a novel OFDM-based receiver algorithm that addresses perspective distortion.
- PixNet presents a blur-adaptive error correction code that is robust to distortions in the LCD-camera channel.
- We present an extensive empirical study of data communication over LCD-camera links that addresses distances, view angles, camera focus, and different ap-

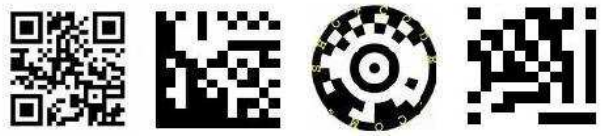


Figure 2: A few examples of 2D barcodes used in mobile tagging. From left: QR Codes, Data Matrix Codes, Shot Codes and EZ Codes.

proaches for encoding information, e.g., 2D OFDM and 2D barcodes.

2. RELATED WORK

Communication over visual channels has intrigued people for the last two decades. For instance, in 1994, Timex released the Data Link series of watches embedded with photodetectors to capture information transmitted visually via a CRT monitor [23]. This product however delivered only very low data rates (up to two bytes per frame). More recent forms of visual communication have employed barcodes, which are visual markers printed on commercial goods for stock control or painted on walls or roads to convey location information. Barcodes are typically decoded using flying spot scanning lasers and a single photodetector. This minimizes perspective distortions and blur [35]. Recent work on pervasive computing has developed a new class of 2D barcodes that are captured and processed using cell phone cameras. A few such barcodes are shown in Fig. 2 [2, 3, 4, 1]. Today there is a spectrum of 2D barcodes. Some 2D barcodes have been proposed for information storage [15]. Though they advertise high information density, due to their proprietary nature, little is known about their tolerance to blur, perspective distortion, and distances. Other 2D codes (e.g., QR codes, Data Matrix, and BoKodes) are shown to be robust to visual distortion and blur, but are limited in terms of their information density [27]. There has also been some work on time-multiplexing such 2D barcodes to allow decoding across multiple frames [19]. However, these schemes deliver only a few hundred characters per minute and do not deal with large distances or wide view angles. PixNet's design differs from all of the above as it encodes information in the frequency domain using spatial OFDM. As a result, it can achieve higher data rates and wider view angles as shown in §8.

PixNet builds on a proposal by Hranilovic and Kschischang [16] that also advocate using OFDM to transmit between an LCD-camera pair. Their approach however uses OFDM as-is and does not deal with perspective distortion, blur, or frame synchronization. As a result, it applies only to scenarios in which the camera and LCD are perfectly aligned and perfectly focused. Additionally, the work in [16] is based on emulation, i.e., the channel function is measured empirically but no data is transmitted. It is then applied to the OFDM encoded data in simulation. In contrast, PixNet generalizes OFDM's sampling correction algorithm to address perspective distortion, and augments it with a blur-adaptive error correction code. PixNet also presents the first implementation of an LCD-camera communication system that delivers high data rates over multi-meter distances and wide view angles.

PixNet is also related to work on free-space optics (FSO) and visible light communication (VLC). Free space optics (FSO) use laser beams that transmit light in the atmosphere without the use of physical fiber. This technology was originally developed by the military but has since been used in optical interconnects for on-board and cross-board communication [14, 25]. FSO, however, is not applicable in the end user scenarios we are targeting, because it requires almost perfect alignment of the transmitter and the receiver.

For indoor environments, visible light communication (VLC) has been proposed [29], where light sources (e.g., light bulbs) are modulated to enable data communication in addition to illumination [21, 26]. Such light sources are highly diffusive in nature, resulting in very low throughput communication. Although there has been some work on specialized sources and photodetectors [26] that address this problem, fundamentally VLC is not appropriate for the high bandwidth scenarios targeted by PixNet.

Last, there is extensive work by the imaging community that has looked at the problem of perspective distortion in the visual domain [10, 11, 41]. PixNet takes this work a step further; it shows for the first time that wireless transmission algorithms can naturally address the problem of perspective distortion. In particular, PixNet exploits OFDM that operates in the frequency domain to detect and correct such distortions. This approach is not only novel but also computationally more efficient when compared to existing perspective correction algorithms proposed in the literature [10, 11, 41].

3. PROBLEM DOMAIN

PixNet focuses on information transmission over LCD camera links, i.e., the development of modulation and coding schemes that provide high throughput over large distances and wide view angles. To put the scope of this work into perspective, we start by listing the assumptions under which we have developed PixNet.

- We consider scenarios where the camera is already setup to take pictures of the communicating LCD screen. In static environments the user may focus the camera on the screen. In dynamic scenarios, one can leverage prior work on steerable projector-camera systems [12, 34] and use automatic focusing. Note that this does not mean that the LCD and camera are aligned; the camera can still be viewing the LCD from any view angle.
- We focus on one-way communication from an LCD to a camera. Higher layer protocols and applications may require two-way communication (e.g., TCP Acks). A full-featured integration of LCD-camera links into a general purpose networking framework has to consider these issues in more detail. In this paper however we assume that Acks and uplink traffic are sent over the RF WLAN and focus only on one-way transmission.

4. PixNet'S TRANSMITTER

Analogous to the way that a standard OFDM transmitter translates bits into voltages, the PixNet transmitter translates bits into pixel luminance values on the LCD. There are two important differences however. First, RF communication uses high frequency carrier waves to transmit both real and imaginary components whereas pixel luminance values are purely real and support no imaginary components. Second, traditional OFDM is one dimensional but in contrast,

LCD screens are two dimensional (3D including the time dimension).

To handle these differences, PixNet generalizes the standard OFDM algorithm by making some key modifications. In the modified algorithm, bits are first modulated into complex numbers and then broken down into symbols just like standard OFDM. However, instead of feeding these into a one dimensional IFFT, each symbol is first arranged into a two dimensional Hermitian matrix. This matrix is then fed into a 2-D IFFT. The special properties of the Hermitian arrangement ensure that the output from the IFFT is entirely real. Additionally, by using a 2D IFFT we also get a two dimensional output appropriate for display on an LCD. To actually transmit the data, the transmitter stacks in each frame as many OFDM symbols as the LCD can support. It then transmits frame after frame until the transfer is complete.

4.1 Modulation

PixNet uses 4QAM modulation which produces a complex number for every pair of bits. Fig. 3(a) illustrates this process and shows the four complex numbers that correspond to all possible values of a sequence of 2 bits.

4.2 Maintaining Real Output

As mentioned above, PixNet ensures that the output of the transmitter is purely real by taking advantage of the intrinsic property of Fourier transforms, which says that if the input function is Hermitian, its IFFT is real.[1] To turn the modulated bits into Hermitian matrices, PixNet first breaks the stream of modulated values into symbols. For example, say we have a transmitter block of 5×5 pixels. We can use it to transmit 12 real data values as follows: We first organize the values into a Hermitian form, i.e., we assign to pixel $(-i, -j)$ the complex conjugate of the value assigned to pixel (i, j), as shown in Fig. 3(b). We then apply it to a 2D IFFT. The resulting 25 values are all real and can be sent as pixels' luminance. We refer to such a 2D block of encoded pixels as a 2D OFDM symbol.

One may think that, by transmitting only 12 complex numbers in 25 real values, we have reduced the efficiency to 50%. Recall however that each complex number is composed of two real numbers. Hence, 12 complex numbers already have 24 real values. As for the center frequency, OFDM naturally does not transmit data in the DC frequency. Furthermore, in PixNet the DC frequency is the average pixel luminance (by definition of IFFT). Upon reception, the average pixel luminance typically gets badly distorted due to the ambient light in the environment. PixNet filters the distortion caused by ambient light by simply ignoring the DC frequency.

4.3 Lack of pixel-to-pixel matching between the LCD and Camera

In addition to modifying OFDM to meet the special requirements imposed by the LCD-camera channel, PixNet takes advantage of the standard properties of OFDM to solve problems created by the LCD-camera channel. Specifically, there is no one-to-one mapping between the pixels on the LCD and the pixels on the camera. Each pixel on the LCD is a light source and light is diffusive in nature. The camera's lens attempts to concentrate the light from each LCD pixel onto a small region in the image. In practice, focus can never be perfect and hence, each camera pixel receives light from multiple nearby LCD pixels as shown in Fig. 4.[2] As a result, the transmitter pixels bleed into each other at the receiver. This bleeding of the pixels creates two effects on the received samples: First, pixels are subjected to interference from pixels belonging to neighboring symbols (inter-symbol interference). Second, pixels from the same symbol interfere with each other creating a blur effect (inter-pixel interference).

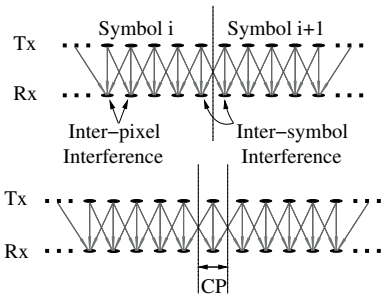


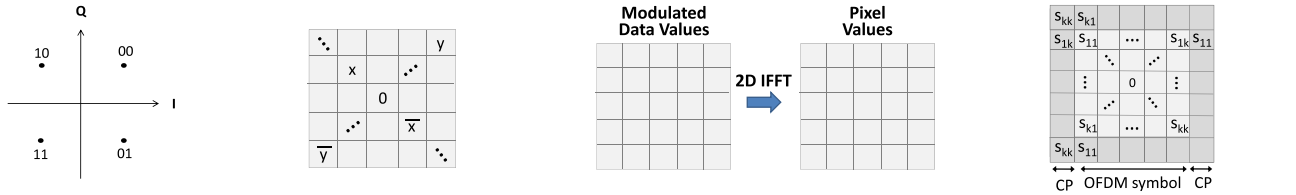
Figure 4: Inter-pixel and Inter-symbol interference encountered in the LCD-camera channels. Each pixel at the transmitter is shown to interfere with its neighboring pixel at the receiver. After the introduction of the cyclic prefix, the Inter-symbol interference disappears.

The above problem is analogous to multipath in RF wireless channels. Specifically, in environments with multipath, multiple copies of the signal traverse different paths and arrive at the receiver with different delays. As a result, each received signal sample is a linear combination of a few consecutive transmitted samples. RF transmitters deal with multipath by employing OFDM and appending a *cyclic prefix* to their signal [8]. We adopt a similar approach in PixNet. Adding a cyclic prefix (CP) means repeating the first few signal samples at the end of the symbol [8]. In PixNet, we append the CP around the entire symbol, as shown in Figure 3(d). Specifically, we copy the first few rows from the top of the symbol to the bottom of the symbol and vice versa. We also do the same for a few columns at the left and right edges of the symbol. Since the cyclic prefix introduces a separation between the two symbols as shown in Fig 4, it prevents pixels belonging to different symbols from interfering with one another. Inter-pixel interference is dealt with by the fact that OFDM transforms the samples into the frequency domain. The seemingly complex operation of pixel bleeding which is a convolution in the spatial domain turns into a simple multiplication operation in the frequency domain. This makes it easy to address blur from inter-pixel interference as described in Section 6.1.

The operator can pick the size of the OFDM symbol and the cyclic prefix length. The cyclic prefix length chosen should be greater than the amount by which pixels bleed

[1] A Hermitian function is a complex function with the property that its complex conjugate is equal to the original function with the variable changed in sign, i.e., $f(-x) = \overline{f(x)}$.

[2] The pixels on the camera (RX) and the LCD (TX) are shown to be spatially aligned only for demonstration purposes. Typically, the number of pixels on the camera and the LCD are different and the pixels are not aligned.



(a) 4QAM maps a bit pair to one of four complex numbers

(b) A Hermitian OFDM symbol

(c) IFFT maps modulated bits to pixel values

(d) An OFDM symbol with a Cyclic Prefix (CP)

Figure 3: Components of PixNet's Transmitter.

into each other. Empirically, we found that an OFDM symbol of 81×81 pixels and a cyclic prefix of 5 pixels (introduced on all sides of the symbol) works reasonably well.

5. PixNet'S RECEIVER

PixNet's receiver works by first extracting the LCD's coded frame from the frame captured by the camera, using a corner detection algorithm. It then corrects for perspective distortion present in this extracted frame and obtains the right samples to be passed through an FFT for normal OFDM decoding.

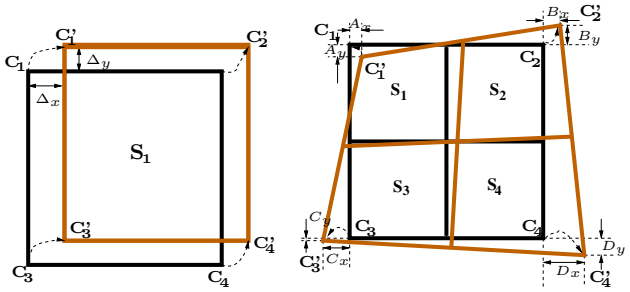
5.1 Corner Detection

Before the receiver can start decoding, it needs to separate the coded frame from the background in the image. This requires the receiver to detect the four corners of the coded frame. Corner detection is a widely studied problem [33]. In particular, the literature on 2D barcodes has a variety of corner detection algorithms. PixNet can work with any of them. Our implementation uses the Data Matrix corner detection algorithm described in [3].

5.2 Perspective Correction

When a camera takes a picture of an object, it projects that object on to the plane of its sensors. This projection includes shifting the object with respect to its surroundings, scaling the object, and in general, distorting the geometry of the original object. This problem of perspective transformation is widely studied in computer vision [5]. Our context however differs from the traditional context studied in computer vision. On the one hand, our constraints are stricter since we cannot use offline algorithms or tolerate minor distortions which are typically acceptable to the human eye [5]. On the other hand, we have more flexibility in our design since we can encode the imaged object in a manner that simplifies correcting for perspective distortion. Because of these differences, we do not use traditional perspective correction algorithms. Instead we develop our own algorithm which generalizes the OFDM sampling correction algorithm.

The intuition underlying our approach is simple: We approach perspective distortion as a sampling problem. Specifically, the LCD pixel values refer to the signal samples at the transmitter. The camera pixel values refer to the signal samples at the receiver. When the LCD and camera are at an angle, some parts of the LCD are closer to the camera, and hence occupy a relatively bigger space in the image. This emulates the effect of sampling these parts at a higher rate. Parts of the LCD that are further away from the camera, occupy a relatively smaller space in the image, and hence it is as if they were sampled at a slower rate. To correct for perspective distortion, PixNet needs to find the relationship



(a) A 2D OFDM symbol with shift offsets

(b) A 2D OFDM symbol with perspective distortion

Figure 5: Example distortions of the LCD-Camera channel.

between the sampling points on the camera and those on the LCD, and re-sample the received image at the locations that best reflect the original LCD samples.

Sampling differences between sender and receiver occur in RF channels too. The DAC at the transmitter and the ADC at the receiver are not synchronized to sample the signal at the same instants. Furthermore, they typically differ in their sampling frequency. As a result, the receiver's samples are different from the transmitter's samples. To decode the signal properly, the receiver needs to resample the signal as closely as possible to the transmitter's samples. However, in RF channels, the sampling is regular and hence there is no analogy to the case of geometric distortion.

In this section, we build on this intuition to provide a generalized OFDM sampling correction algorithm that works with perspective distortion. Before dealing with the general case of irregular sampling, we motivate our approach with a simpler special case, where the samples at the receiver are all shifted by the same amount with respect to the samples at the transmitter.

5.2.1 A Constant Shift

Consider a 2D OFDM symbol sampled with shifts Δ_x on the x-axis, and Δ_y on the y-axis, as shown in Fig. 5(a). A basic property of the Fourier transform is that shifts in the signal domain translate into phase offsets in the frequency domain [8].[3] Given this property, it is relatively simple to figure out how a constant sampling shift effects the encoded data. Specifically, since a 2D OFDM symbol is generated by taking a 2D IFFT of the complex numbers, $s_{k,l}$, that represent the modulated bits, a sampling shift introduces phase shifts in these complex numbers as follows:

[3]This argument is subject to the shift being less than the cyclic prefix of the OFDM symbol since any larger shifts cause interference from nearby OFDM symbols.

PROPOSITION 5.1. *Let $\Delta\theta_x$ be the difference between the phase shifts experienced by $s_{k,l}$ and $s_{k',l}$, and $\Delta\theta_y$ the difference between the phase shifts experienced by $s_{k,l}$ and $s_{k,l'}$. Then:*

$$(\Delta\theta_x, \Delta\theta_y) = \left(\frac{2\pi(k-k')\Delta_x}{L_s}, \frac{2\pi(l-l')\Delta_y}{L_s}\right) \quad (1)$$

where $L_s \times L_s$ is the size of the 2D-FFT.

PROOF. The proof is relatively simple. Specifically, the 2D OFDM symbol can be expressed as:

$$S_{m,n} = \sum_{k=0}^{L_s-1}\sum_{l=0}^{L_s-1} s_{k,l}\, e^{j\frac{2\pi km}{L_s}} e^{j\frac{2\pi ln}{L_s}}$$

If The receiver samples the OFDM symbol with a shift (Δ_x, Δ_y), the resulting samples are given by

$$S'_{m,n} = \sum_{k=0}^{L_s-1}\sum_{l=0}^{L_s-1} s_{k,l}\, e^{j\frac{2\pi k(m+\Delta_x)}{L_s}} e^{j\frac{2\pi l(n+\Delta_y)}{L_s}}$$

To decode the OFDM symbol from these samples, the receiver takes a 2D FFT. Given however that the samples are shifted, the FFT does not reproduce exactly the original complex numbers; it produces a phase shifted version $z_{p,q}$ of the original complex numbers as follows:

$$\begin{aligned}
z_{p,q} &= \frac{1}{L_s^2}\sum_{\substack{m=0;\\ n=0}}^{L_s-1} S'_{m,n}\, e^{-j\frac{2\pi mp}{L_s}} e^{-j\frac{2\pi nq}{L_s}} \\
&= \frac{1}{L_s^2}\sum_{\substack{m=0;\\ n=0}}^{L_s-1}\sum_{\substack{k=0;\\ l=0}}^{L_s-1} s_{k,l}\, e^{j\frac{2\pi m(k-p)}{L_s}} e^{j\frac{2\pi n(l-q)}{L_s}} e^{j\frac{2\pi k\Delta_x}{L_s}} e^{j\frac{2\pi l\Delta_y}{L_s}} \\
&= \frac{1}{L_s^2}\sum_{\substack{k=0;\\ l=0}}^{L_s-1} s_{k,l}\, e^{j\frac{2\pi k\Delta_x}{L_s}} e^{j\frac{2\pi l\Delta_y}{L_s}} \sum_{m=0}^{L_s-1} e^{j\frac{2\pi m(k-p)}{L_s}} \sum_{n=0}^{L_s-1} e^{j\frac{2\pi n(l-q)}{L_s}} \\
&= \sum_{k=0}^{L_s-1}\sum_{l=0}^{L_s-1} s_{k,l}\, e^{j\frac{2\pi k\Delta_x}{L_s}} e^{j\frac{2\pi l\Delta_y}{L_s}} \delta(k-p)\delta(l-q) \\
&= s_{p,q}\, e^{j\frac{2\pi p\Delta_x}{L_s}} e^{j\frac{2\pi q\Delta_y}{L_s}}
\end{aligned}$$

Thus, each complex number, $s_{k,l}$ experiences a phase shift $\frac{2\pi k\Delta_x}{L_s} + \frac{2\pi l\Delta_y}{L_s}$. Thus, the difference in phase shift between $s_{k,l}$ and $s_{k',l}$ is $\frac{2\pi(k-k')\Delta_x}{L_s}$ and that between $s_{k,l}$ and $s_{k,l'}$ is $\frac{2\pi(l-l')\Delta_y}{L_s}$. □

The above proposition allows the receiver to estimate the sampling shifts (Δ_x, Δ_y) from differences in the phase shifts of the modulated complex numbers, $(\Delta\theta_x, \Delta\theta_y)$. But how does the receiver obtain $(\Delta\theta_x, \Delta\theta_y)$? To enable the receiver to compute $(\Delta\theta_x, \Delta\theta_y)$, we transmit a few known complex numbers in each OFDM symbol, which we refer to as *pilots*. Since the receiver knows the original phase of each transmitted pilot, it can easily compute the phase shift experienced by each pilot. The receiver then computes differences in phase shifts between two pilots in the same row, $\Delta\theta_x$, and two pilots in the same column $\Delta\theta_y$, and substitutes them in eq. 1 to obtain the sampling shifts. Once it has the sampling shifts, it resamples the OFDM symbol at the correct locations, and computes the FFT again to obtain the correct modulated complex numbers. Finally the receiver demodulates the complex numbers to obtain the bits using standard 4QAM demodulation.

5.2.2 Perspective Correction Algorithm

We now consider a general case where we sample a 2D OFDM symbol with a perspective distortion as shown in Fig. 5(b). Following from the observation above, we suspect that these corner shifts should similarly result in phase offsets at the receiver.

However, unlike the previous case where we needed to estimate 2 unknowns (Δ_x, Δ_y), in this case we need to estimate 8 unknowns $(A_x, A_y), (B_x, B_y), (C_x, C_y)$ and (D_x, D_y). If we consider any particular symbol we would have only 2 equations, one for phase shifts θ_x, and another for θ_y. But we need at least 8 equations to estimate these 8 offsets. We solve this problem by considering a super-symbol consisting of 4 symbols as shown in Fig. 5(b). Now, we can generate 2 equations for each symbol, resulting in 8 equations overall. In the Appendix, we prove the following:

PROPOSITION 5.2. *Consider a 2×2 2D OFDM super-symbol generated by taking the IFFT of the complex numbers $s_{k,l,r}$, where k and l are integers between 0 and $L_s - 1$ and $r \in \{1, 2, 3, 4\}$ denotes the symbol index. Let us assume that this symbol is sampled with relatively small x and y corner offsets of $(A_x, A_y), (B_x, B_y), (C_x, C_y)$ and (D_x, D_y) at its four corners, as show in Fig. 5(b). Let $\Delta\theta_{x,r}$ be the difference between the phase shifts experienced by $s_{k,l,r}$ and $s_{k',l,r}$, and $\Delta\theta_{y,r}$ be the difference between the phase shifts experienced by $s_{k,l,r}$ and $s_{k,l',r}$. Then:*

$$\begin{pmatrix}\Delta\theta_{1,x}\\ \Delta\theta_{2,x}\\ \Delta\theta_{3,x}\\ \Delta\theta_{4,x}\end{pmatrix} = \frac{2\pi(k-k')}{16L_s}\begin{pmatrix}9&3&3&1\\3&9&1&3\\3&1&9&3\\1&3&3&9\end{pmatrix}\begin{pmatrix}A_x\\B_x\\C_x\\D_x\end{pmatrix}$$

$$\begin{pmatrix}\Delta\theta_{1,y}\\ \Delta\theta_{2,y}\\ \Delta\theta_{3,y}\\ \Delta\theta_{4,y}\end{pmatrix} = \frac{2\pi(l-l')}{16L_s}\begin{pmatrix}9&3&3&1\\3&9&1&3\\3&1&9&3\\1&3&3&9\end{pmatrix}\begin{pmatrix}A_y\\B_y\\C_y\\D_y\end{pmatrix} \quad (2)$$

PROOF. Refer to the Appendix.

Note that eq. 2 is a generalized version of eq. 1 for $(A_x, A_y) = (B_x, B_y) = (C_x, C_y) = (D_x, D_y) = (\Delta_x, \Delta_y)$. After estimating the offsets $(A_x, B_x, C_x, D_x), (A_y, B_y, C_y, D_y)$ with the help of eq. 2 using the pilot bins as described in sec. 5.2.1, the receiver re-samples each OFDM symbol at the correct sampling points and computes the FFT again. The resulting complex numbers are demodulated to obtain the transmitted bits.

6. PixNet CROSS TX-RX TECHNIQUES

PixNet takes advantage of two techniques which are implemented partially in the transmitter and partially in the receiver. These are blur-adaptive coding and frame synchronization.

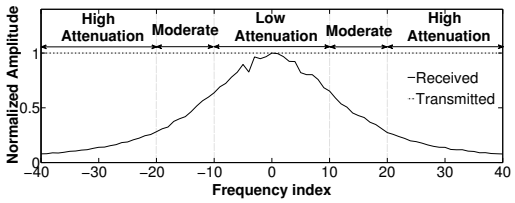


Figure 6: Received signal amplitude across different frequencies with a DELL 30" flat display and a Nikon D3X camera at a distance of 2 meters.

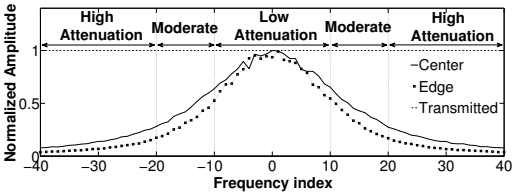


Figure 7: Received signal amplitude across different frequencies for symbols that are in the focal plane and out of the focal plane.

6.1 Blur-Adaptive Coding

Blur eliminates sharp transitions in an image and causes nearby pixels to blend together. Thus, its impact is similar to that of a low pass filter, i.e., it attenuates the high frequencies in an image. Fig. 6 shows the signal amplitude as a function of the frequency index for a transmitted signal and its received version. The transmitted signal is chosen to have the same energy in all frequencies. We can see from the figure that the high frequencies (i.e., those above 20 or below -20) are heavily attenuated. Thus, these frequencies cannot be used for transmitting information. Conversely, the lowest frequencies (between -10 and 10) are mostly preserved and can deliver information with almost no error. Finally, the frequencies in between, experience significant attenuation but can still be used to transmit some information.

Since PixNet operates in the frequency domain, it can naturally deal with different frequencies experiencing different attenuation. PixNet completely suppresses very high frequencies, and does not use them to transmit information. Frequencies that experience low or moderate attenuation are used to transmit information, but protected with a Reed Solomon error correcting code. An error correction code is chosen for each frequency with redundancy commensurate with the attenuation that frequency experiences.

We find that the amount of attenuation varies not only with the frequency, but it also varies with the position of the symbol in a given frame. Specifically, cameras focus on a particular plane. Objects that are either nearer or farther from the focus plane experience blur (this is typically known as limited depth-of-field (DoF)) [36]. Thus, when the LCD and camera have a view angle with respect to each other, only the center of the LCD will be in focus. Symbols away from the center of the frame are not in the plane of focus and hence experience increased attenuation due to blur (in addition to perspective distortion) as shown in Fig. 7. PixNet exploits this information while optimizing the redundancy in the error correcting code. Specifically, the symbols at the center of the frame will have lower redundancy and the symbols away from the center of the frame will have higher redundancy.

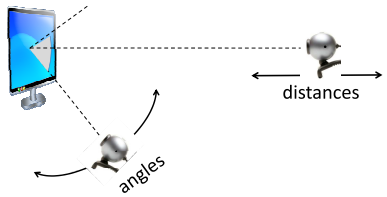


Figure 8: Experimental Setup

The configuration of the Reed Solomon code in PixNet is as follows: The code operates on a block size of 255, where each element in the block is 8 bits [6]. Based on empirical calculations, PixNet uses codes that correct up to 3% (weak), 7% (medium), and 15% (strong) errors. Frequencies labeled *low attenuation* in Fig. 6 are always coded with the weak code. The *moderate attenuation* frequencies, however, are coded with the medium code for symbols in the focal plane, and with the strong code for symbols out of the focal plane.

6.2 Frame Synchronization

Unsynchronized operation between LCD and camera can cause the camera to capture an image while the LCD is in the process of rendering a new frame. We have performed detailed experiments to identify specific distortions arising from the lack of transmitter-receiver synchronization. We observed the following two distortions:

- *Frame Splitting*: This distortion causes two consecutive LCD frames to form the top and bottom parts of one frame captured by the camera. This occurs because the graphics card writes the next frame to the LCD's frame buffer while the LCD has not completely rendered the current frame. This is the so-called 'vertical sync' problem that occurs when copying of frames by the graphics card is out of sync with the rendering of frames by the LCD screen.
- *Frame Shadowing*: In this case, the image captured by the camera is a linear combination of two consecutive LCD frames. Specifically, camera captures light for the duration that the shutter is open, and integrates over this time period. Thus, if the LCD moves from rendering one frame to the next while the camera's shutter is open, the camera integrates light from both LCD frames. The net effect of this integration process is a shadowed frame.

PixNet tackles these problems as follows. For frame splitting, we leverage the OpenGL graphics library [32]. OpenGL provides the `GLFlush` and `GLFinish` interface commands as part of its API to force the graphics card to wait for the LCD screen to complete rendering of the previously copied frame before the next frame is copied to its frame buffer. This effectively eliminates the frame splitting problem. To tackle frame shadowing, we calibrate the shutter speed to ensure that the camera's frame rate is at least twice that of the LCD. This configuration ensures that each frame displayed by the LCD is captured without shadowing in at least one of the camera's frames. Note that it is common in communication systems to run the receiver at twice the speed of the transmitter to improve robustness to synchronization errors [7].

7. EVALUATION ENVIRONMENT

Hardware: For our transmitter, we use a DELL 30" flat display at a resolution of 2560 × 1600 pixels. At the receiver end, we experiment with a Casio EX-F1, which has a resolution of 6 megapixels, and a Nikon D3X, which has a resolution of 24 megapixels. In some experiments, we use a Nokia N82 phone, which has a 5 megapixel camera.

Compared Schemes: We compare two encoding schemes:

- **PixNet:** We have implemented a prototype of PixNet in C. All our experiments use a symbol size of 81 × 81 pixels with a cyclic prefix of 5 pixels.
- **QR codes:** We experiment with a baseline scheme that encodes data using QR code. Our experiments use the open source reference implementation of QR code from [42].

Experimental Setup: Fig. 8 illustrates our experimental setup. We fix the location of the LCD and move the camera. We experiment with different distances between the LCD and camera as well as different view angles. We pack as many QR codes or PixNet OFDM symbols that can fit on the screen. We set the camera exposure time to 1/60 seconds, hence allowing for a frame rate of 60 fps. Since at most half of the frames show shadowing (see §6.2), the effective frame rate is 30 fps. We use the autofocus functionality of the cameras, except for the blur experiments in which we control the focus to show the impact of different levels of blur.

Metric: We compare PixNet and QR code schemes in terms of their throughput, that is the number of correctly received bits after Reed-Solomon decoding.[4] The throughput is computed as the average per frame throughput multiplied by 30 frames/s. In all experiments we use a shutter speed (exposure time) of 1/60 second, which corresponds to a frame rate of 60 fps and an effective frame rate of 30 fps. The Casio EX-F1 has a burst mode that captures 60 fps. The Nikon D3X FX does not have a burst mode and hence cannot deliver 60 fps in realtime. However by setting the exposure time to 1/60 seconds, we obtain the same quality as if each frame lasted 1/60 second. Our intent however in presenting results for Nikon D3X FX is to show the effect of different sensor types on the performance of PixNet.

Ensuring Fair Comparison: For a fair comparison between PixNet and QR code, we need to ensure that they both have the same upper bound on the amount of information they can pack in a frame. To do so, we note that cameras and LCDs have three color channels: Red, Green, and Blue. QR codes are monochrome and hence set all three channels to the same value. To ensure that PixNet cannot pack more information by using colors, we have PixNet use a single channel, which is the green channel. Also, QR codes are black and white and do not use grey levels. Thus, they allow a maximum of one bit per pixel. PixNet also allows a maximum of one bit per pixel. Specifically, the combination of 4QAM modulation and a hermitian structure results in, on average, one bit per spatial frequency. The number of spatial frequencies is the same as the number of pixels (they are the input and output of IFFT). Thus, both PixNet and QR code can transmit at the same maximum data rates.

[4]For QR code, we use the default "L" error correction level, which allows for maximum throughput per code.

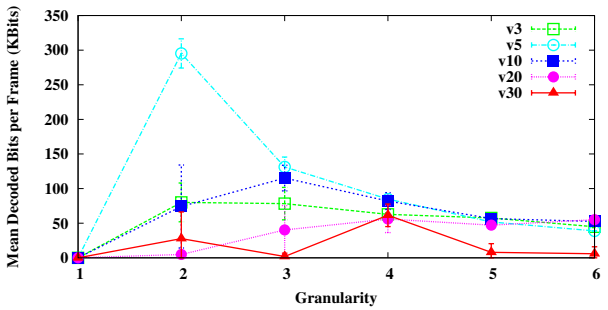


Figure 9: Impact of QR code's granularity and version on its performance.

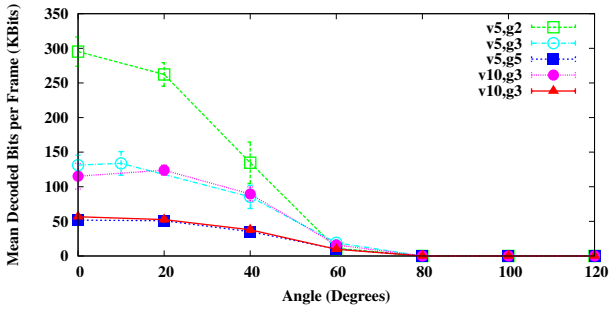


Figure 10: Impact of angle on throughput achieved with different QR Code versions and granularities

The actual throughput they achieve in practice depends on how many bits/s they can deliver reliably.

7.1 Calibration of QR Code

As mentioned earlier, we allow QR code to transmit one bit/pixel, i.e., to use a granularity of 1. In practice, however, such a granularity suffers excessive errors and reduces the throughput. Thus, we take an additional step and calibrate the QR code parameters for optimal performance. The QR code [2] has two parameters: granularity and version. The granularity defines the number of pixels used to convey a single bit. For instance, a granularity of 2 means that each bit is expressed using 2 × 2 pixels. The version defines the number of bits encoded in a single QR code. For instance, v5 encodes 864 bits per QR code.

Method. For each version and granularity, we fit as many QR codes as we can on the 30" Dell 3007WFP LCD screen we use in all our experiments. We perform two types of experiments. First, we study the interaction between the version and granularity. To do so, we align the LCD-camera pair (i.e., ensure zero view angle) and fix the distance between them to two meters. We take multiple pictures of the screen using the Nikon D3X camera and decode each frame using the QR code reference implementation [2]. Second, we repeat the above experiment but with different view angles and distances.

Results. Fig. 9 shows the effect of different QR code granularities and versions (For clarity, we plot the subset of the versions that showed better performance). In principle, increasing the version or decreasing the granularity increases the data rate but decreases robustness to errors. The figure shows that QR code version 5 with a granularity of 2 represents a sweet spot in this tradeoff, and yields the best

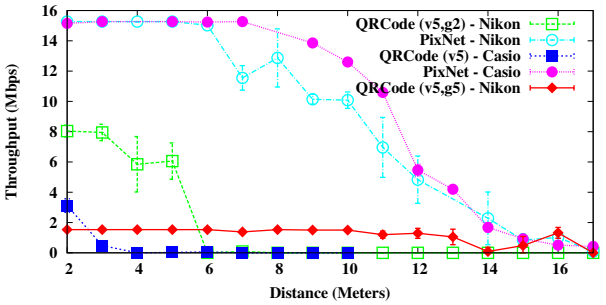


Figure 11: Impact of distance on throughput of PixNet and QR Code

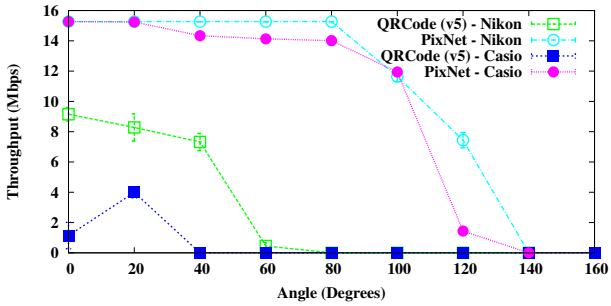


Figure 12: Impact of viewing angle on PixNet and QR Code

performance. Higher versions pack too many bits in a single QR code, and hence fail to support the fine granularities, resulting in lower overall throughput. Lower versions pack too little information in a QR code and hence are not effective at amortizing per QR code metadata (e.g., corner and synchronization markers).

Fig. 10 shows the performance of different QR code versions and granularities as a function of the viewing angle (again, we plot the subset of configurations that showed better performance). It confirms that version 5 at a granularity of 2 performs the best independent of the viewing angle.

We also study the impact of distance and find that there is no clear winner. Specifically, the optimal QR code granularity increases with distance. For the distances that we experiment with (i.e., [2m-16m]) v5 with granularity 2 works well for the shorter distances and version 5 with granularity of 5 works well for the longer distances. Hence, for our distance experiments, we show results for both in §8. Finally, note that though we allow QR code to use both configurations to deal with distances, all of the experiments in this paper use a single configuration for PixNet (a symbol size of 81×81 pixels with a 5 pixels cyclic prefix).

8. RESULTS

(a) Impact of Distance

First, we examine whether an LCD-camera link can deliver data over reasonable indoor distances.

Method. In this experiment, we keep the transmitter and receiver aligned but move the receiver away in increments of a meter, each time allowing the camera to refocus before capturing the frames. This experiment is performed with Nikon and Casio cameras as receivers. For each location, the camera zooms to fill the camera's view with the transmitted frame.

Results. Fig. 11 shows the throughput of PixNet and QR code as a function of distance between the transmitter and the receiver. We show results for QR code v5 at granularities of 2 and 5. We observe that in comparison with QR code, PixNet delivers up to $2x - 9x$ higher throughput depending on the distance. We also observe that PixNet continues to achieve throughputs of up to 12 Mbps at a distance of 10m which is representative of a large open space, such as a conference room or auditorium. In contrast, QR code (v5,g2) fails to work beyond 5m. QR code (v5,g5) however is able to support distances comparable to PixNet, but with an order of magnitude worse performance at short distances.

Interestingly, the figure also shows that while PixNet achieves a higher throughput with Casio EX-F1, QR code achieves a higher throughput with Nikon D3X. This is due to the two cameras having different pixel sizes. Nikon D3X has a CCD pixel size of 34.46 μm^2 where as Casio EX-F1 has a CCD pixel size of 193.54 μm^2, which is around 5.6 times bigger than that of Nikon. Having a bigger pixel size allows the pixels to gather more light quickly, and reduces the amount of shot-noise (noise that arises because of small number of photons being incident on the sensor). However, bigger pixel size also means that the light being gathered at each receiver pixel is potentially a combination of light from multiple transmitter pixels, and hence images from Casio have higher blur (inter-pixel interference).

As described in section 6.1, PixNet has an effective way of dealing with blur. This, combined with the fact that Casio can collect more light makes PixNet perform better with Casio. QR codes on the other hand are highly sensitive to blur and hence perform poorly with Casio when compared with Nikon.

(b) Impact of Viewing Angle

We evaluate PixNet and QR code with perspective distortion.

Method. In this experiment, the distance between transmitter and receiver is maintained at 2m and the view angle is changed to span the range $[0°, 140°]$ (view angles are measured as solid angles, where $140°$ refers to $-70°$ to $70°$).

Results. Fig. 12 shows the throughput as a function of the view angle, both for PixNet and QR Code. PixNet is able to provide up to half its peak throughput (8 Mbps) at angles as wide as $120°$, corresponding to a $3x$ gain in view angles compared with QR Code.

We note that while PixNet achieved longer distances with Casio, it achieves wider angles with Nikon. The reason for this behavior is that at very high angles, the transmitted rectangular image becomes a trapezoid at the receiver. The image size shrinks heavily on one side. As a result, only a few camera pixels are available to sample that portion of the image. Since Nikon has smaller pixels and a higher resolution than Casio, it can sample much more finely and retain more information. The performance of Casio is once again inferior for QR codes because of higher blur and coarser sampling.

(c) Impact of Blur

Blur is a common distortion in LCD-camera channels. It occurs because of focus imperfection, jerks while capturing images, or low resolution at the camera [36].

Method. To measure the throughput achieved as a function of the level of blur in the image, we use the same setup as in the previous experiments. We ensure that the frame

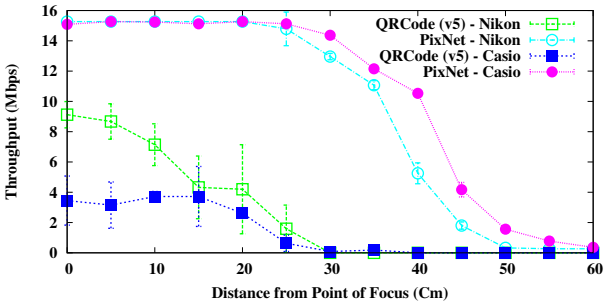


Figure 13: Impact of Blur on throughput achieved, using PixNet and QR Code

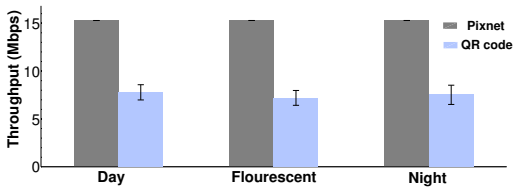


Figure 14: Impact of ambient light on Throughput achieved, using PixNet and QR Code

displayed on the LCD is in sharp focus when the camera is at a distance of 2m from the LCD. We then progressively move the camera back away from the plane of focus. We take measurements at increasing distances without re-focusing the camera before capturing frames.

Results. Fig. 13 shows how PixNet's throughput changes with the amount of blur. PixNet provides constant throughput for distances up to 40 cm away from the plane of focus. On the other hand, QR code is highly sensitive to blur and rapidly deteriorates in performance even with a little amount of blur. More importantly, this result shows that PixNet does not degrade in performance even if the image experiences a reasonable amount of blur either due to handshakes or because it is not perfectly in focus.

We once again find that Casio performs better than Nikon for PixNet and performs worse than Nikon for QR code. The reason for this behavior is the difference in pixel sizes for Nikon and Casio as discussed before.

(d) Impact of Ambient Light

For an LCD-camera link to provide high throughput in a variety of settings, its performance must remain stable over varying degrees of illumination.

Method. We fix the camera at a distance of 2m from the LCD. We keep the LCD and camera planes at zero view angle. We experiment with this setup for both PixNet and QR Code, under natural light, fluorescent light and in a dark room.

Result. In Figure 14, we present the throughput achieved by both PixNet and QR code. We observe that PixNet's performance is independent of the ambient light in the environment.

We observe that QR code's performance is also stable across different lighting conditions. QR code achieves this by employing a histogram based light-balancing technique that calculates an optimal intensity threshold in order to distinguish black pixels from white pixels. QR code's technique involves binning different intensity values and finding successive peaks and a valley in an intensity histogram. PixNet's scheme for light balancing however, is much simpler and requires no additional computation. It comes as an added advantage of operating in the frequency domain. Since illumination changes only effect the DC frequency, PixNet does not transmit any information in the DC frequency and simply discards the DC frequency at the receiver to prevent the variation in illumination from effecting the transmitted data.

Camera Model	PixNet	QR code v5, g2	QR code v5, g5
Nokia N82	4.301 Mb/s	NILL	1.544 Mb/s
Nikon D3X	15.28 Mb/s	7.6673 Mb/s	1.55 Mb/s
Casio EX-F1	15.15 Mb/s	3.47 Mb/s	1.539 Mb/s

Table 1: Nominal throughput in Mbps of PixNet and QR codes (version 5, granularities 2 and 5) across different receivers.

(e) Smart Phones

Smart phone cameras have smaller apertures than standard cameras [36]. This implies that the amount of light reaching a smart phone camera's sensors is typically an order of magnitude low (the intensity of light varies as a square of the aperture's diameter). And hence, the picture quality is usually poorer and has a lot more noise. Here, we would like to see what performance can be achieved by smart phone cameras.

Method. We use the setup in Fig. 8 with no appreciable angle between the camera and the LCD. We use Nokia N82 as our receiver. The LCD and the smart phone are separated by a distance of 2m.

Results. In Table 1, we show the throughput achieved for Nokia N82 for PixNet and QR code. We present results for Nikon and Casio receivers as well for comparison. Note that due to increased noise in the images captured by N82, QR codes yield no throughput at a granularity of 2. However, N82 is able to realize the maximum possible throughput that can be achieved with QR codes at a granularity of 5. We see that PixNet performs well even for smart phones. Also since smart phone technology is rapidly evolving, we can expect to see smart phones equipped with higher quality cameras that will yield even better performance.

9. DISCUSSION

So far, we have shown that PixNet links can deliver multiple Mb/s over significant distances and wide view angles. While the full potential of these links is yet to be discovered, in this section, we discuss possible applications and general technology trends.

9.1 PixNet's Applications

PixNet is useful in scenarios with high user density. Traditional RF-based communication performs poorly in such scenarios because of increased contention between a large number of co-located users. In contrast, the high directionality of LCD-camera links enables a multitude of such links to communicate simultaneously without interfering with each other, and hence allows the throughput to scale with the number of communicating entities.

- **PartyNets:** In a congested party, business reception, or trade show, tens, hundred, and sometimes thousands of pairwise communications occur simultaneously, in which people exchange business brochures, photos, video demos, etc. A technology like PixNet has two advantages for such scenarios. First, it allows all concur-

rent pairwise communications to occur without worrying about interference. Second, it provides a privacy similar to that of standard human communication, that is only nearby cameras looking directly at the screen of a smart phone can capture the information.

- **GalleryNets:** In a museum, gallery, or a mall, there is an interest in providing additional information about the displayed objects. For example, in a museum or gallery the user may access an audio or video presentation about the artifact she/he is viewing. In a mall, a user may see pictures or videos featuring the cloth she is interested in buying. These applications challenge existing RF communications, since at any point in time there may be many users who are interested in learning about different objects. LCD camera links, with their contention-free nature, can address the challenge. Specifically, each artifact in a museum or window in a mall, can have a small LCD screen that transmits video and audio about it. A user who is interested in learning about the object can simply point her phone's camera toward the screen and obtain the desired information.
- **VANETs:** In the area of Intelligent Transportation Systems (ITS) there have been proposals to modulate traffic lights with driver assistance information such as current traffic conditions, distance to destination, GPS coordinates, etc. Low cost cameras mounted on cars are used to capture images and decode the information [17]. There are also proposals to do the same using automobile head lights that display information comprised of driving speed, collision warnings, driver intentions etc., [31]. Neighboring cars use cameras to capture this information. This approach is particularly useful on congested highways where the traffic density and the resulting contention challenge traditional RF communications.
- **Augmented Reality:** The applications above require the use of active transmitting elements such as an LCD or a projector. PixNet also works with passive sources where encoded frames are painted as 2D barcodes. Digital Lane Marking is a development in this direction where information comprising road curvature, highway numbers, GPS coordinates etc. are painted on road lanes. Cameras installed on automobiles can decode this information as they drive along the road and thereby provide a safe and enhanced driving experience. Another application lies in stock control where information about each item is encoded in the form of 2D barcodes printed on the product. Today such barcodes need to be scanned at a close distance with no view angles. With PixNet these barcodes can be scanned by a camera at large distances and view angles.

9.2 Current trends

Most experiments in this paper use standalone cameras and LCDs. These devices have better capabilities than cameras and LCDs embedded in cell phones and PDAs. However, the cost and specifications of digital cameras and LCDs are improving at a fast pace. Specifically, digital cameras are getting cheaper every day. The number of "pixels per dollar" has been growing exponentially over the past decade and is expected to continue along this trend [20]. Further, new technologies such as microlens arrays provide sophisticated optical capabilities in significantly smaller form factors. Hence, even advanced cameras would no longer require large optical separation between the sensor and the lens, making them more compact [40, 39]. Also with the advent of electronic shutters, cameras are now able to support up to hundreds or thousands of frames per seconds [28]. Even consumer cameras support a range of frames per second from 10 to 1000 fps [9], and these values continue to increase.

Not only are image sensors becoming increasingly cheap but the LCD market has also seen a sharp decrease over the past decade, primarily due to the demand from the TV market. Furthermore, with the development of pico projectors, large area displays are not only going to become cheaper but also embedded on handheld devices [13, 24].

10. CONCLUSIONS

We present PixNet, a novel approach to transmit data from an LCD to a camera, and show that such links can deliver high throughput over multi-meter distances and wide view angles. LCD-camera pairs provide a new point in the design space of wireless communication. On the one hand, visible light signals cannot traverse walls and physical obstacles, thus limiting their communication range. On the other hand, LCD-camera links are robust against interference, making them attractive in scenarios with dense pairwise communications. We believe that spatial OFDM and the suite of tools introduced in this paper offer a new perspective on the capability of these communication channels. LCD-camera links also enable an exciting array of applications ranging from using large LCD displays for broadcasting digital content to supporting point-to-point data transfer between handheld devices such as smart phones. We believe our work merely scratches the surface and a rich body of research lies ahead in this new and exciting field.

11. ACKNOWLEDGMENTS

We thank Prof. Ramesh Raskar for his insights on optical systems. We also thank the reviewers and our shepherd, Alex Snoeren, for their comments on the paper. This work is funded by an NSF grant.

12. REFERENCES

[1] Automatic identification and data capture techniques – Aztec code bar code symbology specification. ISO/IEC 24778:2008.
[2] Automatic identification and data capture techniques – QR code 2005 bar code symbology specification. ISO/IEC 18004:2006.
[3] Automatic identification and data capture techniques - Data Matrix bar code symbology specification. ISO/IEC 16022:2006.
[4] International symbology specification – maxicode. ISO/IEC 16023:2000.
[5] *Digital Image Warping*. IEEE computer society Press, 1990.
[6] *Reed-Solomon Codes and Their Applications*. Wiley-IEEE Pres, 1999.
[7] *Digital Communication Receivers: Synchronization, Channel Estimation, and Signal Processing*. John Wiley and Sons, 2001.

[8] *Multi Carrier Digital Communications: Theory and Applications of OFDM*. Springer New York, 2004.

[9] C. D. Cameras. http://www.casio.com/news/content/56fef8fb-6b1f-470c-9372-29fbd4b4c146/.

[10] H. Chen, R. Sukhthankar, G. Wallace, and T. jen Cham. Calibrating scalable multi-projector displays using camera homography trees. In *In Computer Vision and Pattern Recognition*, pages 9–14, 2001.

[11] H. Chen, R. Sukthankar, G. Wallace, and K. Li. Scalable alignment of large-format multi-projector displays using camera homography trees. In *Proceedings of IEEE Visualization*, pages 339–346, 2002.

[12] S. Gupta and C. Jaynes. Active pursuit tracking in a projector-camera system with application to augmented reality. In *CVPR*, 2005.

[13] A. Hang, E. Rukzio, and A. Greaves. Projector phone: A study of using mobile phones with integrated projector for interaction with maps. In *Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI)*, September 2008.

[14] P. R. Haugen, S. Rychnovsky, A. Husain, and L. Hutcheson. Optical interconnects for high speed computing. *OPT. ENG.*, 25(10):1076–1085, October 1986.

[15] HCCB. http://research.microsoft.com/en-us/projects/hccb/.

[16] S. Hranilovic and F. R. Kschischang. A pixelated mimo wireless optical communication system. In *IEEE Journal of Selected Topics in Quantum Electronics*, 2006.

[17] N. K. Kanhere and S. T. Birchfield. A taxonomy and analysis of camera calibration methods for traffic monitoring applications. In *IEEE Transactions on Intelligent Transportation Systems*, volume 11, pages 441–452, June 2010.

[18] H. Kato et al. 2d barcodes for mobile phones. In *2nd Int'l Conf. Mobile Technology, Applications and Systems (MTAS 2005)*.

[19] T. Langlotz and O. Bimber. Unsynchronized 4d barcodes: coding and decoding time-multiplexed 2d colorcodes. In *ISVC'07: Proceedings of the 3rd international conference on Advances in visual computing*, pages 363–374, Berlin, Heidelberg, 2007. Springer-Verlag.

[20] H. Law. http://en.wikipedia.org/wiki/digital_camera.

[21] C. G. Lee, C. S. Park, J. H. Kim, and D. H. Kim. Experimental verification of optical wireless communication link using high-brightness illumination light-emitting diodes. *Fiber Optics and Optical Comm.*, 47, December 2007.

[22] H. Lee and J. Kim. Retrospective correction of nonuniform illumination on bi-level images. *Opt. Express*, 17, 2009.

[23] T. D. Link. http://en.wikipedia.org/wiki/timex_datalink.

[24] D. Madden. CES 2008: DLP shows off working pico projector prototype. 7.1.2008. http://www.pocket-lint.com/news/12082/pico-projector-prototype-shown-off.

[25] D. A. Miller. Optical interconnects to silicon. *IEEE Journal on selected topics in quantum electronics*, 6(6), 2000.

[26] H. L. Minh, D. C. O'Brien, G. E. Faulkner, L. Zeng, K. Lee, D. Jung, and Y. Oh. High-speed visible light communications using multiple-resonant equalization. In *IEEE Photonics Technology Letters*, 2008.

[27] A. Mohan, G. Woo, S. Hiura, Q. Smithwick, and R. Raskar. Bokode: imperceptible visual tags for camera based interaction from a distance. In *ACM SIGGRAPH*, 2009.

[28] MotionBLITZ. http://www.compumodules.com/image-processing/high-speed-camera/.

[29] D. O'Brien, H. L. Minh, L. Zeng, G. Faulkner, K. Lee, D. Jung, Y. Oh, and E. T. Won. Indoor visible light communications: challenges and prospects. In *SPIE Optics and Photonics*, volume 7091, pages 709106–709106–9, 2008.

[30] E. Ohbuchi, H. Hanaizumi, and L. A. Hock. Barcode readers using the camera device in mobile phones. In *International Conference on Cyberworlds (CW'04)*, 2004.

[31] R. O'Malley, E. Jones, and M. Glavin. Rear-lamp vehicle detection and tracking in low-exposure color video for night conditions. In *IEEE Transactions on Intelligent Transportation Systems*, volume 11, pages 453–462, June 2010.

[32] OpenGL. http://www.opengl.org.

[33] J. R. Parker. *Algorithms for Image Processing and Computer Vision*. Wiley, 1996.

[34] C. Pinhanez. Using a steerable projector and camera to transform surfaces into interactive displays, 2001.

[35] M. A. Q. Packaging history: The emergence of the uniform product code (upc) in the united states. *History and Technology*, 11(1):101–111, 1994.

[36] R. Raskar and J. Tumblin. *Computational Photography: Mastering New Techniques for Lenses, Lighting, and Sensors*. A K Peters, 2010.

[37] J. Rekimoto and Y. Ayatsuka. Cybercode: Designing augmented reality environments with visual tags. In *DARE*. ACM Press, 2000.

[38] M. Rohs. Real-world interaction with camera phones. *Lecture Notes in Computer Science*, 2005.

[39] R. Shogenji, Y. Kitamura, K. Yamada, S. Miyatake, and J. Tanida. Multispectral imaging using compact compound optics. In *Optics Express*, volume 12, pages 1643–1655, 2004.

[40] J. Tanida, T. Kumagai, K. Yamada, S. Miyatake, K. Ishida, T. Morimoto, N. Kondou, D. Miyazaki, and Y. Ichioka. Thin observation module by bound optics (TOMBO): concept and experimental verification. In *Appl. Opt.*, volume 40, pages 1806–1813, 2001.

[41] R. Yang, D. Gotz, J. Hensley, H. Towles, and M. S. Brown. Pixelflex: A reconfigurable multi-projector display system, 2001.

[42] Zxing. http://code.google.com/p/zxing/.