Challenge: Mobile Optical Networks Through Visual MIMO

Ashwin Ashok, Marco Gruteser, Narayan Mandayam
WINLAB, Rutgers University, 671 Route 1 South, North Brunswick, NJ, USA
aashok@rutgers.edu, {gruteser, narayan}@winlab.rutgers.edu

Jayant Silva, Michael Varga, Kristin Dana
Department of ECE, Rutgers University, 94 Brett Road, Piscataway, NJ, USA
{jayantms, mfvarga}@rutgers.edu, kdana@ece.rutgers.edu

ABSTRACT

Mobile optical communications has so far largely been limited to short ranges of about ten meters, since the highly directional nature of optical transmissions would require costly mechanical steering mechanisms. Advances in CCD and CMOS imaging technology along with the advent of visible and infrared (IR) light sources such as (light emitting diode) LED arrays presents an exciting and challenging concept which we call as visual-MIMO (multiple-input multiple-output) where optical transmissions by multiple transmitter elements are received by an array of photodiode elements (e.g. pixels in a camera). Visual-MIMO opens a new vista of research challenges in PHY, MAC and Network layer research and this paper brings together the networking, communications and computer vision fields to discuss the feasibility of this as well as the underlying opportunities and challenges. Example applications range from household/factory robotic to tactical to vehicular networks as well pervasive computing, where RF communications can be interference-limited and prone to eavesdropping and security lapses while the less observable nature of highly directional optical transmissions can be beneficial. The impact of the characteristics of such technologies on the medium access and network layers has so far received little consideration. Example characteristics are a strong reliance on computer vision algorithms for tracking, a form of interference cancellation that allows successfully receiving packets from multiple transmitters simultaneously, and the absence of fast fading but a high susceptibility to outages due to line-of-sight interruptions. These characteristics lead to significant challenges and opportunities for mobile networking research.

MobiCom'10, September 20–24, 2010, Chicago, Illinois, USA.

Categories and Subject Descriptors

C.2.1 [**Computer-Communication Networks**]: Network Architecture and Design—*Wireless Communication*

General Terms

Design, Measurement, Experimentation, Performance

Keywords

Visual MIMO, Optical Communication, Computer Vision, RF Communication, Line of Sight(LOS)

1. INTRODUCTION

Radio frequency based wireless communications and networking has seen tremendous growth over the last several years serving as the foundation for myriad applications. With such increased adoption, the non-line of sight and ubiquitous propagation characteristics of wireless communications at typical radio frequencies, while often an advantage, are also leading to many unmitigated challenges. For example, they lead to increased co-channel interference, eavesdropping, and spoofing risks that make it hard to meet stringent reliability or security requirements. For many applications ranging from household and factory robotics to vehicular networks to pervasive computing, wireless communication in the optical spectrum can address such challenges through directional transmissions with narrow beamwidths and line-of-sight restrictions. These directional transmissions reduce co-channel interference through improved spatial reuse and make it difficult for an eavesdropper to detect the presence of communications. In contrast, achieving similar beamwidths in the RF spectrum is impractical as it would require inordinately large antennas, due to the larger wavelength.

Advances in CMOS imaging technology along with the advent of visible and infrared (IR) light sources such as light emitting diode (LED) arrays or LCD screens present an exciting and challenging concept to enable *mobile* optical networking. In this concept, which we call as Visual-MIMO (visual multiple-input multiple-output), optical transmissions by multiple transmitter elements are received by an array of photodiode elements (e.g. pixels in a CMOS camera). This paper brings together the networking, communications and computer vision fields to discuss the feasibility of this concept as well as the underlying opportunities and challenges in PHY, MAC, and Network layer research.

Mobile Optical Limitations. Optical wireless communications with narrow beams, however, has hitherto been impractical in most mobile settings, because both the sender and receiver need to operate with very narrow beams and angles-of-view, respectively, to achieve transmission ranges greater than a few tens of meters. Except for short-range diffuse IR transmitters with a range of about 10m [1, 18] wireless optical transmissions are thus largely confined to stationary building-to-building transmission links. Free-space optics transceivers designed for this purposes can achieve ranges of a few kilometers under good weather conditions [15]. Due to the extremely narrow beamwidths used, any application with some mobility would require costly mechanical steering systems for transmitter and receiver [11].

Optical wireless requires very narrow beams to achieve longer ranges because the signal-to-noise ratio is limited by several factors. First, transmission power levels are lower than in the RF spectrum because of output power limitations of LED technology and eye safety restrictions for laser transmitters. The optical spectrum is characterized by high background noise typically by sunlight in the infrared and visible light wavelengths and other IR heat sources in vicinity. The radiation of visible light emitted by the sun is many orders of magnitude higher than the power it emits in radio frequencies. Third, in addition to this background noise, optical receivers experience shot noise. Shot noise is caused by the random arrival of photons at the receiver. Since the energy of one photon in the optical spectrum is much higher than in the radio spectrum, fewer photons will be generated at the same transmission power level and at these lower quantities variations from random arrival patterns have a more significant effect on signals. Thus, in optical systems shot noise dominates thermal noise, which is a main receiver noise source in RF systems.

The Visual MIMO Approach. This paper argues that it is now becoming feasible to overcome the transmission range limitations of conventional wireless optics through camera receivers and LED transmitter arrays and that developing protocols and mobile computing systems with this technology presents many exciting new research challenges for the mobile networking and computing community. The image sensor in a camera is essentially an array of photodiodes and the camera lens provides a different narrow field of view for each photodiode. This creates a large number of highly directional receive elements (the camera pixels), which allows reducing interference and noise and thereby can achieve large ranges, yet still maintain the wide field-of-view necessary for mobile communications. The tradeoffs in the visual MIMO system, however, are a limited receiver sampling frequency (e.g., hundreds to thousand frames per second for lower end cameras and a million frames per second for high-end models) and, as in all optical wireless communications, strong line-of-sight (LOS) requirements. To address the rather limited frame rate (sampling frequency) of current cameras, the system can use a visual MIMO approach, i.e. transmit with multiple LEDs and record the signal with multiple camera pixels. As we will discuss, this approach can also allow many "parallel" communication channels, similar in concept to RF MIMO systems [13], albeit over a channel with very different characteristics. At the physical layer, the visual MIMO approach requires techniques to acquire and track signals from a transmitter as they are captured by different photodiodes (pixel) during movement. We show how a physical layer could rely on computer vision/image analysis as opposed to traditional baseband signal processing, opening many avenues for interdisciplinary research. At the PHY layer, visual MIMO can also benefit from exploiting the multiplexing/diversity tradeoff as a function of the resolvability of multiple images on the image plane at various distances between the transmitter and receiver. This differs from the channel fading dependent multiplexing/diversity gain tradeoffs in RF-MIMO systems where distance is not a key concern. At the MAC layer, visual MIMO can also benefit from novel channel access mechanisms that adapt between parallel transmissions when "interference cancellation" is possible and separate channel access when it is not. The reliance on line-of-sight communications and the fact that mobility (e.g. such as in vehicular or robotic networks) may present intermittent links, as well as the perspective-dependent achievable throughput also calls for new visions for MAC and networking layer protocols that can keep track of network geometry.

Applications. Several key applications in the mobile computing field could benefit from visual MIMO. First, safety applications in vehicular networks such as emergency electronic brake lights (EEBL) [37] and cooperative collision warning (CCW) [10] require reliable communications under potentially high co-channel interference because vehicle position and dynamics information needs to be shared among nearby vehicles in potentially very dense highway scenarios. Visual MIMO could reduce interference because it's directional and line-of-sight transmissions allow for increased spatial reuse.

Communications in military applications can be enhanced by the increased security of visual MIMO channels. The line-of-sight requirement greatly reduces the potential for interception and jamming that is inherent in RF communication. Additionally, the source of the signal interception can be more easily determined, so the potential for spoofing signals is reduced. Longer range communication could be accomplished by a network of visual MIMO channels consisting of cameras/monitor relay stations.

The ubiquitous placement of LCD screens and surveillance cameras in urban environments create numerous opportunities for practical applications of visual MIMO channels. LCD screens for electronic signage can have dual functionality by transmitting embedded signals via intensity modulation, so that visual observation for human observers would coexist with a visual MIMO wireless communications channel. Alternatives to intensity modulation include the use of angle-based modulation where observation of the screen at different angles enables different visual observation. Angle-based modulation can be accomplished via polarization methods or digital mirror arrays. Such embedded signals may also enable new user interface, for example by facilitating recognition of pointing or gestures with a camera-equipped mobile device.

Visual MIMO also may find application in computer vision, where camera networks refer to the cooperation of numerous cameras viewing a scene in order to create a 3D environment map. The key challenges in these networks is (1) accurate camera calibration so that each camera has a known position/orientation and (2) accurate point correspondences in order to compute geometry via stereo or structure-from-motion algorithms. Camera networks can utilize visual MIMO protocols to transmit/receive a temporal pattern to uniquely identify key scene points to provide unambiguous point correspondences and enable robust camera calibration even in low light conditions. An interesting merger of computer vision recognition algorithms with communications protocols can be explored by recognizing not static passive objects, but objects that are communicating known temporal pilot sequences and headers.

To focus our discussion, the remainder of this paper will discuss the visual MIMO concept in the context of vehicular network communications.

2. RELATED

While there is a large body of work in optical networking [25] and free space optics [24, 5], it largely focuses on stationary rather than mobile networks. Except for recent spherical FSO transceiver designs for mobile ad hoc networks [36] and optical satellite communications with physical steering [11, 28], mobile optical communications research has primarily focused on short range infrared communications for mobile devices [18, 33]. While earlier work has used cameras to assist in steering of FSO transceivers [35],the visual MIMO approach differs by directly using cameras as receiver to design an adaptive visual MIMO system that uses multiplexing at short distances but still can achieve ranges of hundreds of meters in a diversity mode. It exploits advances in CMOS imagers that allow higher frame rates compared to earlier CCD designs.

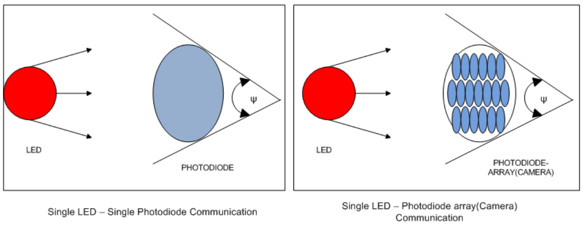
IR has a small range (typically up-to 10m), the effective power of the IR beam has to be restricted to not damage human tissues, and IR transmitters are relatively costly to build. Thus, more recently, research has also explored using the visible light spectrum for communication [20, 30, 3, 27, 19]. Low-speed audio communication systems using LED transmitters have already been demonstrated [29]. In Japan, a consortium of 21 research groups called the Visible Light Communication Consortium (VLCC) has been formed to research into areas of VLC [3]. Since 2008, the Smart Lighting research group at Boston University [2] has been investigating visible light communication systems for indoor lighting and outdoor vehicle to vehicle applications [9]. All this work generally uses photodiodes at the receiver to convert the optical signals to electrical signals. Though photo diodes can convert pulses at very high rates, they suffer from large interference and background light noise. This results in very low signal-to-noise ratios (SNR), which leads to the short range of typical IR communication systems, even with more sophisticated receiver processing and modulation techniques as studied in [32].

Only a few sporadic projects have recently begun to investigate cameras as receivers, particularly for inter-vehicle communications [34] and traffic light to vehicle communications [8]. Their analytical results show that communication distances of about 100 m with a BER $\leq 10^{-6}$ are possible. Other work has investigated channel modeling [20] and multiplexing [4]. More recently, researchers of the MIT Bokode project [23] have applied computational photography to camera based communications. Building on such results and directions, this paper argues that the novel concept of visual MIMO is becoming feasible and that it presents exciting opportunities and challenges to the mobile computing and networking community.

3. LED-CAMERA COMMUNICATIONS

Photodiode arrays of a camera can provide a wide receiver field of view that allows for node mobility without the need to realign the receiver. Yet, by virtue of the camera design, each single photodiode element has a very narrow field of view, allowing high gain communication. The camera lens creates the effect of each photodiode being angled to a slightly different part of the scene, so that the combination of all diodes generates an image with a wide field of view. Other research groups have recently proposed variations of such designs [33]. For example, if larger receiver sizes are practical, the lens can be eliminated by using a photodiode array on a spherical receiver structure [26].

3.1 Capacity Analysis



Single LED – Single Photodiode Communication

Single LED – Photodiode array(Camera) Communication

Figure 1: LED-Photodiode/Camera Communication Illustration

We analyze a stationary communication model where a single LED with output power P_t transmits to an optical receiver over a wireless channel as shown in figure 1. This is a conservative model, because it does not include the effect of scene noise due to motion and achievable gains from multiple parallel transmission (from multiple LEDs). The two types of optical receivers we consider in our analysis are, (a) a conventional photodiode receiver and (b) a photodiode array (camera) receiver.

In an optical wireless channel, since the frequency of the optical signal is very large compared to the rate of change of the impulse response, multipath fading and doppler shift are negligible. As described by Kahn and Barry [18], the received signal power follows $P_r = (RhP_t)^2$ where h is a channel parameter called channel DC gain and R is the receiver's responsivity or the optical power to current conversion ratio. However the received signal is corrupted by noise from the optical channel which is typically dominated by shot noise from background light sources and modeled as an additive white Gaussian process (AWGN) with a two sided power spectral density per unit area $S(f) = qRP_n$ [34, 18]. Here, q is the electron charge and P_n quantifies the power in background light per unit area. Hence, for a receiver sampling rate of W, the noise power is $P_N = qRP_nAW$ where A is the area of the photodiode. The signal to noise ratio for a single LED-single photodiode communication is,

$$SNR_{pd} = \frac{P_r}{P_N} = \frac{\kappa P_t^2 d^{-4}}{qRP_nAW} \tag{1}$$

where κ is a function of parameters such as the LED's lambertian radiation pattern, irradiance angle, field-of-view and optical concentration gain of the receiver [18].

Applying the model to the photodiode array receiver, we observe that the key difference between a conventional photodiode receiver and an array receiver lies in the detector area. When using the array, we assume the receiver can select the subset of diodes that actually observe a strong signal from the transmitter. This effectively reduces the detector area size and consequently reduces the noise. For the camera receiver (with a fixed-focus setting of the camera lens), we estimate the area of the array actually used through perspective projection [17]. Given a focal length f, a round LED of diameter l and the distance d between camera and LED, the LED will occupy a circle of diameter $l' = \frac{fl}{d}$ on the photodetector array. To conservatively account for the quantization effects, we assume that it will occupy a square area of size l'^2. This noise reduction gain is, however, limited by camera resolution. When the LED moves away from the camera, the projected diameter l' will eventually become smaller than the size of a photodiode. From this point on, the camera cannot further reduce the number of photodiodes that are used in the reception process and its performance becomes similar to a single conventional photodetector (having the size of one pixel). We refer to distance where the LED generates an image that falls onto exactly one pixel as the critical distance $d_c = fl/s$, where s is the edge-length of a pixel.

Following this analysis, the signal to noise ratio for a single LED-photodiode array(camera) communication is,

$$SNR_{cam} = \begin{cases} \frac{\kappa P_t^2 d^{-2}}{qRP_n W f^2 l^2} & \text{if } d < d_c \\ \frac{\kappa P_t^2 d^{-4}}{qRP_n W s^2} & \text{if } d \geq d_c \end{cases} \tag{2}$$

We observe from equations (1) and (2), for $d < d_c$, that a camera receiver has gain in SNR over a single photodiode receiver in the order of d^2. Thus at larger distances a camera would be a more resourceful option than a single photodiode receiver. Also for $d > d_c$, though the camera receiver is equivalent to a single photodiode the gain in performance can be achieved by reducing the pixel size s which is not possible in a photodetector. Since current off-the-shelf camera implementations are more limited in sampling rate (which equates to frame rate in camera) than photodetectors, a camera system will likely achieve even higher SNRs than a photodetector with a high-sampling rate. The lower framerate, however, also directly limits achievable rates.

To understand this tradeoff, given that the noise model is AWGN, we plot the Shannon capacity $C = Wlog_2(1+SNR)$ over a range of distances in figure 2 for a single photodiode receiver and three different camera receivers. We set the sampling rate at 100MHz for the photodiode and 1000fps for the Basler Pilot piA640 machine vision camera & 100fps for SONY PS3eye webcam (two off-the-shelf cameras which use a CCD image sensor). We also consider a hypothetical camera which could sample at a rate of 1M fps. The parameter values underlying this result are summarized in Table 1. The graph shows that even at the low sampling rates of a toy webcam the camera system can still outperform the single photodiode due to its SNR advantage at larger distances. Moreover, the capacity of the camera system can be increased considerably by using an array of LED transmitters (appropriately spaced) where the capacity at short distances can be scaled by a number equal to the number of LEDs and in some cases at longer distances too. We also see that the capacity of a camera system is more consistent over distance than for a single photodiode system for which it falls off rapidly (relatively) over distance.

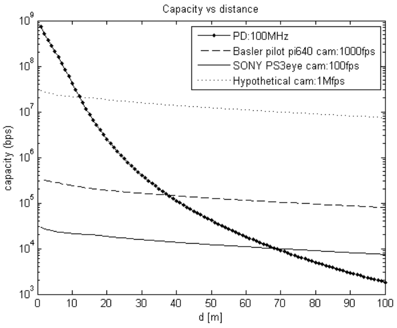


Figure 2: Capacity versus distance for the proposed system with Photodiode and Camera receivers

To further illustrate the camera advantage of eliminating noise by selecting only the photodiodes that receive the signal, we conduct an experiment with a blinking LED positioned 2m from the camera. The camera records a sequence of images in this completely stationary scenario. Figure 3 shows two histograms of the mean pixel value, one computed over a 10×10 area centered on the LED and one computed over the complete 640×480 image. These represent a single photodiode approach and a camera with the ability to eliminate background noise as discussed. The figure shows that in the first case the on and off state can be clearly distinguished through pixel values while in the second case the distinction is difficult since the signal is masked by shot noise.

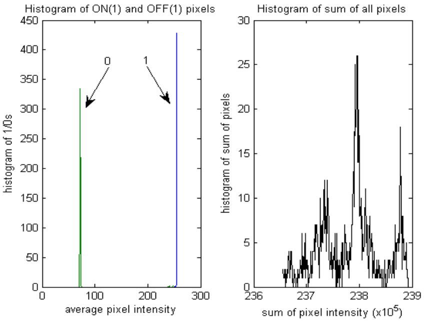


Figure 3: Histogram plots of Basler Pilot piA640 camera snapshots in medium sunlight(left:10×10, right:640×480)

Note that in a mobile transmitter-receiver scenario the camera's SNR gain (and hence the capacity gain) over a single photodiode can be expected to be pronounced because of scene noise, for example in a situation where the 'scene' has a strong reflector such as a white body. By extracting only those areas of the image that observe a strong transmitter signal, a camera can also selectively eliminate these distractors (noise) which is not possible with a single photodiode.

Parameter	**PD**	**B**	**S**
P_t[mW]	100	100	100
$FOV\psi$[deg]	50	50	50
$A[mm^2]$	15.7	15.7	15.7
$P_n[mW/cm^2]$	600	600	600
l[mm]	6	6	6
f[mm]	–	21	6.5
$s[\mu]$	–	7.1	6

Table 1: Table of parameter values for photodiode and camera(PD Photodiode,B Basler Pilot piA640, S SONY PS3eye)

4. TOWARDS A VISUAL MIMO PHY- A COMPUTER VISION APPROACH

To realize the potential capacity gains described in the previous section, the visual MIMO system needs to identify which set of photodiodes receive the signal, or equivalently, which region of the image contains LED transmitters. The output of the photodiode array in this case is equivalent to an image, where each pixel is analogous to a single photodiode. This task of identifying which region in the image contains LED transmitters is analogous to antenna selection in RF MIMO systems. Conventional techniques such as known pilot sequences are not suitable for the visual MIMO system because of the framerate limitations of cameras. High framerates are usually achieved by reading data only from one or more small regions of interests (a limited set of photodiodes). When the set of photodiodes that receive the signal is not yet known the complete array of photodiodes must be read out, which is only possible at lower frame rates. Due to node mobility and a lower framerate, the set of photodiodes receiving the signal can change before the pilot sequence is completed, rendering the pilot sequence approach ineffective.

We propose to draw from techniques in the computer vision community to develop receiver-side processing techniques that can identifying and tracking the pixels that contain the image. Visual imagery is rich in detail and objects in images can be represented computationally via *feature vectors.* Given a computational representation for LED transmitters, feature-based recognition can be used for localization, or signal acquisition, even with the complexity of dynamic traffic scenes.

Challenges of Real World Scenes. The challenges presented include: (1) camera motion , (2) illumination variation and (3) background distractors such as other vehicles on the road. Camera motion is inherently present in the visual MIMO communications system because the camera at the receiver and LED transmitters are on different mobile nodes. Consequently, the geometry of the image formation process varies, i.e. the position and orientation of the camera center with respect to the scene varies. As the camera moves further, the object of interest appears to become smaller. Because of this perspective projection, the LED transmitter undergoes arbitrary scaling, and the standard communications approach of template matching with matched filters or with correlation-based detectors is insufficient. The computer vision literature has numerous methods for achieving scale invariance in object recognition. In our prototype system we employ the popular approach of SIFT-matching [21], scale invariant feature transform, for representing and recognizing the LED transmitter.

While camera motion creates geometric issues in matching, illumination variation causes photometric issues to overcome for LED transmitter localization. The appearance of the LED transmitter changes with illumination variation in the scene. Therefore, simple intuitive methods (such as detecting the red region) do not work in practice. The problem of color constancy is well documented in the computer vision literature [16, 12, 6]. While human perception creates a constant color representation of objects, the color measurement varies dramatically and is not a reliable method for detection. In this system, the feature-based representation for the LED transmitter is robust to photometric variations due to illumination changes (e.g. sun vs. shade) as well as geometric variations due to camera motion.

The use of CV algorithms also helps to locate the LED transmitter in the presence of background distractors. Recall that the primary advantage of the visual MIMO channel over standard communication channels is the ability to focus attention at the correct portion of the scene. The photodiode approach is not a viable option for communications with LED's due to the significant noise increase with distance. For the Visual MIMO system, the background portions of the image can be discarded and therefore do not contribute to channel noise. The spatial focus achieved by the CV algorithms is obtained using two methods: (1) *recognition* and (2) *tracking*. The two methods can be interpreted as two modes of operation for the module that locates the LED transmitter. For the recognition mode, there is no assumption of the LED transmitter's location and the entire image is searched in order to find the current location. Once recognized, the LED transmitter region can be tracked in subsequent frames. The tracking mode has lower computational cost than recognition mode because a smaller image region is processed. However, both modes have computational algorithms that typically run in real time.

4.1 First Experiments

As a preliminary prototype of the transmitter, we have implemented an array of LEDs that can be connected via a USB interface to a PC. The LED array is controlled by an array of Field-effect Transistors (FET) with signals generated by a microcontroller. The microcontroller receives the LED constellations via the USB connection and generates the corresponding LED signals based on its internal timer. A Basler Pilot pi640 camera was mounted on the dashboard of a car and used to capture video at 640×480 resolution and 60 fps of the car ahead while driving at 25km/h. The image of the LEDs was then rendered onto the license plate of the car in the video using motion estimation and image warping. The video was then used to test the recognition and tracking of the LED transmitter using computer vision techniques.

Figure 4 illustrates the recognition and tracking of LED transmitters for signal acquisition. The recognition task is implemented using the scale invariant feature transform (SIFT) method [21] by comparing the current image with a template image in a manner that is robust to scale. Standard SIFT matching used in the experiment runs in a time of 1.34 seconds for a 640×480 image. While the standard implementation was used for this prototype, the computational speed of SIFT is not expected to be a bottleneck for a version for two reasons. First, the SIFT algorithm can be modified for speed reduction. For example, an approximate SIFT algorithm has been developed by [14], which runs in 0.180 seconds for a 400×320 image. Another example of SIFT variations is the SURF method [7] which achieve a 200ms computation speed for a typical image. Additionally, recent developments such as [31] show SIFT implemented on hardware such as a Field Programmable Gate Array (FPGA) to improve its speed by an order of magnitude. Tracking in real-time is accomplished and implemented in many vision tasks. Here the tracking is based on the Lucas and Kanade [22] implementation in OpenCV where the tracker runs in 30 frames per second.

Computer vision techniques not necessarily need to process every frame; which can simplify computation and further enhance the processing speeds of the system. But, apart from the computation complexity of the CV algorithms the system implementation has a few important constraints such as cost and power requirements (refer Table 2) especially when considering a mobile transceiver design. While cost can be traded off with complexity of the system, power management in mobile devices is still a big challenge.

Parameter	**L**	**B**	**S**
Cost(per unit)	\$25	\$1700	\$40
Power consumption	0.1W	5.5W	2.5W

Table 2: Implementation constraints in our preliminary prototype (L LED array (3x3) transmitter, B Basler Pilot piA640, S SONY PS3eye)

5. NETWORKING CHALLENGES

Characteristics of a visual MIMO network such as highly directional transmitters and receivers, strict line-of-sight requirements, and a perspective dependent multiplexing gain also raise many challenges for research at the link and network layers.

Perspective-Dependent Multiplexing Gain In contrast to RF wireless channels, the visual MIMO system will not be subject to multipath fading. It's bitrate can therefore be expected to change at slower timescales. Achievable bitrate does depend, however, on perspective and distance between transmitter and receiver. The receiver can distinguish all LEDs when it has a full frontal view on the transmitter array at close distance. At a large distance or at from an angled view, the LEDs will blend together in the image. Thus, in the first case, information can be multiplexed over all LEDs and the system can achieve high datarates while in the second case the system should operate in a diversity mode at lower datarates. Exploiting this property will require new methods of diversity-multiplexing and bitrate adaptation.

Use of Geometric Information: Since achievable bitrates are primarily dependent on receiver perspective and LOS availability, visual MIMO protocols could benefit from knowledge of the network geometry (as opposed to maintaining only topology and/or SNR information). Such knowledge is useful both at the physical and network layer. At the physical layer, for example, the transmitter could provide the receiver with information about the transmitter LED array geometry (i.e., an LED template) to assist the receiver in recognition, tracking, and demodulation. Geometry is also useful at the network layer because, unlike for RF wireless channels, link bitrates are quite predictable given network geometry. Since interference, multipath and doppler effects are negligible, link capacity is largely defined by the distance between and orientation of two nodes (given known transmitter and receiver configuration), unless the line-of-sight path is obstructed. In addition, location and orientation vary relatively slowly and predictably compared to RF wireless channels even in automotive highway settings (at least for cars moving in the same direction). Thus, it is sufficient for the receiver to send distance and angular information ev-

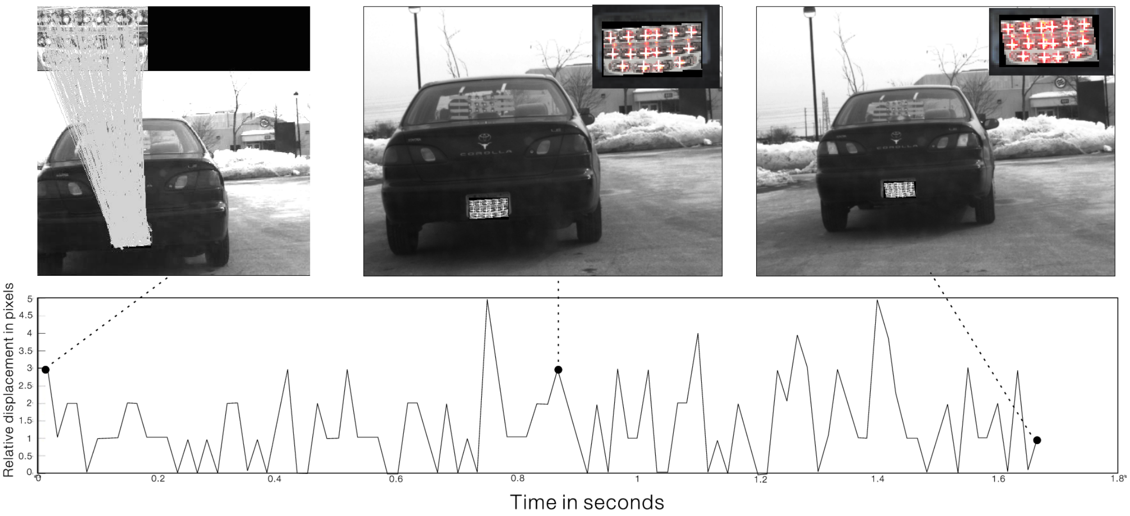


Figure 4: Left: SIFT feature-based recognition showing corresponding features between the LED template and the query image. Center and Right: LK tracking of an LED Transmitter. The inset shows a magnified view of the tracked LEDs as indicated by the white crosses. Bottom: Relative displacement of an LED due to the motion of the car over 120 frames.

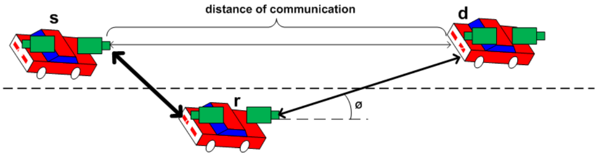


Figure 5: Multi-path transmission strategies using geometric information

ery few hundred milliseconds, it is not needed on a per packet basis.

Visual ranging and network localization: Visual MIMO can also give rise to different localization techniques, which could be used to track the network geometry information described before. Given a known LED template, distance and angle information can be generated through camera pose estimation, an image analysis techniques. It is worth studying whether accuracy can be improved through particular signaling techniques or additional information from the transmitter, particularly under partial occlusion or FOV-clipping of the LED array. The pairwise pose estimates can also be refined through network localization algorithms. Unlike RF-based network localization, these algorithms should take into account that accuracy of individual pose estimates declines with increased angles.

Visual multi-path transmissions. Visual MIMO can also perform interference cancellation to simultaneously receive packets from multiple nodes. This calls for novel MAC protocols that can allow higher spatial reuse and adapt between separately scheduled and parallel transmissions when possible. This characteristic of the network, however, also allows multi-path transmission strategies that are similar in concept to cooperative communications but carry less coordination overhead. Consider a scenario with three nodes as shown in figure 5: a source, a destination,and one potential relay, which is positioned in-between the two other nodes but closer to the source than the destination (without obstructing line of sight between source and destination). Because shorter distances allow higher multiplexing gain, it is likely that the link capacity between source and relay is greater than the others, $C_{sr} > C_{rd} > C_{sd}$. Thus, the multi-hop path sr, rd has higher capacity than the direct link sd, but the highest throughput can be achieved through simultaneous transmission through the relay and on the direct link. Since transmission through the relay is limited by C_{rd} and $C_{sr} > C_{rd}$, the source can use its excess capacity to transmit information directly to d, which can receive both the relay transmission and the source transmissions in parallel. Transmitters can again use geometric information to decide on transmission strategies.

With relatively low rates, forwarding through multiple hops leads to potentially large delays. Fortunately, the practically non-interfering nature of the transmissions and the predictable channel can be exploited with cut-through techniques to achieve lower forwarding delays. It also allows setup of longer-lived virtual circuits, with a set of LEDs at each transmitter allocated for each circuit.

6. CONCLUSIONS

We argued that advances in CMOS and LED technology have enabled the concept of a visual MIMO system. Visual MIMO allows communication range of hundreds of meters with a relatively wide field-of-view compared to free-space optics, thereby enabling a higher a degree of node mobility. Our analysis showed that even visual MIMO system using a toy webcam can achieve close to order of magnitude gains in bitrate over a conventional photodetector receiver with the same field-of-view and there is significant room for improvement through more specialized image sensors. This visual MIMO system presents a broad spectrum of opportunities and challenges for mobile computing and networking research. At the physical layer, it can take advantage of computer vision-inspired techniques to recognize, track, and extract the transmitting LEDs from an image. Its characteristics of directional transmission, interference cancellation, line-of-sight disruptions, a perspective-dependent multiplexing gain, and a flat (no fading) channel also call for reexamining MAC and network layer protocols.

7. REFERENCES

[1] Irda. `http://http://www.irda.org/index.cfm/`.

[2] Smart lighting erc at boston university (slc/bu). `http://smartlighting.bu.edu/research/index.html`.

[3] Visible light communication consortium. `http://vlcc.net`.

[4] S. Arai, S. Mase, T. Yamazato, T. Endo, T. Fujii, M. Tanimoto, K. Kidono, Y. Kimura, and Y. Ninomiya. Experimental on hierarchical transmission scheme for visible light communication using led traffic light and high-speed camera. In *Vehicular Technology Conference, 2007. VTC-2007 Fall. 2007 IEEE 66th*, pages 2174–2178, 30 2007-Oct. 3 2007.

[5] J. c. Arun Majumdar. *Free-Space Laser Communications:Principles and Advances*. Springer, Nov 2007.

[6] K. Barnard, G. Finlayson, and B. Funt. Color constancy for scenes with varying illumination. *Computer Vision and Image Understanding*, 65(2):311–321, February 1997.

[7] H. Bay, T. Tuytelaars, V. Gool, and L. Surf: Speeded up robust features. In *9th European Conference on Computer Vision*, Graz Austria, May 2006.

[8] H. Binti Che Wook, T. Komine, S. Haruyama, and M. Nakagawa. Visible light communication with led-based traffic lights using 2-dimensional image sensor. In *Consumer Communications and Networking Conference, 2006. CCNC 2006. 3rd IEEE*, volume 1, pages 243–247, Jan. 2006.

[9] J. Carruthers, S. Carroll, and P. Kannan. Propagation modelling for indoor optical wireless communications using fast multi-receiver channel estimation. *Optoelectronics, IEE Proceedings -*, 150(5):473–481, Oct. 2003.

[10] T. ElBatt, S. K. Goel, G. Holland, H. Krishnan, and J. Parikh. Cooperative collision warning using dedicated short range wireless communications. In *Proceedings of the 3rd international workshop on Vehicular ad hoc networks*, pages 1–9, Los Angeles, CA, USA, 2006. ACM.

[11] E. Fischer, P. Adolph, T. Weigel, C. Haupt, and G. Baister. Advanced optical solutions for inter-satellite communications. *Optik - International Journal for Light and Electron Optics*, 112(9):442 – 448, 2001.

[12] B. Funt, M. Drew, and J. Ho. *International Journal of Computer Vision*, 6(1):5–24, april 1991.

[13] A. Goldsmith. *Wireless Communications.* Cambridge, 2005.

[14] M. Grabner, H. Grabner, and H. Bischof. Fast approximated sift. In *In 7th Asian Conference of Computer Vision*, pages 918–927, 2006.

[15] D. Heatley, D. Wisely, I. Neild, and P. Cochrane. Optical wireless: the story so far. *Communications Magazine, IEEE*, 36(12):72 –74, 79–82, dec 1998.

[16] J. Ho, B. V. Funt, and M. S. Drew. Separating a color signal into illumination and surface reflectance components: Theory and applications. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(10):966–977, October 1990.

[17] B. K. P. Horn. *Robot vision.* MIT Press, Cambridge, MA, USA, 1986.

[18] J. Kahn and J. Barry. Wireless infrared communications. *Proceedings of the IEEE*, 85(2):265–298, Feb 1997.

[19] S. Kitano, S. Haruyama, and M. Nakagawa. Led road illumination communications system. In *Vehicular Technology Conference, 2003. VTC 2003-Fall. 2003 IEEE 58th*, volume 5, pages 3346–3350, Oct. 2003.

[20] T. Komine and M. Nakagawa. Fundamental analysis for visible-light communication system using led lights. *IEEE Transactions on Consumer Electronics,*, 50(1):100–107, Feb 2004.

[21] D. G. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60:91–110, 2004.

[22] B. D. Lucas and T. Kanade. An iterative image registration technique with an application to stereo vision. *Proceedings of Imaging understanding workshop*, pages 121–130, 1981.

[23] A. Mohan, G. Woo, S. Hiura, Q. Smithwick, and R. Raskar. Bokode: imperceptible visual tags for camera based interaction from a distance. In *SIGGRAPH '09: ACM SIGGRAPH 2009 papers*, pages 1–8, New York, NY, USA, 2009. ACM.

[24] S. Muhammad, P. Kohldorfer, and E. Leitgeb. Channel modeling for terrestrial free space optical links. In *Transparent Optical Networks, 2005, Proceedings of 2005 7th International Conference*, volume 1, pages 407–410 Vol. 1, July 2005.

[25] B. Mukherjee. *Optical WDM Networks.* Springer, Jan 2006.

[26] B. Nakhkoob, M. Bilgi, M. Yuksel, and M. Hella. Multi-transceiver optical wireless spherical structures for manets. *IEEE J.Sel. A. Commun.*, 27(9):1612–1622, 2009.

[27] M. S. Navin Kumar, Nuno Lourenco and R. Aguiar. Visible light communication systems conception and vidas. *IETE Tech. Review*, 25:359–367, 2008.

[28] W. Ng, A. Walson, G. Tangonan, J. Lee, and I. Newberg. Optical steering of dual band microwave phased array antenna using semiconductor laser switching. *Electronics Letters, IEEE-*, 26(12):791–793, June 1990.

[29] G. Pang, K.-L. Ho, T. Kwan, and E. Yang. Visible light communication for audio systems. *Consumer Electronics, IEEE Transactions on*, 45(4):1112–1118, Nov 1999.

[30] G. Pang, T. Kwan, H. Liu, and C.-H. Chan. Led wireless. *Industry Applications Magazine, IEEE*, 8(1):21–28, Jan/Feb 2002.

[31] S. Se, H.-K. Ng, P. Jasiobedzki, and T.-J. Moyung. Vision based modeling and localization for planetary exploration rovers. *55th International Astronautical Congress, Vancouver, Canada*, 2004.

[32] H. Sugiyama and Nakagawa. Experimental investigation of modulation methods for visible light communications. *IEEE Transactions on Communications,*, 89(12):3393–3400, Dec 2006.

[33] A. Tang, J. Kahn, and K.-P. Ho. Wireless infrared communication links using multi-beam transmitters and imaging receivers. In *Communications, 1996. ICC 96, Conference Record, Converging Technologies for Tomorrow's Applications. 1996 IEEE International Conference on*, volume 1, pages 180–186 vol.1, Jun 1996.

[34] H. T.Saito and Nakagawa. Inter-vehicle communication and ranging method using led rear lights. *Proceedings of the fifth IASTED international conference on Communication Systems and Networks*, 5:278–283, Aug 2006.

[35] H. Willebrand and B. Ghuman. *Free Space Optics:Enabling Optical Connectivity in Today's Networks*. Sams, 2002.

[36] M. Yuksel, J. Akella, S. Kalyanaraman, and P. Dutta. Free-space-optical mobile ad hoc networks: Auto-configurable building blocks. *Wirel. Netw. 15*, 15(3):295–312, 2007.

[37] Y. Zang, L. Stibor, H. Reumerman, and H. Chen. Wireless local danger warning using inter-vehicle communications in highway scenarios. In *Wireless Conference, 2008. EW 2008. 14th European*, pages 1–7, 2008.