

Rate Adaptation in Visual MIMO

Ashwin Ashok[‡], Marco Gruteser[‡], Narayan Mandayam[‡], Taekyoung Kwon^{‡*}

Wenjia Yuan[†], Michael Varga[†], Kristin Dana[†]

[‡]WINLAB, Rutgers University, 671 Route 1 South, North Brunswick, NJ, USA

[†]Department of ECE, Rutgers University, 94 Brett Road, Piscataway, NJ, USA

^{*}School of Computer Science and Engineering, Seoul National University, Seoul, Korea

Email: [‡]{aashok,gruteser,narayan,tkkwon}@winlab.rutgers.edu,[†]{wenjiay,mfvarga}@rutgers.edu,[†]kdana@ece.rutgers.edu

Abstract—We propose a rate adaptation scheme for visual MIMO camera-based communications, wherein parallel data transmissions from light emitting arrays are received by multiple receive elements of a CCD/CMOS camera image sensor. Unlike RF MIMO, multipath fading is negligible in the visual MIMO channel. Instead, the channel is largely dependent on receiver perspective (distance and angle) and visibility issues (partial line-of-sight availability and occlusions). This allows for slower adaptation but requires the adaptation algorithm to choose among a more complex set of modes. In this paper, we define a set of operating modes for visual MIMO transmitters and propose a rate adaptation scheme to switch between these modes. Our Visual MIMO Rate Adaptation (VMRA) is a packet based rate adaptation protocol that bases its rate selection decisions on the packet error rate feedback. Using trace-based simulation results for a vehicle-to-vehicle communication scenario, we illustrate how our VMRA algorithms can adapt over distance as well as visibility variations in an optical link and achieve a higher average throughput.

I. INTRODUCTION

The increasingly ubiquitous use of cameras and light emitting devices, for example, in cell phones, cars, laptops, music players, and surveillance systems creates an exciting novel opportunity to build camera based optical wireless communication systems and networks based on a concept we call Visual MIMO [1]. In this concept, optical transmissions by an array of light emitting devices, which we refer to as light emitting array (LEA), are received by an array of photodiode elements (e.g. pixels in a CMOS camera). Examples of LEAs are arrays of light emitting diodes, pixel arrays on LCD or plasma screens, or digital micromirror devices combined with a light source. The pixels 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 and by transmitting using multiple elements of the LEA such camera based communications can take advantage of data rate gains from multiplexing and/or diversity using such an inherent multiple input multiple output (MIMO) setting in the system. Visual MIMO promises to achieve higher information capacity [1] than conventional optical wireless systems that use photodiode receivers especially in mobile settings where ranges greater than tens of meters are required. Thus, one potential application area is vehicle-to-vehicle communications, where LED lights could act as transmitters and in-vehicle cameras as receivers. Visual MIMO techniques, however, could also facilitate communications be-

tween large screens and handheld devices or communications between handheld devices such as mobile phones. This may be particularly attractive in dense settings, such as conferences, since visual MIMO communications are virtually interference-free due to its high directionality. The PixNet project [2] has demonstrated that using an LCD monitor as transmitter that data rates of the order of Mbps are possible. Such higher data rates are typically achieved by multiplexing information data streams over several transmitter elements.

Due to negligible multipath fading in optical channels the data-rates achievable (i.e., the degree of multiplexing) in visual MIMO depend primarily on the distortions in the visual channel rather than multipath fading, unlike RF. These distortions are typically observed as distortions in the size and shape of the image, partial visibility of an image and even interference between images of two different transmitter elements in the scene due to perspective projection onto the camera sensor and image blurring. In mobile settings, the quality of the visual MIMO link varies significantly with the variation in these distortions which depends on the camera receiver perspective.

This suggests that the throughput of visual MIMO links can be significantly improved through rate adaptation techniques, which adapt the transmission data rate to the receiver perspective. Particularly in a vehicular setting with front and rear facing visual MIMO transceivers the receiver could provide feedback and make rate adaptation possible. Rate adaptation has, of course, been the focus of extensive study for RF communication systems (e.g., [3]–[5]). The visual MIMO rate adaptation challenge differs in that (i) the primary challenge lies in MIMO mode adaptation, and (ii) visual MIMO modes present a more complex set of choices than what RF rate adaptation algorithms have explored. Mode adaptation is the more significant problem in visual MIMO, since visual MIMO transmitters can employ a much larger number of transmitter elements than typical RF MIMO systems (due to their operation in the optical spectrum). Selection of the mode requires more complex decisions than the typical rate-up or rate-down decisions of many RF rate adaptation algorithms because the adaptation algorithm has to choose a perspective-appropriate subset of transmitter elements for multiplexing. This subset has to be chosen such that all of the elements are visible to the camera (within field-of-view and not occluded) and the transmitter elements do not interfere with each other.

Interference between transmitter elements typically occurs when link distance increases and the images of the elements start to overlap in the camera view.

To address these challenges, this paper first defines a set of visual MIMO transmission modes for an $N = N_r \times N_c$ LED array transmitter and develops adaptation algorithms to switch to perspective-appropriate modes. The scheme uses packet error feedback to choose the appropriate set of LEDs both over changing distance and changing partial visibility conditions. To identify the set of LEDs suitable for multiplexing, we propose a probing scheme that uses certain spatial patterns and a block CRC scheme that uses separate CRCs for the blocks of information sent from each transmitter element. Using trace-based simulations, we compare their performance with a baseline solution that uses an exhaustive probing search through all LED elements. The simulations are based on a car-following video sequence where the car brake light LEDs are assumed to be the transmitter elements.

II. RELATED WORK

Rate adaptation protocols have been largely studied, designed, experimented and implemented for the RF channel over the years and especially for 802.11 based networks [3]–[5]. In IEEE 802.11 networks, the current standard feedback from a receiver to a transmitter is only the presence or absence of an ACK frame. Many rate adaptation schemes also try to rely on physical layer metrics such as signal-to-interference noise ratio (SINR) and bit error rate (BER) to obtain the channel condition [6]. Such schemes typically apply to cellular networks (e.g WiMAX and 3GPP) since cellular networks have a wide range of SINR values [7]. There have also been both simulation and implementation efforts in 802.11 networks to find out what is the best rate along with the choice of spatial diversity or spatial multiplexing for the wide range of channel state among MIMO antennas [8], [9].

Similar to a few RF mechanisms an optical transmitter may also seek to adapt to the link condition change based on a few physical layer parameters. For optical channels, Diana and Kahn [10] investigated how to adjust the parameters of FEC techniques, such as repetition codes and rate-compatible punctured convolutional (RCPC) codes, for Infra-Red links (IR) based on BER metric. In their paper, Garcia-Zambrana [11] do not adapt the bit-rate directly but seek to enhance the peak-to-average optical power ratio (PAOPR) by inserting the silence period while keeping the average optical transmitted power for FSO links. More recently, Grubor et al. [12] investigated how the power and information bits can be allocated among OFDM subcarriers for throughput maximization. Applying above techniques such as repetitive codes, RCPC, silence periods to visual MIMO where the camera sampling rates are limited add complexity and significant overheads which can depreciate the effective throughput of the system.

III. PERSPECTIVE DEPENDENT DATA RATES

In Visual MIMO, the achievable data rate depends largely on receiver perspective. In RF MIMO communication systems,

multipath fading can lead to independent parallel channels between antenna pairs. This allows multiplexing of information over these independent channels. With N independent channels used for multiplexing, N symbols can be transmitted simultaneously, leading to an N -fold gain in data rate. Although multipath fading is negligible in the optical spectrum considered here, independent parallel channels also exist in visual MIMO. Consider an ideal full frontal view onto a light emitting array at close distance. The light from different transmitter elements will fall onto different pixels in the camera image. These pixels can be independently read out, which allows the same multiplexing of information across different transmitter-pixel pairs. In this ideal case, the Shannon capacity for the visual MIMO system with multiplexing can be characterized as in RF MIMO by $C_m = NW \log_2(1 + \gamma)$, where W is the sampling rate (frame-rate of the camera) and γ is the signal-to-noise ratio (SNR) in a single LED-camera communication system, as discussed in earlier work [1]. This assumes that SNR differences from LED to LED are negligible. As in RF MIMO, operation in a diversity mode is also possible. In this mode the same bits are signaled on all (or a subset) of transmitter LEDs. This leads to a stronger signal at the receiver and usually less errors. Note that this is also possible when LEDs are blended together in the image, the signals from multiple LEDs will simply be combined on the receiver pixels. This leads to a capacity of $C_d = W \log_2(1 + N_d \gamma)$, where N_d denotes the number of LEDs transmitting in this diversity mode. The key difference to RF MIMO lies in larger N and very different channel distortions introduced by the optical channel.

The SNR γ of the visual MIMO system with single LED transmitter is given as,

$$\gamma = \begin{cases} \frac{\kappa P_t^2 d^{-2}}{q R P_n W f^2 l^2} & \text{if } d < d_c \\ \frac{\kappa P_t^2 d^{-4}}{q R P_n W s^2} & \text{if } d \geq d_c \end{cases} \quad (1)$$

where P_t is the transmit power, q is charge of an electron, R is the receiver responsivity, P_n is the photon shot-noise power density (Watts/area), f is the focal length of the camera, W is the camera frame rate, l is LED diameter, s is the pixel side length and d is distance. The parameter $d_c = fl/s$ is the distance at which one LED projects onto exactly one pixel and 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 [13].

A. Modeling Channel Distortions

In practice, the availability of parallel channels will be affected by visibility issues, perspective distortions, and lens blur.

Visibility: Like any other optical wireless system, visual MIMO requires line-of-sight. An outage will generally occur when none of the transmitter elements are directly visible in the camera image. Only rarely will reflections of the transmitter image be strong enough to be detected by the receiver. A key difference of visual MIMO systems is, however, that

only *some* of the transmitter elements may be visible. This can occur when random objects partially obstruct the line of visibility between the camera and the transmitter LEDs. It can also occur when the transmitter is only partially within the field of view of the camera or due to weather effects such as snow flakes and rain drops. Such partial visibility means, that fewer parallel channels are available and the maximum achievable gain will be degraded. We model such visibility issues through an index function $V(n)$, which for each LED $n \in 1 \dots N$ takes a value 1 when the LED is visible or 0 when it is obscured. The instantaneous multiplexing and diversity capacities C_m and C_d can then be obtained by replacing the total number of LEDs N with the number of visible LEDs $\sum V(n)$. Clearly, visibility often changes over time. Modeling such visibility changes is beyond the scope of this article.

Perspective Distortion: Changes in viewing angle or distance lead to perspective distortions that can also affect the availability of such independent transmitter-pixel channels. Consider again the full-frontal view onto a transmitter array, but now from larger distance. As distance increases, the image of the transmitter will become smaller. Eventually, light from multiple transmitter LEDs will shine onto the same pixels. At this point the light from these transmitters can no longer be independently read out and the achievable multiplexing gain is again reduced. With changes in viewing angle, the image of the transmitter will be skewed. This can lead to situations where part of the transmitter LEDs shine onto the same pixels, while other transmitter LEDs can still be independently received.

Given the camera parameters as well as the location of the transmitters and camera in 3D space, perspective projection analysis [14] can be used to determine which pixels detect light from which transmitters. For simplicity, let us focus here on the effect of distance. Given a fixed-focal length f of the camera, a spatial distance α between the centers of two adjacent LEDs, and the distance to the camera d , we can calculate the separation of the LEDs on the camera image plane using projection. To be able to independently read the signal from two LEDs, let us assume that a minimum image separation η is required. Thus, multiplexing over all N parallel channels is only possible for distances d below the threshold $d^* = \frac{f\alpha}{\eta}$. In practice, α and η are likely to be fixed for a visual MIMO system, since it will be difficult to dynamically increase the spacing between LEDs or improve the resolution of the camera. It is possible, however, to indirectly modify α by leaving some LEDs unused. This effectively increases the separation between LEDs in use but decreases the multiplexing gain.

Lens Blur: In addition to perspective distortions, lens blur can lead to blending of the images from two different transmitters. The amount of blur in a camera image is a characteristic of the camera lens and specific to the type of lens used in the system. Such blur is often modeled with a Gaussian blur filter. That is, a (blurred) image Z_{im} is the output of a Gaussian blur filter whose input is an ideal image Z_{ideal} , that is $Z_{im} = Z_{ideal} * g_{blur}$ where $'*$ ' represents a 2D-convolution operation and g_{blur} is a 2D-Gaussian function with zero mean

and standard deviation σ_{blur} measured using experiments [14]. In this paper, we will assume that the Gaussian blur from two LEDs can be separated and independently read out, if the distance between the centers is greater than the full-width-at-half-maximum (FWHM) of the Gaussian blur function. FWHM is often used as a parameter for image resolution in analyzing fine detailed astronomical and medical images [15], [16]. That means, we define the minimum necessary separation in the image plane η as follows.

$$\eta = 2\sqrt{2\ln 2}\sigma_{blur} \quad (2)$$

The rate adaptation problem then is to choose transmission modes that exploit the available parallel channels while keeping the error rate low. Multiplexing across more transmitter LEDs will lead to higher data rates, but including an LED that is occluded in the image, for example, would lead to bit errors. We will further discuss and analyze different possible transmission modes next.

B. Transmission Modes

A transmission mode is a certain assignment of multiplexing and diversity functions to the set of LEDs. In one mode, which we refer to as full multiplexing, bits are multiplexed over all LEDs. In another mode, all LEDs would be used to transmit the same bits. We refer to this as full diversity mode. In between these extremes, lie many other possibilities where only subsets of LEDs are used for multiplexing or some subsets of LEDs are grouped for diversity operation. We therefore define a transmission *mode* as a set of m non-overlapping subsets chosen from the N LEDs of the array where each set $i = 1, 2, \dots, m$ is a group of D_i LEDs such that $\sum_{i=1}^m D_i \leq N$. Data bits are transmitted using the LED array in a way that LEDs in each subset transmit the same bits (diversity) and information is multiplexed over the different subsets. The *modes* ($m = N, D_i = 1 \forall i = 1, 2, \dots, m$), ($m = 1, D_i = N \ \& \ i = 1$) correspond to full-multiplexing and full-diversity cases respectively.

As an example, let us discuss some possible modes that can be obtained on a 3×3 transmitter array by choosing subsets of LEDs for multiplexing. Assume that each LED is separated from the next LED in the same row and column by α units. Recall that the full multiplexing mode (mode 1 in Fig. 1) can be used only up to a critical distance of d^* and would provide a multiplexing gain of 9. If we now consider mode 2, which leaves LEDs $\{(1, 2), (2, 1), (2, 3), (3, 2)\}$ unused, the spatial separation between active LEDs increases to $\alpha\sqrt{2}$. This increases the maximum distance to $\sqrt{2}d^*$, albeit at a reduced multiplexing gain of 5. In mode 3, we also switch off LED $(2, 2)$, which allows communication for all $d \leq 2d^*$. The system can multiplex over the remaining LEDs $\{(1, 1), (1, 3), (3, 1), (3, 3)\}$, yielding a multiplexing gain of 4. The largest range is provided by the full diversity mode.

Other modes can be required to address visibility of the transmitter to the camera. Weather conditions such as fog, rain or snow can significantly reduce the resolvability due to

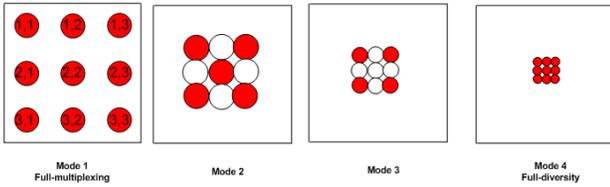


Fig. 1. Illustration of 3×3 LED array modes

modes	(d_{min}, d_{max})	(N_m, N_d)	Rate
1	$(0, d^*]$	(9,1)	$C_{mimo}(d^*)$
2	$(d^*, \sqrt{2}d^*]$	(5,1)	$C_{mimo}(\sqrt{2}d^*)$
3	$(\sqrt{2}d^*, 2d^*]$	(4,1)	$C_{mimo}(2d^*)$
4	$(2d^*, d_{max}]$	(1,9)	$C_{mimo}(d_{max})$

TABLE I
MODES AND RATES FOR $N = 3 \times 3$ LED ARRAY

occlusions over time and blurring. For example, if the right half of the LED array was obscured, a multiplexing mode should only include LEDs from the left half. The resolvability is also reduced when the camera is at an angle to the transmitter. One possible mode to address the resulting skewed images is to use a combination of multiplexing and diversity where a group of LEDs could coordinate to attempt to provide sufficient brightness for a particular bit, but individual groups could be spaced sufficiently far apart to reduce the chances of blur among the groups. In the vehicular application context, we expect that visibility and distance distortions are more prominent than such angular distortions and the remainder of the paper will focus on these.

C. The Rate Adaptation Problem and Error Model

Due to the large number of possible transmission modes, the visual MIMO rate adaptation problem lies in efficiently choosing a transmission mode that maximizes throughput. We assume an on-off-keying communication system where feedback in the form of acknowledgments is available. The feedback channel could be realized through a reverse visual MIMO link.

We base our design and simulations on the following packet error model. Recall that for an independent and identically distributed (i.i.d) stream of bits framed into L bit packet sequences the packet error rate (PER) is given as $PER = 1 - (1 - P_e)^L$, where P_e is the bit error probability. A received packet is erroneous if at least one bit in the packet or equivalently one LED is in error. Bit errors may be caused due the AWGN background light noise and also due to visual distortions we discussed. In this context, we consider that LEDs will be in error when their centers cannot be resolved from any adjacent LED in the image space. Let us consider a visual MIMO LED array with N LEDs such that the transmission modes are defined using m multiplexing sets and each set containing D_i LEDs ($i = 1, 2 \dots m$) used for diversity, as discussed in section III-B. Then the packet error ratio can be expressed as,

$$PER = \begin{cases} \frac{1}{m} \sum_{i=1}^m 1 - \sum_{j=1}^{D_i} ([1 - Q(\sqrt{D_i \gamma})]^L) V(j) & \text{if(A)} \\ 1 & \text{if(o.w)} \end{cases} \quad (3)$$

where (A) is the condition $\min(\alpha_{(im)}) > 2\eta$. $\alpha_{(im)}$ is the set of the image separation values (in pixels) between any two multiplexing sets and η is the resolvability factor from equation 2. Given the spatial separation α the separation in the image at a distance d be found by perspective projection equations [14] as $\alpha_{(im)} = \frac{f\alpha}{ds}$. The decisions on the number of sets m and LEDs D_i for each set depends on this resolvability condition (A) which infact depends on the distance. In this way this model accounts for errors due to distance based distortions. The errors resulting from AWGN background noise are addressed through the term $Q(\sqrt{\gamma})$ which is the average BER for a single transmitter (single LED) with an SNR γ for OOK modulation in an AWGN channel (shot-noise optical channel). The errors due to visibility are factored into this model using the $V(i)$ factor which denotes if an LED i is visible (1) or obscured (0).

IV. VMRA-RATE ADAPTATION ALGORITHMS

In this section we detail our proposed algorithms for our receiver-based rate adaptation protocol VMRA that adapts its transmission data rate over distance and visibility variations in the visual channel. The algorithms use the packet error feedback information to choose the appropriate set of LEDs both over changing distance and changing partial visibility conditions. In our design the LED array transmitter sends a continuous stream of packets - each appended with a CRC - that are decoded at the camera receiver. Upon each successful packet receipt the receiver sends back an acknowledgment (ACK signal) back to the transmitter over a reverse visual MIMO feedback link. The transmitter then flags the transmission as erroneous based on a packet error ratio computed over a time window of T sec (can be of the order of tens of frame-time),

$$PER = 1 - \frac{(\#_ACKs_in_time_T)}{\#_packets_transmitted_in_time_T} \quad (4)$$

The transmission is termed successful as long as the PER is below a preset threshold PER^* (typically 10-15%). In our protocol the transmitter data rate is adapted to the distance variations in the channel by switching to the perspective-appropriate mode as described in section III-B. Since the data rates in our system is primarily dependent on the number of multiplexing/diversity LEDs over each iteration of the adaptation, we design our algorithms to output the set of indices of LEDs that can be multiplexed β_m ($|\beta_m| = N_m$) and that can be used for diversity β_d ($|\beta_d| = N_d$). To adapt the transmissions to the visibility variations in the channel each of our algorithms use different techniques to determine the set of visible LEDs over each iteration of the adaptation. Such

techniques will be discussed along with a detailed description of the algorithms in the following sections.

A. Exhaustive Visibility search VMRA

This algorithm uses an elementary approach to find the erroneous LEDs (bits) by exhaustively searching over all the LEDs ($N_r \times N_c$ of them) in the transmitter array and output the indices of LEDs that are to be multiplexed β_m or use diversity β_d .

- To find the erroneous LEDs, each LED is set to transmit a unique l bit training sequence (common bit-sequence + unique ID of each LED), one after the other, such that each bit is transmitted (each LED is ON) only once over each image frame and those LEDs that are not in error are indexed into a set V (visible set of LEDs). Since errors may also occur due to even distance variations the algorithm tries to determine the most appropriate *mode* for transmission based on an estimate of the distance between the transmitter and receiver it calculates in the next step.
- The algorithm first finds a pair of LEDs within the set V that are separated the farthest in space. Using the image separation between these two corner LEDs, an estimate of the distance d is calculated using perspective projection theory [14]. Though not addressed in detail this paper this method may also be extended to using all four corner LEDs to also detect skew from different viewing angles to adapt to angular distortions. The perspective appropriate *mode* is determined by checking if the distance estimate is within the distance ranges (d_{min}, d_{max}) of a particular *mode* (as discussed in section III-B). The distance estimation to determine the perspective appropriate mode is necessarily only when the system is using any of the multiplexing modes. At a large distance LEDs may merge together or the LED signal is too weak to be decoded. In such a case where all the LEDs are marked erroneous by the exhaustive search the algorithm defaults the mode to a full-diversity mode.
- When an error occurs in the transmission ($PER > PER^*$) or when a time-out (t_{out} set to a large value like 10-20sec) occurs the algorithm reinitates the exhaustive search and adaptation procedure. If transmissions are successful for the period of t_{out} the algorithm steps up the transmission rate to next *mode* with a higher data rate.

B. Framing based algorithms

The exhaustive search to detect erroneous LEDs may prove wasteful particularly when the the size of array is very large. In such cases rate adaptation may not perform as fast as it is needed to especially in mobile scenarios. Given the spatial setting of the LED array it may be possible to find erroneous LEDs by framing packet transmissions in a spatially coordinated manner over the array. By coordinating such packet transmissions over space and time it may be possible to ‘track’ the bit transmissions from packets not only in time

but spatially as well. In this aspect, we propose two possible techniques of such spatio-temporal framing,

Bit per LED : The LED transmitter array is set to transmit packets of constant size L bits such that each LED transmits one bit. Each packet contains a C bit CRC for error detection. When is the system uses ‘multiplexing’ each LED transmits an independent bit from the data packet, while when using ‘diversity’ all the LEDs of the transmit array transmit the same bit. The significance of such a framing technique is that it is practical and easily implementable.

Block per LED: The data packets are split into blocks of data bits. These blocks are spatially framed such that each LED transmits bits corresponding to individual blocks from a packet. Each block is also appended with a C bit CRC for error detection. Only when the system uses spatial diversity on a set of LEDs then the transmitter frames the packets such that the diversity LEDs will transmit the same bits from the same block. Though such a configuration may be relatively complex when compared to **Bit per LED** the advantage is that detecting the erroneous LEDs (and hence bit errors) becomes very easy as packet errors can be detected just by indexing the erroneous blocks and mapping it to the corresponding LED indices.

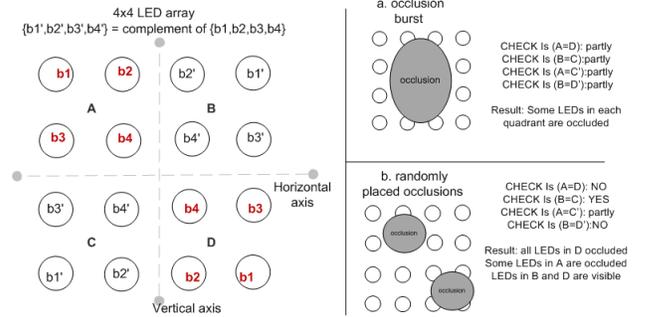


Fig. 2. Probevisibility pattern for a 4×4 LED array in *Probe VMRA*

1) *Probe-VMRA*: In this section we propose a *Probe VMRA* algorithm that uses the *bit per LED* framing technique for packet transmissions. This algorithm uses a unique probe function *ProbeVisibility()* to detect occlusion in the visual channel by using a smart spatial patterning of bits on each row and column of the array. The spatial patterning for *Probevisibility()* is such that any square LED array can be reconfigured to transmit similar bits on each spatial quadrant of the array and complement bits on each side of the horizontal and vertical axis at the center of the array (as shown in Fig. 2). The fundamental idea behind using such a pattern is that, by having a copy of the bit and its complement the detection of bit-flips double efficient. This simplifies the detection of erroneous LEDs’ locations. Other possible patterns to detect occlusions are to use all ones/zeros or alternating ones/zeros. The issue with such approaches is that, in cases of occlusions in the channel the pixel intensity of a bit depends on the object occluding the camera view. If the object is white in color then the pixel intensity will remain high and a bit 1 is retained as bit 1 thus not flagging an error. Alternating ones and zeros

may prove helpful but it may only detect if an occlusion has occurred but may not be possible to exactly reveal the occluded LED positions most of the time. In Fig. 2 we illustrate two practical cases in which such a complement bit pattern can work (shown for a 4×4 array but can be extended to any square array). But we also realize that such a probing may not be always error-free. For example, a false alarm can occur when bit b1 in quadrant A gets corrupted due to stronger background noise. In this case the probing detects an occlusion while actually there is no occlusion but noise. Also, the probing may miss an error such as when all the b1 and b1' bits are corrupted such that b1 is detected as b1' and b1' as b1. In this case the probing returns *no-error* while actually there has been occlusion at four locations. But such occurrences are very rare in reality because it is highly uncommon that an occlusion is of the form that can create exactly complementary effects on different spatial regions of the array. False alarms are also rare because the ambient photon noise is typically uniformly averaged over the detector area. Hence we rule out such possibilities and consider the most typical cases in our algorithm design.

- When a transmission *error* is declared in the system ($PER > PER^*$), the algorithm first checks for the occlusion by initiating the *ProbeVisibility()* where the LEDs are set to transmit bits based on the spatial pattern mentioned earlier. The function returns the set of visible LEDs indices V and a probe flag $pFLAG$ (*true* when occluded).
- If the occlusion is full then the probe function would return $V = \emptyset$. In other cases, the algorithm increments its *mode* so as to accommodate for any distance variation.
- If the error was due to both partial occlusions and distance variations the algorithm first sets its *mode* and then re-probes using *ProbeVisibility()* in the next iteration.
- If there is no occlusion then the function returns a probe-error $pError$ and the algorithm increments to the next transmission *mode* and checks for errors in transmission.
- In case of multiplexing modes the algorithm sets the LEDs in the visible set V for multiplexing. In diversity mode the algorithm sets all the LEDs for transmitting similar bits (regardless of whether they are occluded or not). If the algorithm is already in its diversity mode then the algorithm resets itself back to full multiplexing and reinitiates the probing for occlusions.

2) *Index-VMRA*: Here we present a method that obviates the need for exhaustive search or spatial probing to detect the presence of occlusion. This approach uses the *block per LED* framing technique of sending packets in the form of independent blocks of bits for each LED. Since each block is appended with a CRC the system can keep track of the erroneous LEDs on the fly by indexing the erroneous block of each packet during a transmission error and indexing those LEDs into a set E . We denote the set of usable LEDs in a *mode* as $U(\mathbf{mode})$. Using this approach we propose the *Index VMRA* algorithm (Algorithm 1).

- In each iteration, the algorithm first indexes all the erroneous LEDs of the set $U(\mathbf{mode})$ into E .
- If the set E is full (all LEDs used in a particular *mode* are in error) then the algorithm shifts to the next *mode* and updates $U(\mathbf{mode}) = \text{set of usable LEDs in the mode}$.
- If the set E is not full, then all LEDs in $U(\mathbf{mode})$ that are not in error ($U(\mathbf{mode}) - E$) are indexed into the set I .
- If the transmission is in multiplexing mode then all the LEDs in set I are used for multiplexing. In diversity mode, since all LEDs transmit the same bits it may happen that the CRC bits may be corrupted resulting always in error. In this case we start with using all the LEDs in diversity mode and reduce the set by one (LED in any row r and column c) in each subsequent iterations until the transmission is successful ($PER < PER^*$).
- Since it is possible to determine the LEDs that are erroneous in each iteration, over the adaptation period, the erroneous LEDs are set to transmit 'training packets' (packets containing alternate ON-OFF sequences 101010...L bits) to determine if the channel is less 'noisy'. Upon successful reception of these bits the receiver acknowledges by sending back an ACK over the visual MIMO feedback channel and those LEDs are indexed to be used for data transmission again.
- The algorithm reinitiates the adaptation procedure when an error occurs in the transmission ($PER > PER^*$) or when a time-out (t_{out} set to a large value like 10-20sec) occurs. If all transmissions are successful for the period of t_{out} the algorithm steps up the transmission rate to next *mode* with a higher data rate.

V. PERFORMANCE EVALUATION

We evaluate the performance of our VMRA protocol in terms of the average throughput achieved by its candidate algorithms over the distance and visibility variations in the channel and compare it with an oracle solution (referred to as *ideal*, that has the power to adapt over the visible set of LEDs to any type of occlusion and distance variation by using the best *mode* for maximizing the throughput). We then elaborate the adaptation behavior of our two framing based algorithms; *Probe VMRA*, *Index VMRA*. Our evaluation uses a trace-driven simulation using traces of input derived from a realistic vehicle-vehicle communication setting.

A. Obtaining trace inputs

As our source for our traces, we used the video of a real car on a highway with a 3×3 LED array configuration in its brake-lights. The video was captured using our Basler Pylon piA640 camera fitted onto another car at a frame-rate of 60fps and 640×480 resolution. Short sequences of image frames in the video were analyzed partly manually and partly using software to generate the two dataset traces (a. *distance* and, b. *distance - occlusion*) that were used as our test inputs for simulation. To obtain the *distance* trace we used a basic tracking technique from computer vision [14]

Algorithm 1 Index VMRA

INPUT: packet error rate PER , total# of modes μ , $U(mode)$
OUTPUT: LED index sets β_m (multiplexing) and β_d (diversity)
while iteration $i \neq 0$ **do**
 if $t_{out} \& PER < PER^*$ **then**
 if $mode \neq 1$ **then**
 $mode = mode - 1$
 end if
 end if
 if $mode = \max(mode)$ **then**
 if $PER < PER^*$ **then**
 $\beta_d = \{all\ LEDs\}$
 continue
 end if
 else
 for all $(r, c) \in U(mode)$ s.t. $PER(r, c) > PER^*$ **do**
 index (r, c) in E
 end for
 if $E = U(mode)$ **then**
 $mode = \text{modulo}(mode + 1, \mu)$; $I = U(mode)$
 else
 $I = U(mode) - E$
 end if
 end if
 if $mode \neq diversity$ **then**
 $\beta_m = I$; $\beta_d = \emptyset$
 else
 $\beta_m = \emptyset$; $\beta_d = \beta_d - (LED(r, c))$
 end if
end while

to estimate the distance between the LED two brakelights x of the car in each image frame of the video. This inter-brakelight separation in the image (in pixels) was used to compute the distance d between the camera and the car in each frame using perspective projection mapping $d = f \frac{l}{x}$, where f is the calibrated camera focal length and $l = 1.5\text{m}$ is the typical spatial distance of separation between two brakelights in a car. The *distance – occlusion* trace was obtained by analyzing a set of frames from the video and manually creating a dataset where the number of distinguishable LEDs visible in each frame was noted. We also noted down if the transmitter was fully visible ($visibility = 1$), fully occluded ($visibility = 0$) or partially visible ($visibility = 0.25$ or 0.5 or 0.75) depending on the image area of the array visible. Based on the number of LEDs distinguishable in each frame, the transmission *modes* were manually estimated and used as ground-truth. The data samples in both traces were spaced by one frame time (1/60 secs). Figure 3 shows a few samples of images that were analyzed.

B. Simulation Methodology

We simulated the adaptation behavior and computed the performance of our VMRA candidate algorithms in MATLAB for the two types of trace inputs using a common simulation methodology. In the simulation each algorithm set to adapt its transmission rate parameter R_{tx} to the number of visible LEDs and the transmission *mode* based on the PER determined (equation 3) in each iteration. Since SNRs in a visual MIMO channel are typically very large [1] we assume the probability of bit error due to AWGN background noise is zero. Thus the PER takes values depending only on the errors due to visibility and/or distance change at any LED i . Hence PER from equation 3 reduces to,

$$PER = \begin{cases} 1 & \text{if } \{\min(\alpha_{(im)}) < 2\eta\} \text{ or } V(i) = 0 \\ 0 & \text{if otherwise} \end{cases} \quad (5)$$

While detecting error due to visibility is done independent of adaptation (probing or indexing using framing) the error due to distance change is detected by comparing the *mode* determined by the algorithm with that estimated based on distance and a resolvability threshold ($\eta = \text{FWHM}$ of Gaussian blur) using perspective projection theory as,

$$mode = \begin{cases} 1 & \text{if } d \leq \frac{f\alpha}{\eta_s} \\ 2 & \text{if } \frac{f\alpha}{\eta} \leq d \leq \sqrt{2} \frac{f\alpha}{\eta_s} \\ 3 & \text{if } \sqrt{2} \frac{f\alpha}{\eta} \leq d \leq 2 \frac{f\alpha}{\eta_s} \\ 4 & \text{if } d > \frac{f\alpha}{\eta_s} \end{cases} \quad (6)$$

When $PER = 1$ a transmission *error-flag* is raised and the adaptation algorithm is initiated which would set a rate R_{tx} based on the *mode* and the number of visible LEDs specified by the algorithm. The data rates for each transmission *mode* are summarized in the look-up table II for a 3×3 array configuration. To retain uniformity in all cases we set the rate $R_{tx} = 0$ if the LED array is fully occluded.

In order to understand the behavior of the algorithms we recorded a few parameters output at each iteration, such as, a.the transmission *modes*, b.if transmission *error* flag, c.*ProbeError* flag (only for *Probe VMRA*), and d.transmission rate R_{tx} set by the algorithm. We then computed the average throughput ρ for each algorithms and for each of the two traces analyzed as,

$$\rho = \frac{T}{B} \sum_i^{B/T} R_{tx}(i)(1 - error(i))(1 - ProbeError(i)) \quad (7)$$

where B is the time width of the window of data-trace. As *Exhaustive Visibility search* and *Index VMRA* do not use the probing to detect occlusion we set $ProbeError = 0$ when computing the average throughput for these algorithms.

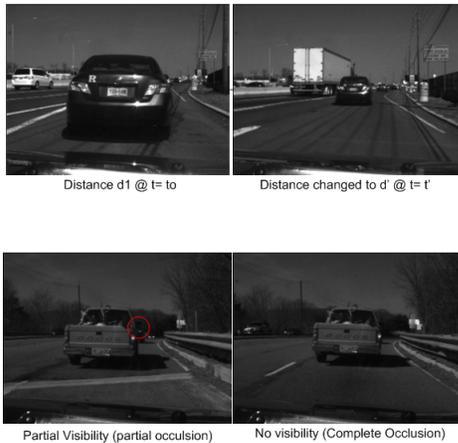


Fig. 3. Sample video frames analyzed for data trace

<i>mode</i>	(d_{min}, d_{max}) [m]	(N_m, N_d)	Rate [kbps]
1	(0, 17.6]	(9, 1)	1192
2	(17.6, 24.91]	(5, 1)	543.9
3	(24.91, 35.2]	(4, 1)	338.54
4	$d > 35.2$	(1, 9)	59.24

TABLE II

RATE CHOICE FOR EACH *mode* FOR $N = 3 \times 3$ LED ARRAY WITH $\alpha = 2$ CM INTERLED SPACING

C. Trace-driven Results

In Fig 4 we plot the average throughput ρ from equation 7 for our proposed VMRA algorithms and compare it with the *ideal* performance from the *oracle* solution. We clearly see that our framing-based approaches achieve close to an ideal performance for distance variations as well as visibility variation traces.

We now elaborate on the adaptation behavior of VMRA protocol using the trace-based results for its framing-based algorithms a. Probe VMRA and, b. Index VMRA. We illustrate each algorithm's adaptation behavior over time by plotting the output of each iteration of the algorithm for the *distance* trace and then repeat the same for the *distance - occlusion* trace. Fig. 5 shows the performance of the *Probe VMRA* algorithm for the *distance* trace over time. We see that whenever an *error* is declared then the algorithm increments its *mode* until no more *error* is declared. Once the system reaches the mode 4 (diversity) then the algorithm resets back to mode 1 (full-multiplexing). As *Index VMRA* also uses the same approach for adaptation over distance the performance will be the same. In Figure 6 and 7 we show the performance of the *Probe VMRA* and *Index VMRA* algorithms for the *distance - occlusion* trace. Observe that, in the region A where the transmitter is partially visible, the *Probe VMRA* always first initiates the probe for detecting occlusion of the LEDs and only if a *ProbeError* is declared the algorithm changes its *mode*. Thus this algorithm always incurs a 'one-iteration' delay when adapting to the partial-occlusion of the array. The *Index VMRA*

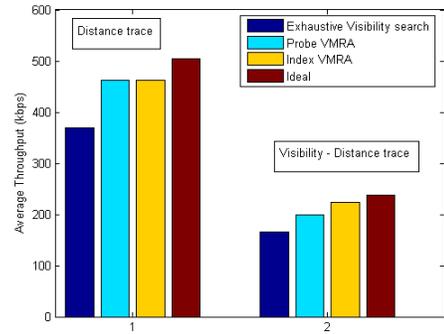


Fig. 4. Summary of throughput performance

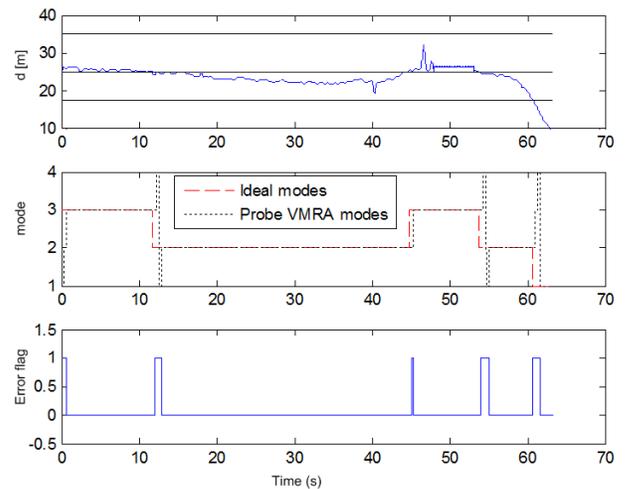


Fig. 5. Probe VMRA performance over *distance* trace

on the other hand does not incur any delay or error in detecting the partial-occlusion as the indices of erroneous LEDs are logged over each iteration. Also observe that the outcome of complete occlusion such as in point B ($t \approx 0.4$ sec) is that all the packets become erroneous. Since the system has no knowledge if the reason for such packet corruption is complete occlusion or distance variation, both these algorithms first check if the system is in a diversity (*mode* 4) and if not then the algorithm sets the system into a full-multiplexing mode (*mode* 1) by transmitting at the highest rate and then wait for the channel to get better (array being visible).

VI. CONCLUSION

We proposed a rate adaptation mechanism for a novel concept called visual MIMO that uses camera based receivers and light emitting transmitter arrays. We discussed how two important factors such as distance between the transmitter and receiver and occlusion of the receiver's view can govern the quality of the optical link and highlighted the necessity to revisit classical rate adaptation methodologies in RF channels when applied to visual MIMO. We proposed a scheme VMRA

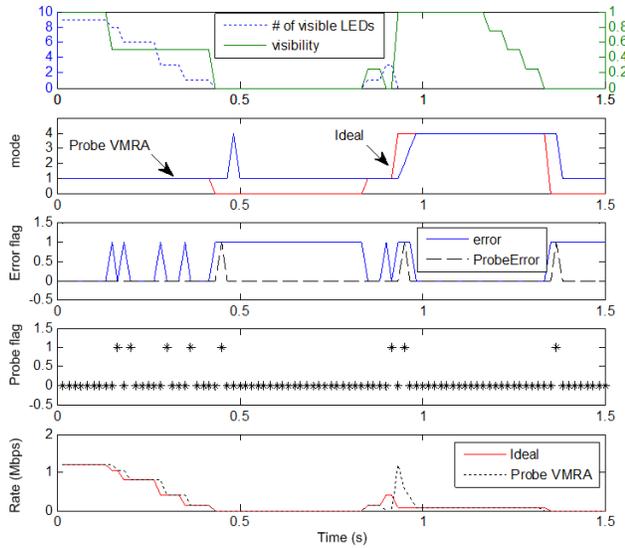


Fig. 6. Probe VMRA-distance – occlusion trace

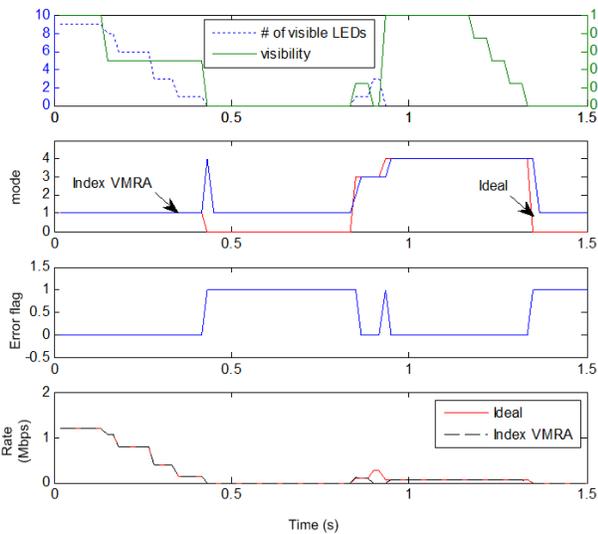


Fig. 7. Index VMRA-distance – occlusion trace

that uses packet error feedback to choose the appropriate set of LEDs both over changing distance and changing partial visibility conditions. We presented three algorithms (*Exhaustive Visibility search VMRA*, *Probe VMRA*, *Index VMRA*) applicable for our VMRA rate adaptation protocol in an exemplary visual MIMO system that uses LED array transmitter and camera receiver. The algorithms adapt to distance variations by setting a rate corresponding to the best possible spatial pattern of the elements in the transmitter array (*modes*) at that distance. The *Probe VMRA* and *Index VMRA* use special probe and a block-CRC indexing scheme respectively to detect occlusion efficiently. The *Index VMRA* approach shows the

best performance (close to ideal) among the three as the algorithm, unlike the other two, offers an error-free detection of occlusion in the channel and hence adaptation to occlusion incurs minimal overhead.

ACKNOWLEDGMENT

We thank all the anonymous reviewers for their valuable comments on this paper. This work is supported in part by the US National Science Foundation under grant CNS-1065463. T. Kwon is supported by the Korean Government MEST, Basic Research Promotion Fund (NRF-2010-013-D00063).

REFERENCES

- [1] A. Ashok, M. Gruteser, N. Mandayam, J. Silva, M. Varga, and K. Dana, "Challenge: Mobile optical networks through visual mimo," in *MobiCom '10: Proceedings of the sixteenth annual international conference on Mobile computing and networking*. New York, NY, USA: ACM, 2010, pp. 105–112.
- [2] S. Perli, N. Ahmed, and D. Katabi, "Pixnet: Lcd-camera pairs as communication links," in *SIGCOMM*, 2010, pp. 451–452.
- [3] S. H. Y. Wong, H. Yang, S. Lu, and V. Bharghavan, "Robust rate adaptation for 802.11 wireless networks," in *ACM Mobicom*, 2006, pp. 146–157.
- [4] K. Ramach, H. Kremo, M. Gruteser, and P. Spasojevi, "Scalability analysis of rate adaptation techniques in congested IEEE 802.11 networks: An orbit testbed comparative study," in *Proc. of IEEE WoWMoM*, 2007.
- [5] B. Sadeghi, V. Kanodia, A. Sabharwal, and E. Knightly, "Oar: an opportunistic auto-rate media access protocol for ad hoc networks," *Wirel. Netw.*, vol. 11, pp. 39–53, January 2005. [Online]. Available: <http://dx.doi.org/10.1007/s11276-004-4745-x>
- [6] R. T. Morris, J. C. Bicket, and J. C. Bicket, "Bit-rate selection in wireless networks," Masters thesis, MIT, Tech. Rep., 2005.
- [7] S. Nanda, K. Balachandran, and S. Kumar, "Adaptation techniques in wireless packet data services," *Communications Magazine, IEEE*, vol. 38, no. 1, pp. 54–64, Jan. 2000.
- [8] I. Pefkianakis, Y. Hu, S. H. Wong, H. Yang, and S. Lu, "Mimo rate adaptation in 802.11n wireless networks," in *Proceedings of the sixteenth annual international conference on Mobile computing and networking*, ser. *MobiCom '10*. New York, NY, USA: ACM, 2010, pp. 257–268. [Online]. Available: <http://doi.acm.org/10.1145/1859995.1860025>
- [9] Q. Xia, M. Hamdi, and K. Ben Letaief, "Open-loop link adaptation for next-generation IEEE 802.11n wireless networks," *Vehicular Technology, IEEE Transactions on*, vol. 58, no. 7, pp. 3713–3725, 2009.
- [10] L. Diana and J. Kahn, "Rate-adaptive modulation techniques for infrared wireless communications," in *Communications, 1999. ICC '99. 1999 IEEE International Conference on*, 1999.
- [11] A. García-Zambrana, C. Castillo-Vázquez, and B. Castillo-Vázquez, "Rate-adaptive fso links over atmospheric turbulence channels by jointly using repetition coding and silence periods," *Opt. Express*, vol. 18, no. 24, pp. 25 422–25 440, Nov 2010. [Online]. Available: <http://www.opticsexpress.org/abstract.cfm?URI=oe-18-24-25422>
- [12] J. Grubor, V. Jungnickel, and K.-D. Langer, "Rate-adaptive multiple sub-carrier-based transmission for broadband infrared wireless communication," in *Optical Fiber Communication Conference, 2006 and the 2006 National Fiber Optic Engineers Conference. OFC 2006*, 2006, p. 10 pp.
- [13] J. Kahn and J. Barry, "Wireless infrared communications," *Proceedings of the IEEE*, vol. 85, no. 2, pp. 265–298, Feb 1997.
- [14] B. K. P. Horn, *Robot vision*. Cambridge, MA, USA: MIT Press, 1986.
- [15] "Mipav;" <http://mipav.cit.nih.gov/documentation/HTML%20Algorithms/ FiltersSpatialGaussianBlur.html>.
- [16] "Stan moore astronomy;" http://www.stanmooreastro.com/pixel_size.html.