

Optimal Radiometric Calibration for Camera-Display Communication

Eric Wengrowski Wenjia Yuan Kristin J. Dana Ashwin Ashok Marco Gruteser

Narayan Mandayam

Department of Electrical and Computer Engineering, and
WINLAB (Wireless Information Network Laboratory)

Rutgers University

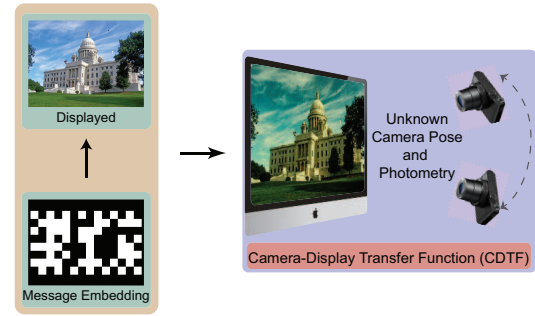
Contact Author: kristin.dana@rutgers.edu

Abstract

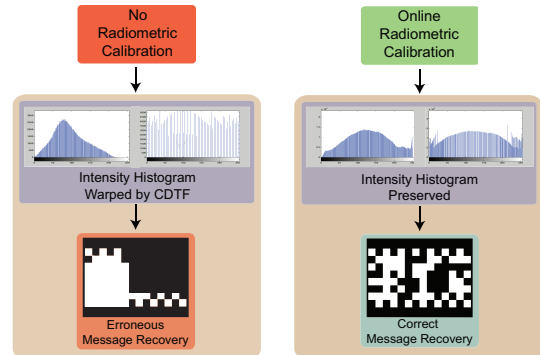
We present a novel method for communicating between a moving camera and an electronic display by embedding and recovering hidden, dynamic information within an image. A handheld camera pointed at the display can receive not only the display image, but also the underlying message. These active scenes are fundamentally different from traditional passive scenes like QR codes because image formation is based on display emittance, not surface reflectance. Detecting and decoding the message requires careful photometric modeling for computational message recovery. Unlike standard watermarking and steganography methods, our message recovery algorithm uses illumination to optically communicate hidden messages in real world scenes. The key innovation of our approach is an algorithm that performs simultaneous radiometric calibration and message recovery in one convex optimization problem. By modeling the photometry of the system using a camera-display transfer function (CDTF), we derive an optimal online radiometric calibration (OORC) for robust computational messaging as demonstrated with nine different commercial cameras and displays.

1. Introduction

While traditional computer vision concentrates on objects that reflect environment lighting (passive scenes), objects which emit light, such as electronic displays, are increasingly common in modern scenes. Unlike passive scenes, *active scenes* can have intentional information that must be detected and recovered. For example, displays with QR codes [13] can be found in numerous locations such as shop windows and billboards. However, QR-codes are very simple examples because the bold, static pattern makes detection somewhat trivial. The problem is more challeng-



(a) Hidden messages rely on small intensity variations and are corrupted by this camera-display transfer function (CDTF).



(b) Online Radiometric Calibration mitigates the distorting effects of the CDTF to enable more accurate message recovery.

Figure 1. From the display to the camera, the light signal is affected by display photometry, camera pose and camera radiometry. In each pair of intensity histograms shown above, the left represents an image histogram before passing through the CDTF, and the right represents the histogram after the CDTF.

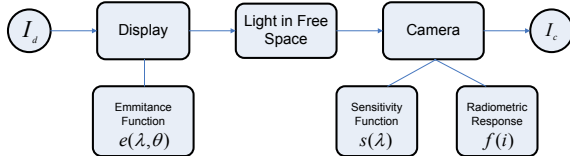


Figure 2. Image Formation Pipeline: The image I_d is displayed by an electronic display with an emittance function e . The display is observed by a camera with sensitivity s and radiometric response function f .

ing when the codes are not visible markers, but are hidden within a displayed image. The displayed image is a light field, and decoding the message is an interesting problem in photometric modeling and computational photography. The paradigm has numerous applications because the electronic display and the camera can act as a communication channel where the display pixels are transmitters and the camera pixels are receivers [2][1][31]. Unlike hidden messaging in the digital domain, prior work in real-world camera-display messaging is very limited. In this paper, we develop an optimal method for sending and retrieving hidden time-varying messages using electronic displays and cameras which accounts for the characteristics of light emittance from the display. The electronic display has two communication channels: 1) the original display image such as advertising, maps, slides, or artwork; 2) the transmission of hidden time-varying messages.

When light is emitted from a display, the resultant 3D light field has an intensity that depends on the angle of observation as well as the pixel value controlled by the display. The emittance function of the electronic display is analogous to the BRDF (bidirectional reflectance distribution function) of a surface. This function characterizes the light radiating from a display pixel. It has a particular spectral shape that does not match the spectral sensitivity curve of the camera. The effects of the display emittance function, the spectral sensitivity of the camera and the effect of camera viewing angle are all components of our photometric model for image formation as shown in Figure 2. Our approach does not require measurement or knowledge of the exact display emittance function. Instead, we estimate the entire system transfer function as a *camera-display transfer function* (CDTF) which determines the captured pixel value as a function of the displayed pixel value. By using online frame-to-frame estimation of the CDTF, no prior calibration is required and the method is independent of the particular choice of display and camera.

Interestingly, while our overall goal has very strong similarities to the field of watermarking and steganography, we present results that are novel and are aligned with the goals of computational photography. Although watermarking literature has many hidden messaging methods, this area

largely ignores the physics of illumination. Display-camera messaging is fundamentally different from watermarking because each pixel of the image is a light source that propagates in free space. Therefore, representations and methods that act only in the digital domain are not sufficient.

The problem of understanding the relationship between the displayed pixel and the captured pixel is closely related to the area of radiometric calibration [22][6][24]. In these methods, a brightness transfer function characterizes the relationship between scene radiance and image pixel values. The characterization of this function is done by measuring a range of scene radiances and the corresponding captured image pixels. Our problem in camera-display messaging is similar but has important key differences. The CDTF is more complex than standard radiometric calibration because the system consists of both a display and a camera, each device adding its own nonlinearities. We can exploit the control of pixel intensities on the display and easily capture the full range of input intensities. However, the display emittance function is typically dependent on the display viewing angle. Therefore, the CDTF is dependent on camera pose. In a moving camera system, the CDTF must be estimated per frame; that is, an online CDTF estimation is needed. Furthermore, this function varies spatially over the electronic display surface.

We show that the two-part problem of online radiometric calibration and accurate message retrieval can be structured as an optimization problem. We present an elegant problem formulation where the photometric modeling leads to *physically-motivated kernel functions* that are used with a support vector machine classifier. We show that calibration and message bit classification can be done simultaneously and the resulting optimization algorithm operates in four dimensional space and is convex. The algorithm is a novel method for *online optimal radiometric calibration* (OORC) that enables accurate camera-display messaging. An example message recovery result is shown in Figure 3. Our experimental results show that accuracy levels for message recovery can improve from as low as 40-60% to higher than 90% using our approach when compared to either no calibration, or sequential calibration followed message recovery. For evaluation of results, 9 different combinations of displays and cameras are used with 15 different image sequences, for multiple embedded intensity values, and multiple camera-display view angles.

The contributions of the paper can be summarized as follows: 1) A new optimal online radiometric calibration with simultaneous message recovery, cast as a convex optimization problem; 2) photometric model of the camera display transfer function; 3) the use of ratex patches to provide continual calibration information as a practical method for online calibration; 4) the use of distribution-driven intensity mapping as a practical method for visually non-disruptive

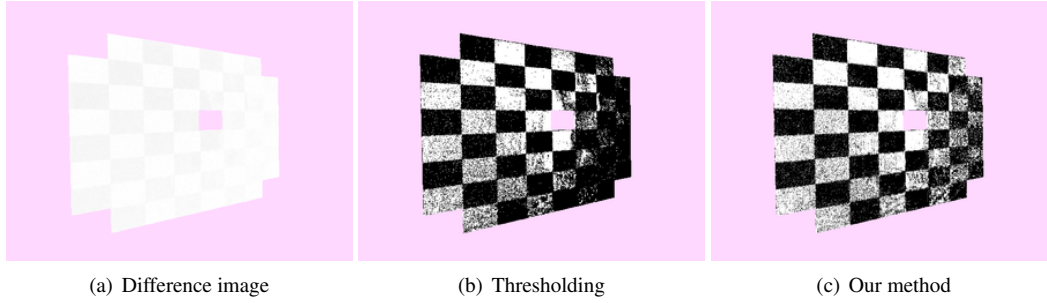


Figure 3. Comparison of message recovery with a naive method and the proposed optimal method (a) Difference of two consecutive frames in the captured sequence to reveal the transmitted message. (b) Naive method: Threshold the difference image by a constant (threshold $T = 5$ for this example). (c) Optimal Method: Bits are classified by a simultaneous radiometric calibration and message recovery.

online calibration.

2. Related Work

Watermarking In developing a system where cameras and displays can communicate under real world conditions, the initial expectation was that existing watermarking techniques could be used directly. Certainly the work in this field is extensive and has a long history with numerous surveys compiled [4] [35] [28] [5] [14] [27]. Surprisingly, existing methods are not directly applicable to our problem. In the field of watermarking, a fixed image or mark is embedded in an image often with the goal of identifying fraudulent copies of a video, image or document. Existing work emphasizes almost exclusively the digital domain and does not account for the effect of illumination in the image formation process in real world scenes. In the digital domain, neglecting the physics of illumination is quite reasonable; however, for camera-display messaging, illumination plays a central role.

From a computer vision point of view, the imaging process can be divided into two main components: photometry and geometry. The geometric aspects of image formation have been addressed to some extent in the watermarking community, and many techniques have been developed for robustness to geometric changes during the imaging process such as scaling, rotations, translations and general homography transformations [7] [29] [8] [34] [19] [28] [30]. However, the *photometry* of imaging has largely been ignored. The rare mention of photometric effects [40] [37] in the watermarking literature doesn't define photometry with respect to illumination; instead photometric effects are defined as "lossy compression, denoising, noise addition and lowpass filtering". In fact, photometric attacks are sometimes defined as jpeg compression [8].

Radiometric Calibration Ideally, we consider the pixel-values in a camera image to be a measurement of light incident on the image plane sensor. It is well known that the

relationship is typically nonlinear. Radiometric calibration methods have been developed to estimate the camera response function that converts irradiance to pixel values. In measuring a camera response, a series of known brightness values are measured along with the corresponding pixel values. In general, having such ground truth brightness is quite difficult. The classic method [6] uses multiple exposure values instead. The light intensity on the sensor is a linear function of the time of exposure, so known exposure times enables ground truth light intensity. This exposure-based method is used in several radiometric calibration methods [22] [24] [6] [21] [17]. Our goal for the display-camera system is related to radiometric calibration; the system converts scene radiance to pixels (the camera), but also converts from pixel to scene radiance (the display) so that the whole camera-display system is a function that maps a color value at the display to a color value at the camera.

The camera response in radiometric calibration is either estimated as a full mapping where i_{out} is specified for every i_{in} or as an analytic function $g(i_{in})$. Several authors [22] [3] [18] use polynomials to model the radiometric response function. Similarly, we have found that fourth order polynomials can be used for modeling the inverse display-camera transfer function. The dependence on color is typically modeled by considering each channel independently [22] [24] [6] [9]. Interestingly, although more complex color models have been developed [16] [20] [36], we have found the independent channel approach suitable for the display-camera representation where the optimality criterion is accurate message recovery.

Existing radiometric calibration methods are developed for cameras, not camera-display systems. Therefore, display emittance function is not part of these prior methods. However, for the camera-display transfer function, this component plays an important role. We do not use the measured display emittance function explicitly, but since the CDTF is view dependent and the camera can move, our approach is to perform radiometric calibration per frame.

Other Methods for Camera-Display Communication

Camera-display communications have precedent in the computer vision community, but existing methods differ from our proposed approach. For example, researchers on the Bocode project [23] presented a system using an invisible message, however the message is a fixed symbol, not a time-varying message. Invisible QR codes were addressed in [15], but these QR-codes are fixed. Similarly, traditional watermark approaches typically contained fixed messages. LCD-camera communications is presented in [25] with a time-varying message, but the camera is in a fixed position with respect to the display. Consequently, the electronic display is not detected, tracked or segmented from the background. Furthermore, the transmitted signal is not hidden in this work. Recent work has been done in high speed visible light communications [32], but this work does not utilize existing displays and cameras and requires specialized hardware and LED devices. Time-of-flight cameras have recently been used for phase-based communication [39], but these methods require special hardware. Interest in camera-display messaging is also shared in the mobile communications domain. COBRA, RDCode, and Strata have developed 2D barcode schemes designed to address the challenges of low-resolution and slow shutter speeds typically present in smartphone cameras [10] [33] [12]. Likewise, Lightsync has targeted synchronization challenges with low frequency cameras. [11].

3. System Properties

In our proposed camera-display communication system, pixel values from the display are inputs, while captured intensities from the camera are output. We denote the mapping from displayed intensities to captured ones as *Camera-Display Transfer Function* (CDTF). In this section, we motivate the need for online radiometric calibration by briefly analyzing factors that influence the CDTF.

Display Emittance Variation Displays vary widely in brightness, hue, white balance, contrast and many other parameters that will influence the appearance of light. To affirm this hypothesis, an SLR camera with fixed parameters observes 3 displays and models the CDTF for each one as shown in Figure 4(a) (Samsung), 4(b) (LG), and 4(c) (iMac). Although each display is tuned to the same parameters, including contrast and RGB values, each display produces a unique CDTF.

Observation Angles Electronic displays emit light with an angular dependence. Consider the image of an electronic display captured by a camera from multiple angles as shown in Figure 5. More oblique observation angles yield lower captured pixel intensities. Additionally, there is a nonlinear

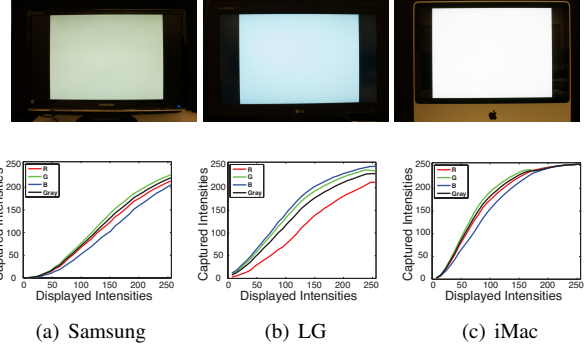


Figure 4. **Variance of Light Output among Displays.** An SLR camera captured a range of grayscale [0,255] intensity values produced by 3 different LCDs. These 3 CDTF curves highlight the difference in the light emittance function for different displays.

relationship between captured light intensity and viewing angle.

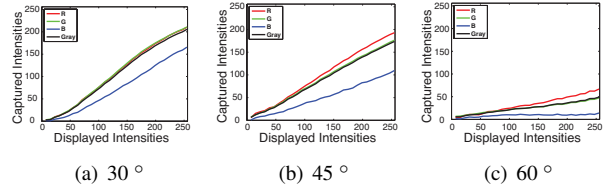


Figure 5. **Influence of observation angles.** Using the Nikon-Samsung pair, a range of grayscale [0, 255] values were displayed and captured from a set of different observation angles. As observation angle became more oblique, the camera-display transfer function changes.

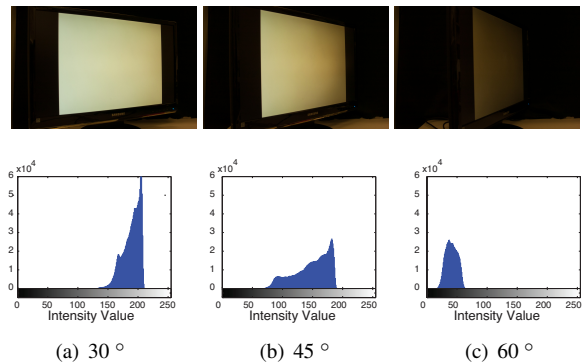


Figure 6. **Histograms of intensities captured from a uniform display.** Notice as observation angle changes, so does the distribution of captured intensities illustrating the angular variation of the display emittance function.

4. Methods

4.1. Photometry of Display-Camera systems

The captured image \mathbf{I}_c from the camera viewing the electronic display image \mathbf{I}_d can be modeled using the image formation pipeline in Figure 2. First, consider a particular pixel within the display image \mathbf{I}_d with red, blue and green components given by $\rho = (\rho_r, \rho_g, \rho_b)$. The captured image \mathbf{I}_c at the camera has three color components (I_c^r, I_c^g, I_c^b) , however there is no one-to-one correspondence between the color channels of the camera sensitivity function and the electronic display emittance function. When the monitor displays the value (ρ_r, ρ_g, ρ_b) at a pixel, it is emitting light in a manner governed by its emittance function and modulated by ρ . The monitor emittance function e is typically a function of the viewing angle $\theta = (\theta_v, \phi_v)$ comprised of a polar and azimuthal component. For example, the emittance function of an LCD monitor has a large decrease in intensity with polar angle (see Figure 6).

The emittance function has three components, i.e. $e = (e_r, e_g, e_b)$. Therefore the emitted light I as a function of wavelength λ for a given pixel (x, y) on the electronic display is given by

$$I(x, y, \lambda) = \rho_r e_r(\lambda, \theta) + \rho_g e_g(\lambda, \theta) + \rho_b e_b(\lambda, \theta), \quad (1)$$

or

$$I(x, y, \lambda) = \rho \cdot e(\lambda, \theta). \quad (2)$$

Now consider the intensity of the light received by one pixel element at the camera sensor. Let $s_r(\lambda)$ denote the camera sensitivity function for the red component, then the red pixel value I_c^r can be expressed as

$$I_c^r \propto \int_{\lambda} [\rho \cdot e(\lambda, \theta)] s_r(\lambda) d\lambda. \quad (3)$$

Notice that the sensitivity function of the camera has a dependence on wavelength that is likely different than the emittance function of the monitor. That is, the interpretation of “red” in the monitor is different from that of the camera. Notice that a sign of proportionality is used in Equation 3 to specify that the pixel value is a linear function of the intensity at the sensor, assuming a linear camera and display. This assumption will be removed in Section 4.3.

Equation 3 can be written to consider all color components in the captured image \mathbf{I}_c as

$$\mathbf{I}_c \propto \int_{\lambda} [\rho \cdot e(\lambda, \theta)] \mathbf{s}(\lambda) d\lambda. \quad (4)$$

where $\mathbf{s} = (s_r, s_g, s_b)$.

4.2. Message Structure

The pixel value ρ is controllable by the electronic display driver, and so it provides a mechanism for embedding

information. We use two sequential frames in our approach. We modify the monitor intensity by adding the value κ and transmit two consecutive images, one with the added value \mathbf{I}_e and one image of original intensity \mathbf{I}_o . The recovered message depends on the display emittance function and camera sensitivity function if the embedded message is done by adding κ as follows:

$$\mathbf{I}_e \propto \int_{\lambda} [(\kappa + \rho) \cdot e(\lambda, \theta)] \mathbf{s}(\lambda) d\lambda. \quad (5)$$

Recovery of the embedded signal leads to a difference equation

$$\mathbf{I}_e - \mathbf{I}_o \propto \int_{\lambda} [\kappa \cdot e(\lambda, \theta)] \mathbf{s}(\lambda) d\lambda. \quad (6)$$

The dependence on the properties of the display e and the spectral sensitivity of the camera s remains. We use additive-based messaging, instead of ratio-based methods, because this structure is convenient for convexity of the algorithm as described in Section 4.3.

The main concept for message embedding is illustrated in Figure 7. In order to convey many “bits” per image, we divide the image region into a series of block components. Each block can convey a bit “1” or “0”. The blocks corresponding to a “1” contain the added value κ typically set to 3 or 5 gray levels on the [0,255] scale, while the zero blocks have no additive component ($\kappa = 0$). The message is recovered by sending the original frame followed by a frame with the embedded message and using the difference for message recovery. The message can also be added to the coarser scales of a image pyramid decomposition [26], in order to better hide the message within the display image content. The display can be tracked with existing methods [38]. This message structure is decidedly very simple, so the methods presented here can be applied to many message coding schemes.

When accounting for the nonlinearity in the camera and display, we rewrite Equation 4 to include the radiometric response function f ,

$$\mathbf{I}_c = f \left(\int_{\lambda} [\rho \cdot e(\lambda, \theta)] \mathbf{s}(\lambda) d\lambda \right). \quad (7)$$

More concisely,

$$\mathbf{I}_c = f(\mathbf{I}_d), \quad (8)$$

and the recovered display intensity is

$$\mathbf{I}_d = f^{-1}(\mathbf{I}_c) = g(\mathbf{I}_c). \quad (9)$$

We use polynomials to represent the radiometric inverse function $g(i)$. The same inverse function g is used for all color channels. This simplification of the color problem is justified by the accuracy of the empirical results.

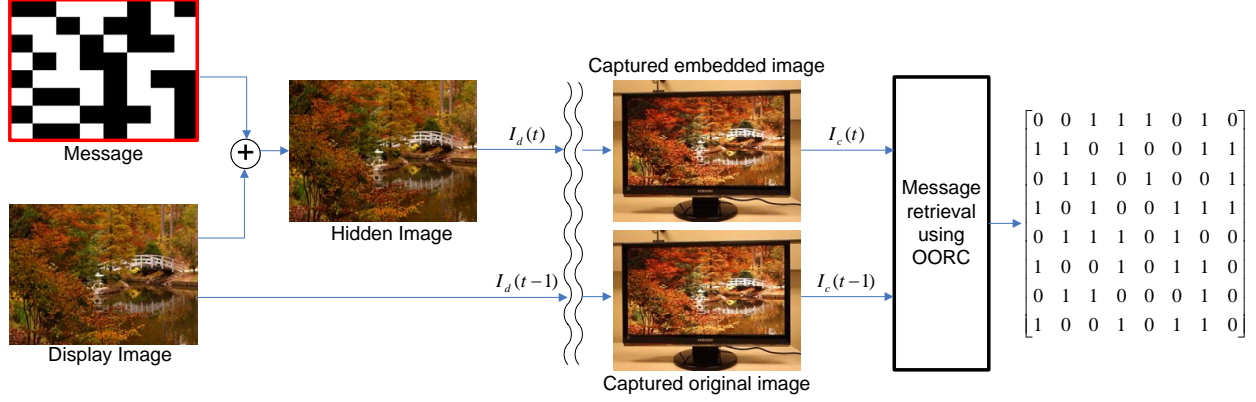


Figure 7. Message Embedding and Retrieval. Two sequential frames are sent, an original frame and a frame with an embedded message image. Simple differencing is not sufficient for message retrieval. Our method (OORC) is used to recover messages accurately.

4.3. Optimal Online Radiometric Calibration

The two goals of message recovery and calibration can be combined to a single problem. While ideal radiometric calibration would provide a captured image that is a linear function of the displayed image, we show that calibrating followed by message recovery only gives a relatively small increase in message accuracy. However, if the two goals are combined into a simultaneous problem we have two benefits: 1) the problem formulation can be done in a convex optimization paradigm with a single global solution and 2) the accuracy increases significantly.

Let $g(i)$ be the inverse function that is modeled with a fourth order polynomial as follows

$$g(i) = a_4 i^4 + a_3 i^3 + a_2 i^2 + a_1 i + a_0. \quad (10)$$

Consider two images frames i_o , where i_o is the original frame and i_e the image frame with the embedded message. Note the use of i_o instead of \mathbf{I}_o for notational compactness. Since we are using an additive message embedding, we wish to classify the message bits as either ones or zeros based on the difference image $i_o - i_e$.

Taking into account the radiometric calibration, we want to classify on the recovered data $g(i_o) - g(i_e)$. We have found empirically that the inverse function can be modeled by a fourth order polynomial, so that the function to be classified is

$$g(i_o) - g(i_e) = a_4(i_o^4 - i_e^4) + a_3(i_o^3 - i_e^3) + a_2(i_o^2 - i_e^2) + a_1(i_o - i_e). \quad (11)$$

In Equation 11, we see that the calibration problem has a physically motivated nonlinear mapping function. That is, we see that the original data (i_o, i_e) can be placed into a higher dimensional space using the nonlinear mapping function Φ which maps from a two dimensional space to

a four dimensional space as follows

$$\Phi(i_o, i_e) = \begin{bmatrix} (i_o^4 - i_e^4) & (i_o^3 - i_e^3) & (i_o^2 - i_e^2) & (i_o - i_e) \end{bmatrix}. \quad (12)$$

In this four dimensional space we seek a separating hyperplane between the two classes (one-bits and zero-bits). Our experimental results indicate that these are not separable in lower dimensional space, but the movement to a higher dimensional space enables the separation. Also, the form of that higher dimensional space is physically motivated by the need for radiometric calibration. Therefore our problem becomes a support vector machine classifier where the optimal support vector weights and the calibration parameters are *simultaneously* estimated. That is, we estimate

$$w^T u + b, \quad (13)$$

where, $w \in \mathbf{R}^4$, b are the separating hyperplane parameters, and u is the input feature vector. Since we want to perform radiometric calibration, the four-dimensional input is given

$$u = \begin{bmatrix} a_4(i_o^4 - i_e^4) & a_3(i_o^3 - i_e^3) & a_2(i_o^2 - i_e^2) & a(i_o - i_e) \end{bmatrix}^T. \quad (14)$$

Notice that the $w^T u + b$ is still linear in the coefficients of the inverse radiometric function. These coefficients and the scale factor w are estimated simultaneously. We arrive at the important observation that accounting for the CDTF preserves the convexity of the overall optimization problem. The coefficients of the function g are scaled by w , so that calibration and classification can be done *simultaneously*, and convexity of the SVM is preserved. We refer to this method as *optimal online radiometric calibration (OORC)* because it recovers radiometric parameters via convex optimization for each frame.

Ratex Patches The standard problem of radiometric calibration is solved by varying exposure so that a range of

scene radiance can be measured. For CDTF calibration in a single frame, patches are placed within the display image that have intensity variation over the full range of display brightness values (a linear variation with pixel values from 0 to 255). These radiometric textured calibration patches or *ratex patches* are placed in inconspicuous regions of the display image such as an image corner. The ratex patches are not used as part of the hidden message, but instead provide training data in each frame for the OORC method of CDTF calibration and message recovery. Consecutive frames of ratex patches toggle between message bit “1” and message bit “0” to provide training data for both message bits.

4.4. Hidden Ratex

We also introduce a method for radiometric calibration that employs visually non-disruptive hidden ratex mapping, since ratex patches can be visually obtrusive and unattractive for applications. Rather than directly measuring the effect that the CDTF has on known intensity values, we measure the effect on the image histogram. Instead of using ratex patches that have a linear variation over the full intensity range, we use display images with intensity values that are well-distributed over the full intensity range. We estimate the CDTF by finding the mapping of the measured histogram to the original histogram. For this approach to work, we need to know the initial intensity histogram of an image before it passes through the CDTF. We perform as simple intensity mapping (equalization) on every image before it is displayed, so the initial intensity histogram is known and uniformly distributed. The camera-captured image is intensity mapped to restore this distribution, after distortion by the CDTF. The inverse CDTF is computed and corrected for in this approach and we refer to this method as *hidden ratex* since no visible patches are used. The hidden ratex method is illustrated in Figure 1.

5. Results

For empirical validation, 9 different combination of displays and cameras are used, comprised of 3 displays: 1) LG M3204CCBA 32 inch, 2) Samsung SyncMaster 2494SW, 3) iMac (21.5 inch 2009); and 3 cameras: 1) Canon EOS Rebel XSi, 2) Nikon D70, 3) Sony DSC-RX100. Fifteen 8-bit display images are used. From each display image, we create a display video of 10 frames: 5 frames with the original display images interleaved with 5 images of embedded time-varying messages. An embedded message frame is followed by an original image frame to provide the temporal image pair i_e and i_o . The display image does not change in the video, only the bits of the message frames. Each message frame has $8 \times 8 = 64$ blocks used for message bits (with 5 bits used for ratex patches for calibration and classification training data). Considering 5 display images, with 5 message frames and 59 bits per frame results

in approximately 1500 message bits. The accuracy for each video is defined as the number of correctly classified bits divided by the total bits embedded and is averaged over all testing videos. The entire test set over all display-camera combinations is approximately 18,000 test bits.

We evaluate 4 methods for embedded message recovery. Method 1 (*Naive Threshold*) has no radiometric calibration, only the difference $i_e - i_o$ is used to recover the message bit via thresholding. Method 2 (*Two-step*) is radiometric calibration using ratex patches followed by thresholding the interframe difference $i_e - i_o$ for message recovery. Method 3 (*OORC*) is the optimal calibration where both radiometric calibration and message recovery are done simultaneously. Method 4 (*Hidden Ratex*) is calibration using hidden ratex intensity mapping followed by simple differencing for message recovery. The methods we introduced here (Methods 2-4) demonstrate significant improvement over naive thresholding. For methods 2 and 3, training data from pixels in the ratex patches are used to train an SVM classifier. For method 4, no visible patches are needed. For each of the 9 display-camera combinations, the accuracy of the 4 message recovery methods was tested with 2 sets of experimental variables: 1) 0° frontal camera-display view; 2) 45° oblique camera-display view; and: 1) embedded message intensity difference of 5; 2) embedded message intensity difference of 3. The results of these tests can be found in Tables 1, 2, 3, and 4. Notice that naive thresholding has low message recover rates (as low as 47.5% for oblique views). Message recovery rates were highest for the OORC method with recovery rates of 98-99% for most camera display combinations even for oblique views. The hidden ratex method also maintained near 90% recognition rates for oblique views and had the advantage of having no visible calibration patches.

6. Discussion and Conclusion

The results indicate a marked improvement in message recovery over naive thresholding for camera-display messaging with our methods. We demonstrate experimental results for nine different camera-display combinations at frontal and oblique viewing directions. We show that naive thresholding, while intuitively simple, is a poor choice because the variation of display intensity with camera pose is ignored. These methods lead to lower message recovery rates, especially for oblique views (45°) and small intensity messages. Prior methods of digital watermarking do not take into account the photometric effects of the camera-display transfer function and the resulting dependence on camera pose. Therefore these prior methods are likewise prone to error. Our experimental results show that hidden, dynamic messages can be embedded in a display image and recovered robustly.

Accuracy (%)	Naive Threshold	Two-step	OORC	Hidden Ratex
Canon-iMac	72.94	75.67	99.17	89.63
Canon-LG	58.94	84.94	98.44	95.74
Canon-Samsung	48.44	64.89	99.39	89.91
Nikon-iMac	60.17	75.50	95.17	90.00
Nikon-LG	49.72	73.39	99.33	94.81
Nikon-Samsung	47.22	72.89	95.00	89.54
Sony-iMac	64.44	76.00	99.06	71.11
Sony-LG	56.11	75.61	98.56	90.93
Sony-Samsung	47.50	79.11	98.89	87.80
Average	56.17	75.33	98.11	88.83

Table 1. Accuracy of embedded message recovery and labeling with additive difference +3 on [0,255] and captured with 45° oblique view.

Accuracy (%)	Naive Threshold	Two-step	OORC	Hidden Ratex
Canon-iMac	85.56	83.06	96.44	91.57
Canon-LG	86.39	90.94	98.67	94.07
Canon-Samsung	87.94	87.78	98.94	91.30
Nikon-iMac	84.06	84.00	96.50	90.27
Nikon-LG	74.67	81.44	99.94	90.09
Nikon-Samsung	77.33	86.06	98.00	91.57
Sony-iMac	89.33	84.22	99.44	70.00
Sony-LG	87.61	95.39	99.72	80.74
Sony-Samsung	80.00	83.78	96.26	84.54
Average	83.56	86.30	98.22	87.13

Table 2. Accuracy of embedded message recovery and labeling with additive difference +3 on [0,255] and captured at 0° frontal view.

Accuracy (%)	Naive Threshold	Two-step	OORC	Hidden Ratex
Canon-iMac	97.06	94.50	99.83	95.37
Canon-LG	87.89	99.00	99.39	99.44
Canon-Samsung	71.67	88.11	100.00	95.37
Nikon-iMac	91.89	93.67	96.00	96.11
Nikon-LG	81.56	95.11	99.94	98.88
Nikon-Samsung	58.78	92.22	99.39	97.41
Sony-iMac	92.28	92.00	99.72	80.37
Sony-LG	77.06	96.22	100.00	91.13
Sony-Samsung	63.28	94.17	99.89	81.67
Average	80.16	93.89	99.35	93.71

Table 3. Accuracy of embedded message recovery and labeling with additive difference +5 on [0,255] and captured with 45° oblique perspective.

Accuracy (%)	Naive Threshold	Two-step	OORC	Hidden Ratex
Canon-iMac	95.28	96.61	99.00	95.74
Canon-LG	97.11	99.72	97.17	97.59
Canon-Samsung	97.39	97.33	98.94	94.35
Nikon-iMac	98.39	99.17	99.22	96.11
Nikon-LG	99.83	100.00	99.83	97.31
Nikon-Samsung	96.33	97.44	98.56	95.74
Sony-iMac	97.72	97.00	99.94	81.67
Sony-LG	99.39	100.00	100.00	90.74
Sony-Samsung	92.50	92.33	98.06	90.28
Average	97.10	97.73	98.97	93.28

Table 4. Accuracy of embedded message recovery and labeling with additive difference +5 on [0,255] and captured at 0° frontal view.

References

- [1] A. Ashok, M. Gruteser, N. Mandayam, and K. Dana. Characterizing multiplexing and diversity in visual mimo. *Information Sciences and Systems (CISS)*, pages pp. 1–6, March 2011. 2
- [2] A. Ashok, M. Gruteser, N. Mandayam, J. Silva, M. Varga, and K. Dana. Challenge: mobile optical networks through visual mimo. *MobiCom: Proceedings of the sixteenth annual international conference on Mobile computing and networking*, pages pp.105–112, 2010. 2

- [3] A. Chakrabarti, D. Scharstein, and T. Zickler. An empirical camera model for internet color vision. *British Machine Vision Conference*, 2009. 3
- [4] A. Cheddad, J. Condell, K. Curran, and P. M. Kevitt. Digital image steganography: Survey and analysis of current methods. *Signal Processing*, 90(3):727 – 752, 2010. 3
- [5] I. Cox, M. Miller, J. Bloom, and C. Honsinger. Digital watermarking. *Journal of Electronic Imaging*, 2002. 3
- [6] P. Debevec and J. Malik. Recovering high dynamic range radiance maps from photographs. *ACM SIGGRAPH*, pages pp. 369–378, 1997. 2, 3
- [7] P. Dong, J. G. Brankov, N. P. Galatsanos, Y. Yang, and F. Divoine. Digital watermarking robust to geometric distortions. *Image Processing, IEEE Transactions on*, 14(12):2140 –2150, 2005. 3
- [8] J. L. Dugelay, S. Roche, C. Rey, and G. Doerr. Still-image watermarking robust to local geometric distortions. *Image Processing, IEEE Transactions on*, 15(9):2831 –2842, 2006. 3
- [9] M. D. Grossberg and S. K. Nayar. What can be known about the radiometric response from images? In *Proceedings of the 7th European Conference on Computer Vision-Part IV, ECCV '02*, pages 189–205, London, UK, UK, 2002. Springer-Verlag. 3
- [10] T. Hao, R. Zhou, and G. Xing. Cobra: color barcode streaming for smartphone systems. In *Proceedings of the 10th international conference on Mobile systems, applications, and services*, pages 85–98. ACM, 2012. 4
- [11] W. Hu, H. Gu, and Q. Pu. Lightsync: Unsynchronized visual communication over screen-camera links. In *Proceedings of the 19th annual international conference on Mobile computing & networking*, pages 15–26. ACM, 2013. 4
- [12] W. Hu, J. Mao, Z. Huang, Y. Xue, J. She, K. Bian, and G. Shen. Strata: layered coding for scalable visual communication. In *Proceedings of the 20th annual international conference on Mobile computing and networking*, pages 79–90. ACM, 2014. 4
- [13] ISO/IEC18004. *Inf. technology, Automatic identification and data capture techniques, Bar code symbology, QR Code*, 2000. 1
- [14] N. Johnson, Z. Duric, and S. Jajodia. *Information Hiding: Steganography and Watermarking - Attacks and Countermeasures*. Springer, 1 edition, 2000. 3
- [15] K. Kamijo, N. Kamijo, and G. Zhang. Invisible barcode with optimized error correction. In *Image Processing, 2008. ICIIP 2008. 15th IEEE International Conference on*, pages 2036 –2039, oct. 2008. 4
- [16] S. J. Kim, H. T. Lin, Z. Lu, S. Susstrunk, S. Lin, and M. S. Brown. A new in-camera imaging model for color computer vision and its application. In *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 2012. 3
- [17] S. J. Kim and M. Pollefeys. Robust radiometric calibration and vignetting correction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(4):562 –576, april 2008. 3
- [18] J.-Y. Lee, Y. Matsushita, B. Shi, I. S. Kweon, and K. Ikeuchi. Radiometric calibration by rank minimization. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 35(1):144 –156, jan. 2013. 3
- [19] C. Y. Lin, M. Wu, J. A. Bloom, I. J. Cox, M. L. Miller, and Y. M. Lui. Rotation, scale, and translation resilient watermarking for images. *Image Processing, IEEE Transactions on*, 10(5):767 –782, 2001. 3
- [20] H. T. Lin, S. J. Kim, S. Susstrunk, and M. S. Brown. Revisiting radiometric calibration for color computer vision. *ICCV*, 2011. 3
- [21] S. Mann and R. Picard. On being undigital with digital cameras: Extending dynamic range by combining differently exposed pictures. *Proc. IST 46th annual conference*, pages 422 – 428, 1995. 3
- [22] T. Mitsunaga and S. Nayar. Radiometric self calibration. In *Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on.*, volume 1, pages 2 vol. (xxiii+637+663), 1999. 2, 3
- [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*. ACM, 2009. 4
- [24] S. K. Nayar and T. Mitsunaga. High dynamic range imaging: spatially varying pixel exposures. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages pp. 472–479, 2000. 2, 3
- [25] S. D. Perli, N. Ahmed, and D. Katabi. Pixnet: designing interference-free wireless links using lcd-camera pairs. *ACM Int. Conf. on Mobile Computing and Networking*, 2010. 4
- [26] J. B. Peter and E. H. Adelson. The laplacian pyramid as a compact image code. *IEEE Transactions on Communications*, 31:532–540, 1983. 5
- [27] F. A. P. Petitcolas, R. J. Anderson, and M. G. Kuhn. Information hiding-a survey. *Proceedings of the IEEE*, 87(7):1062 –1078, jul 1999. 3
- [28] V. M. Potdar, S. Han, and E. Chang. A survey of digital image watermarking techniques. *IEEE International Conference on Industrial Informatics*, pages 709–716, 2005. 3

- [29] A. Sangeetha, B. Gomathy, and K. Anusudha. A watermarking approach to combat geometric attacks. In *Digital Image Processing, 2009 International Conference on*, pages 381–385, 2009. 3
- [30] J. S. .Seo and C. D. Yoo. Image watermarking based on invariant regions of scale-space representation. *Signal Processing, IEEE Transactions on*, 54(4):1537–1549, 2006. 3
- [31] M. Varga, A. Ashok, M. Gruteser, N. Mandayam, W. Yuan, and K. Dana. Demo: Visual mimo-based led-camera communication applied to automobile safety. *Proceedings of ACM/USENIX International Conference on Mobile Systems, Applications, and Services (MobiSys)*, pages pp. 383–384, 2011. 2
- [32] J. Vucic, C. Kottke, S. Nerreter, K. Langer, and J. Walewski. 513 mbit/s visible light communications link based on dmt-modulation of a white led. *Journal of Lightwave Technology*, 28(24):3512–3518, 2010. 4
- [33] A. Wang, S. Ma, C. Hu, J. Huai, C. Peng, and G. Shen. Enhancing reliability to boost the throughput over screen-camera links. In *Proceedings of the 20th annual international conference on Mobile computing and networking*, pages 41–52. ACM, 2014. 4
- [34] X. Wang, L. Hou, and J. Wu. A feature-based robust digital image watermarking against geometric attacks. *Image Vision Comput.*, 26(7), 2008. 3
- [35] P. Wayner. *Disappearing Cryptography: Information Hiding: Steganography & Watermarking*. Morgan Kaufmann Publishers Inc., 3 edition, 2009. 3
- [36] Y. Xiong, K. Saenko, T. Darrell, and T. Zickler. From pixels to physics: Probabilistic color de-rendering. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages pp. 358–365, 2012. 3
- [37] L. Yang and Z. Guo. A robust video watermarking scheme resilient to spatial desynchronization and photometric distortion. In *Signal Processing, 2006 8th International Conference on*, volume 4, 16-20 2006. 3
- [38] W. Yuan, K. Dana, A. Ashok, M. Varga, M. Gruteser, and N. Mandayam. Photographic steganography for visual mimo: A computer vision approach. *IEEE Workshop on the Applications of Computer Vision (WACV)*, pages 345–352, 2012. 5
- [39] W. Yuan, R. E. Howard, K. J. Dana, R. Raskar, A. Ashok, M. Gruteser, and N. Mandayam. Phase messaging method for time-of-flight cameras. In *Computational Photography (ICCP), 2014 IEEE International Conference on*, pages 1–8. IEEE, 2014. 4
- [40] F. Zou, H. Ling, X. Li, Z. Xu, and P. Li. Robust image copy detection using local invariant feature. In *Multimedia Information Networking and Security, 2009. MINES '09. International Conference on*, volume 1, pages 57–61, 2009. 3