Bokode: Imperceptible Visual Tags for Camera Based Interaction from a Distance

Conference Paper · July 2009

DOI: 10.1145/1576246.1531404

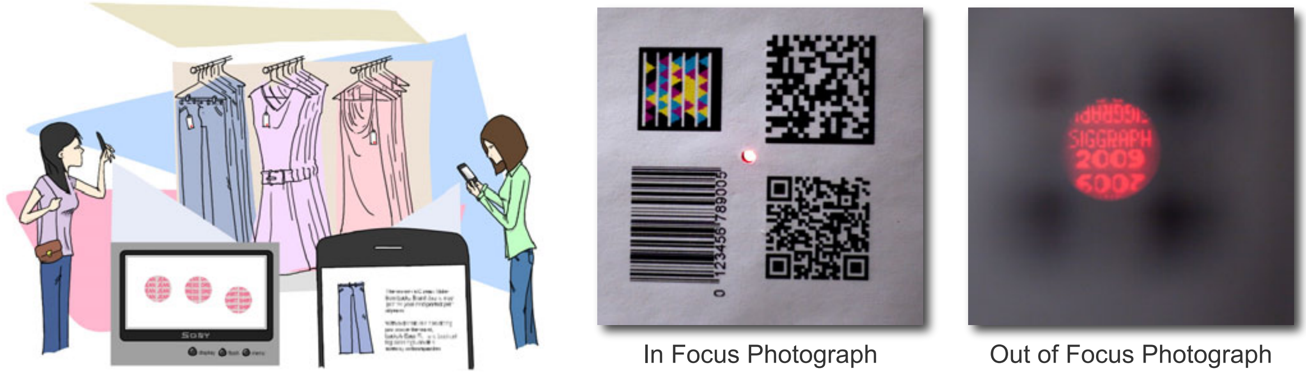
Grace Rusi Woo
Massachusetts Institute of Technology
12 PUBLICATIONS **545** CITATIONS

Shinsaku Hiura
University of Hyogo
112 PUBLICATIONS **806** CITATIONS

Ramesh Raskar
Massachusetts Institute of Technology
552 PUBLICATIONS **20,055** CITATIONS

Bokode: Imperceptible Visual tags for Camera Based Interaction from a Distance

Ankit Mohan Grace Woo* Shinsaku Hiura† Quinn Smithwick Ramesh Raskar
Camera Culture Group, MIT Media Lab
http://cameraculture.media.mit.edu/bokode



In Focus Photograph Out of Focus Photograph

Figure 1: *How can we enable a commodity camera to photograph and capture a tiny barcode (a few millimeters in diameter) from a large distance (several meters away)? Our setup is built on a novel optical barcode design called the Bokode, and a standard camera. The Bokode occupies very few pixels and appears as a dot in an in focus photo captured by a standard camera (center). The barcode information is revealed in an out of focus photo captured by the same camera, occupies several hundred pixels, and is easily decodable. In addition to identity, we also obtain the distance and angle of the camera relative to the Bokode.*

Abstract

We show a new camera based interaction solution where an ordinary camera can detect small optical tags from a relatively large distance. Current optical tags, such as barcodes, must be read within a short range and the codes occupy valuable physical space on products. We present a new low-cost optical design so that the tags can be shrunk to $3mm$ visible diameter, and unmodified ordinary cameras several meters away can be set up to decode the identity plus the relative distance and angle. The design exploits the bokeh effect of ordinary cameras lenses, which maps rays exiting from an out of focus scene point into a disk like blur on the camera sensor. This bokeh-code or *Bokode* is a barcode design with a simple lenslet over the pattern. We show that an off-the-shelf camera can capture Bokode features of $2.5\mu m$ from a distance of over 4 meters. We use intelligent binary coding to estimate the relative distance and angle to the camera, and show potential for applications in augmented reality and motion capture. We analyze the constraints and performance of the optical system, and discuss several plausible application scenarios.

Keywords: computational probes, defocus blur, human-computer interaction, 2D fiducials, augmented reality, motion capture

1. Introduction

Digital cameras, from professional SLRs to cellphone cameras, are ubiquitous and are used not only for photography but also for accessing information by recognizing and understanding scene elements. By adding self-identifying tags, such as barcodes, the recognition problem is greatly simplified. However, typical barcodes are large in size relative to the amount of information they carry, and must be carefully designed for readability by a simple device that usually works only in close proximity to the barcode. Two dimensional fiducial markers used in Augmented Reality (AR), robot navigation, photogrammetry and other applications to compute relative camera pose also suffer from the same problem.

Smart environments of the future require efficient machine-to-machine interaction with aesthetically pleasing visual tagging that is compatible with digital cameras. Original barcodes were intended for machines to recognize tagged objects. Unfortunately, the barcodes are also visible to a human eye, and they take up precious visual real estate. Our motivation is to improve this interaction without further cluttering the human-visible space. But cameras are designed to closely mimic a human eye; how can a commodity camera see the world differently than from the human eye, and decode information not visible to humans?

We present a novel optical barcode design, called the Bokode (Figure 1), which may be read by a camera at relatively large distances. The design exploits the *bokeh* effect of ordinary camera lenses which maps a cone of rays exiting from an out of focus scene point into a disk shaped blur on the camera sensor. From a diffuse point, all these rays have roughly the same radiance, so the bokeh takes the shape of the round aperture. We create a directionally varying set of encoded rays by placing a small diameter ($3mm$), short focal length ($8mm$) lenslet over a printed binary code. The imaged bokeh is hundreds of pixels wide in diameter and it shows a magnified view of the binary code allowing us to decode thousands of bits. The Bokodes appear as *tiny dots to the human eye* (or to a camera in sharp focus), but appear as clearly visible codes to the *out of focus camera*. We observe features as small as $2.5\mu m$ by

---

*now at MIT CSAIL
†now at Osaka University

enabling the Bokode-camera pair to behave like a long distance microscope. We use Bokodes not only for decoding identity, but also for computing the camera pose (distance and angle) with respect to the Bokode.

Contributions: We introduce the concept of a paired display and capture mechanism that exploits directionally varying tags and bokeh of ordinary cameras for relatively long distance communication. We show that we can perform sophisticated geometric operations making them suitable for further use in computer graphics and vision. Our technical contributions are as follows.

- Methods to create and capture angularly encoded information using a lens pair separated by meters.
- A technique for encoding that allows us to recover identification, and relative distance and angle of a camera while observing a single tiny spot.
- Prototype designs for Bokode tags: active using backlighting, passive based on retroreflector, and wide field of view (Krill eye design); and cameras: ordinary out of focus camera, and a translating camera.
- Working systems and performance analysis to demonstrate the concepts for several plausible applications.

Techniques that involve out of focus cameras (e.g. for depth sensing) or devices that create directional rays (e.g. view dependent displays) are both known. But as far as we know, ours is the first system to exploit both creation and capture of bokeh, in a shared setup and enable a new type of machine-to-machine interaction.

2. Related Work

The need for adding intelligence to objects via tags and building human or machine interactions around them is an important area of research and a key requirement for ubiquitous computing [Weiser 1993]. Perhaps the most widely deployed and successful identification methods are barcodes and RFIDs. The problem of creating tags to supplement the physical world has given rise to many passive as well as active solutions which use projectors, bluetooth, LEDs and even IP. Each of these provides a seamless interface between the physical world and the virtual world [Abowd and Mynatt 2000].

Comparison with barcodes: Barcodes are usually decoded using a flying spot scanning laser and a single photodetector which picks up absence or presence of light reflected from the $0-1$ stripes on the barcode [Morton 1994]. This is done to avoid focusing and depth of focus issues common with finite aperture cameras. Newer codes exploit 2D imaging of advanced scanners and camera and pack more information [Pavlidis et al. 1990]. They include Data Matrix [ISO 2006a], QR codes [ISO 2006b], and Aztec codes, all of which use Reed Solomon [MacKay 2003] error correction. Shotcode [de Ipiña et al. 2002] is a circular barcode designed for use with regular cameras. Multiplexed barcodes [Langlotz and Bimber 2007] use multiple color channels and temporally changing codes (using an LCD or a projector) to maximize the data throughput and the robustness of the barcode recognition. Kato and Tan [2007] compared various barcode standards for use with cell phone cameras as the reader. Our method can use any of these encoding schemes as the Bokode pattern. In order for visual codes to gain widespread use, they should be able to be read at a range of distances and printed in different sizes [Pavlidis et al. 1990]. A traditional barcode may be read from far away with a telephoto lens on a standard camera. However a telephoto lens exacerbates the framing problem, making it harder to bring the barcode of interest within the scanner's field of view. A wide angle lens makes the barcode easier to find, but the resulting image has low resolution. On the other hand, Bokodes cannot be printed with standard printers on stickers or on product container; however, anti-counterfeiting tags are a good example of specialized product packaging. To create contrast with background, we use some active light such as a camera flash or a back-lit pattern.

Comparison with RFIDs: Radio Frequency Identification (RFID) tags [Want 2003] are used to determine the presence of an object within a certain range, but do not reveal its location. They suffer from lack of sufficient directionality, and interference with neighboring tags. Although RFID tags can be sensed over distance, they create significant security issues. In the case of passive RFID, the reading distance range may be very limited with the use of a reasonably priced reader. While active RFID might have greater ranges of operation, it requires an on board power source on each tag. However, RFID still has a long way to go before each user is empowered with a simple RFID reader. Localization of RFID tags can be improved by attaching a photosensor [Raskar et al. 2004], but this requires an additional projector.

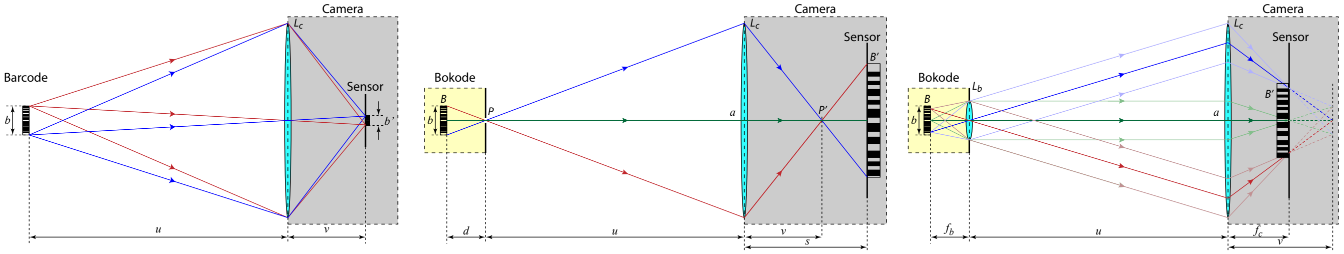
Other self-identifying tags: Optical methods for identifying tags encode in space or time. Barcodes encode in space as described above. Temporal-coding uses blinking lights and gets around limited pixel resolution, dynamic range, and depth of field of a camera. The Sony ID-CAM system [Matsushita et al. 2003] used blinking LEDs to send a unique temporal code that reveals identity. The Hi-Ball tracker [Welch et al. 1999] used temporally multiplexed, ceiling mounted infrared LEDs to estimate the location and pose of a multi-lens sensor. Although temporal codes are easier to detect over larger distances, a limitation is that the camera needs to have fast frame rate or requires complex tracking algorithms [Zhang et al. 2008]. There also exist several non-optical methods based on radio frequency, magnetic and audio systems [Teller et al. 2003].

Camera-pose: Identification information is not sufficient for applications that need to estimate camera-pose with respect to the tagged objects such as AR and motion-capture. To support these kinds of geometric operations from fiducials in a single image requires finding at least four point correspondences. The accuracy of angle estimation is typically a function of the size of the tag. ARTag [Fiala 2005], ARToolKit [Kato and Billinghurst 1999], and Sony EyeToy barcode [EyeToy] use four points on the periphery of the pattern. Tateno et al. [2007] proposed nesting smaller markers within other markers to improve AR performance from a large range of distances. Zhang et al. [2002b] and Claus and Fitzgibbon [2004] perform a survey of several fiducial marker systems, assessing the processing time, identification, and image position accuracy with respect to viewing angle and distance. We define new codes that provide identification, distance, and angle from a single photo. However, the camera that decodes Bokodes is out of focus, so for AR applications, we may need a second camera.

Optical configuration: Our design is based on the defocus blur of a camera. The concept of carefully generating rays in Bokode is closely related to spatial and angular samples of 4D lightfields [Levoy and Hanrahan 1996] and view-dependent displays [Kitamura et al. 2001]. Cameras are often modified to create a desired point-spread function that varies with depth to, for example, extend the depth of field [Levin et al. 2007; Veeraraghavan et al. 2007]. Intentionally taking photos with the camera out of focus is somewhat unusual. Our tag uses a simple lenslet placed at its focal length from a pattern, and this is similar to a Peep-egg or head-mounted display [Azuma et al. 2001] where the virtual image is at infinity. Our design also mimics an infinity corrected microscope, although they are designed to work over a short distance with carefully aligned parallel lenses in a unified optical setup.

3. Optical Communication with Lens Bokeh

We first study the optical factors that make detection of traditional barcodes at a distance challenging. Consider a standard camera



(a) Traditional Camera (b) Pinhole based Bokode setup (c) Lenslet based Bokode setup

Figure 2: *(a) A camera photographing a traditional barcode. The size of the image formed on the sensor reduces as the distance of the camera from the barcode increases. This limits the usable range for such a setup to just a few inches. (b) A pinhole placed in front of a barcode pattern encodes directional rays with the pattern. The camera captures this information by positioning the sensor out of focus. An unbounded amount of magnification is achieved by increasing sensor-lens distance. (c) A small lenslet placed a focal length away from the pattern creates multiple directional beams (ray bundles) for each position in the barcode pattern. The camera lens, focused at infinity images a magnified version of the barcode pattern on the sensor.*

photographing a traditional barcode (Figure 2(a)). A barcode of size b is placed at a distance u from a camera lens. The lens images the barcode on the sensor at a distance v from the lens. The size of the barcode image is $b' = (v/u)b$. The effective magnification scaling is, $M_t = (b'/b) = (v/u)$. For a typical case where a $50mm$ focal length lens takes a photo from about $u \approx 5m$, the focused image is at $v \approx 50mm$ and the magnification is $M_t \approx 0.01$. The magnification reduces as the distance of the camera from the barcode increases.

Our solution is based on defocus blur and provides *depth-independent magnification* of Bokode features. The defocus blur of a point light source on the image sensor is called the *point spread function (PSF)*. The PSF of a camera depends (among other factors), on depth u of the point with respect to the camera, and the camera's plane of focus. A camera focuses by changing the distance between the lens and the image sensor. When the lens focuses on a point source, the image is very close to a point (ignoring lens aberrations). As the plane of focus moves away from the point source, the image expands from a point to fill a circular disc. This disc is also called the circle of confusion. We encode useful information in the bokeh that is visible only when a camera is out of focus.

3.1. Pinhole Bokode

We describe a pin-hole based Bokode which provides a good intuition for bokeh based capture of directionally varying rays (Figure 2(b)). This simple design is useful to understand the relationship among viewable barcode size b, aperture size a, barcode to pinhole distance d and pinhole to camera distance u.

From Figure 2(b), the size of the visible barcode pattern is given by

$$b = ad/u. \tag{1}$$

For a typical camera with an aperture size of $25mm$, distance between the barcode and the pinhole of $5mm$, and the distance of the camera lens from the pinhole of $5m$, the resulting viewable Bokode pattern size is approximately $25\mu m$. Clearly, this barcode size is much smaller than that of traditional printed barcodes.

Next, consider the magnification achieved at the image sensor to observe this tiny code. The lens images the pinhole P to a point P' at a distance v from the lens. According to the thin lens equation, we have

$$\frac{1}{f_c} = \frac{1}{u} + \frac{1}{v}, \tag{2}$$

where f_c is the focal length of the camera lens. Assuming that the pinhole P is infinitely small, the size of the image at a distance v from the lens is also infinitely small. However, as we place the sensor out of focus at a distance s from the lens, we get a highly magnified image of the barcode image. The size of the barcode image is given by

$$b' = (v - s)a/v. \tag{3}$$

In the pinhole model, the magnification properties are easy to understand from the ray diagram, even without Equation 3. As seen in Figure 2(b), the barcode image can be made arbitrarily large by simply moving the lens more out of focus and increasing s. Additionally, a larger lens aperture also gives larger magnification.

With the Bokode's optical setup, the information of the barcode is embedded in the angular and not in the spatial dimension. By throwing the camera out of focus, we capture this angular information in the defocus blur formed on the sensor. The pinhole is blurred, but the information encoded in the bokeh is sharp.

3.2. Lenslet Bokode

The pinhole setup is impractical due to limited light efficiency and diffraction. We replace the pinhole with a small lenslet carefully positioned at a distance equal to its focal length (f_b) away from the barcode pattern as shown in Figure 2(c). The lens collimates the rays coming from a point on the barcode to form a beam or parallel ray bundle. Parallel rays for each point means the virtual image of the barcode is at infinity. The camera focuses at infinity by positioning the camera lens at a distance equal to its focal length f_c from the sensor, and forms an image of the barcode on the sensor.

The viewable part of the barcode produces an image of size

$$b' = (v - f_c)a/v = f_c a/u. \tag{4}$$

Finally, substituting $d = f_b$ and using Equations 1 and 4, we get

$$M_b = f_c/f_b. \tag{5}$$

The resulting optics of a Bokode setup are very similar to that of an infinity corrected microscope. The Bokode lenslet acts like the microscope's objective, and the camera lens is similar to the eyepiece. Unlike a traditional microscope however, there is no tube connecting the two lenses, and we have multiple microscopes sharing the same eye piece in a scene with more than one Bokode.

For a typical setup, we use a Bokode lenslet with focal length of $5mm$, and a camera lens with focal length of $50mm$. The effective magnification factor around $M_b \approx 10$. Compare this to the $M_t \approx 0.01$ obtained in the case of a traditional barcode, the image is 1000 times larger. Furthermore, for a typical viewable Bokode region size of $b = 25\mu m$ (obtained above), we have $b' = 250\mu m$, which implies a coverage of around $50 - 150$ pixels on the image sensor (depending on the sensor pixel size). This also means, for the same sized barcode we can potentially pack 1000 times more bits in both dimension, i.e. million times more data. Equation 1, reminds us that if we reduce lenslet-lens distance u to zero, i.e. by taking photo touching the lenslet (as shown in Figure 12), we can recover all this information.

3.3. Useful Bokode Properties

In addition to the small physical size of a Bokode, the setup has several useful properties that make the it well suited for a range of at-distance tagging, and angle estimation problems compatible with commodity cameras.

Camera focus: In order to image the Bokode pattern, the camera focuses at infinity. This is independent of the Bokode-camera distance. This is a significant advantage over the traditional barcode where cameras require re-focusing when the depth changes. When the scene is in sharp focus, the Bokode ($3mm$ diameter in our prototype), occupies very few pixels in the image and does not intrude into the rest of the scene.

Camera lens aperture: The size of viewable barcode pattern is proportional to the camera aperture size (Equation 1). A relatively large lens aperture is required to see a reasonable part of the Bokode pattern. This explains why the Bokode pattern is effectively 'invisible' to the human eye which has a relatively small pupil size of $2mm$ to $6mm$. In Section 5.2 we discuss ways capturing the Bokode pattern even with a small camera aperture.

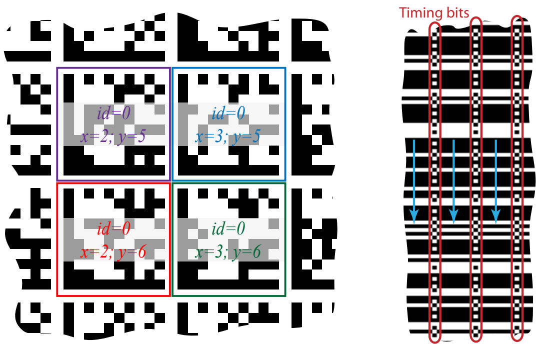
Bokode to camera distance: As shown in Equation 5, the magnification of the optical system is independent of the distance between the Bokode and the camera. This is different from a traditional barcode and makes decoding the Bokode pattern much easier. The size of viewable barcode pattern is inversely proportional to the camera-Bokode distance (Equation 1). Hence, a human eye or small aperture camera, can view the larger pattern by holding it up close. The distance-dependent viewing region may be used to encode hierarchical information into barcodes so that cameras recover more viewable bits as they get closer.

Bokodes orientation: The camera views a different part of the Bokode pattern depending on its position relative to the Bokode. The viewable region of the Bokode is a function of the angle formed between the camera and the Bokode's optical axis. Unlike a traditional barcode, with appropriate pattern design, the Bokode gives completely different pieces of information to cameras in different directions. We use this property of the Bokode for estimating the angle in AR applications.

4. Bokode Pattern Design

The requirements and constraints for the Bokode pattern design are quite different from a traditional 2D barcode. Unlike a traditional barcode, the camera only images a small region of the entire Bokode pattern at a time. This visible Bokode region depends on the distance of the camera from the Bokode and the relative view angle. Our Bokode pattern consists of an array of tiled data matrices such that at least one tile is always imaged by the camera within the working angle and distance range. Instead of simply repeating the same information in each Data Matrix, we vary the data bits across the tiles so that the camera obtains view dependent information, something not possible with traditional barcodes.

The Data Matrix is a two dimensional barcode [ISO 2006a] that uses an matrix of binary cells to store information. As shown in Figure 3(a), we use a tiled array of 10×10 Data Matrix codes with one row/column of silent cells between adjacent codes. The 10×10



(a) Tiled Data Matrix Codes (b) De Bruijn Code

Figure 3: *Bokode pattern designs: (a) A tiled matrix of Data Matrix codes encode identification and angular information. Each* 10×10 *symbol stores its physical position in the Bokode pattern, and a single byte of identification information repeated across each Data Matrix. (b) 1D De Bruijn sequence (along the arrows) accompanied with the timing sequence (highlighted in red). This sequence gives highly accurate angular information but no ID.*

Data Matrix codes encode 3 bytes of data, and another 5 bytes of Reed-Solomon error correcting code. This error correcting code ensures data integrity when up to 30% of the symbol is damaged; we rely on this redundancy to decode overlapping Bokode patterns as discussed in Section 5.3.

Since the visible Bokode pattern depends on the camera angle, the tiled Data Matrix design offers three independent bytes of information for different orientations of the Bokode relative to the camera. While these bits could be assigned arbitrarily according to the applications, we allocate one byte to the ID that is repeated across all the Data Matrix codes, and one byte each for the x and y positions of the Data Matrix in the Bokode pattern matrix. This allows the camera to directly read out the relative position of the viewable Data Matrix code in addition to the Bokode ID from the captured photo. A camera that images the Data Matrix with a position code (x, y) has an angle relative to the Bokode given by the azimuth, $\phi = \arctan((y - y_0)/(x - x_0))$, and the zenith, $\theta = \arctan((\sqrt{(x - x_0)^2 + (y - y_0)^2} \cdot \delta)/f_b)$, where (x_0, y_0) is the Data Matrix code corresponding to the zenith ($\phi = 0$), f_b is the focal length of the lenslet, and δ is the physical size of a single Data Matrix code including the silent space. Additionally, the displacement of each visible Data Matrix from the center of the bokeh circle gives a better estimate of the angular position of the camera using code interpolation. We compute the fractional position (x_f, y_f) of the center of the bokeh circle (p_c) using the coordinates of four corners of the Data Matrix $p_n (n = 0 \cdots 3)$,

$$\begin{pmatrix} x_f \\ y_f \end{pmatrix} = \begin{pmatrix} v_x & v_y \end{pmatrix}^{-1} v_d + \begin{pmatrix} x \\ y \end{pmatrix}, \tag{6}$$

where $v_x = p_1 - p_0$, $v_y = p_0 - p_3$ and $v_d = p_c - (p_0 + p_1 + p_2 + p_3)/4$. We use the average of the fractional codes if multiple Data Matrix codes of the a single Bokode are observed.

Unlike most fiducial based pose estimation techniques, the camera directly reads out the digital angular information from the Bokode, and does not have to estimate the angle based on the local shape of the fiducial. A clear advantage of the Bokode design for angle measurement is the case where the camera is exactly overhead a single planar fiducial (the line from the camera to the fiducial is perpendicular to the fiducial). In this case angle estimates based on distortion may be unstable, but Bokode based angle measurement is robust. Conversely, such fiducials perform very well under grazing

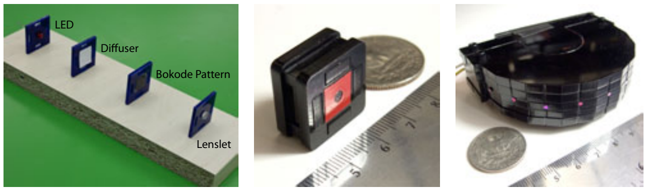


Figure 4: *Photos of Bokode prototypes discussed in the paper: (left) exploded view of our active illumination prototype, (center) an assembled Bokode, and (right) our compound superposition eye (Antarctic Krill) prototype for an ultra wide angular range.*

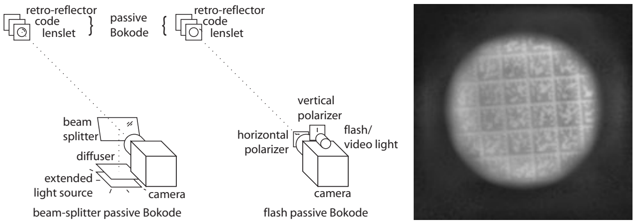


Figure 5: *Passive Bokodes replace the LED behind the Bokode pattern with a retroreflector and use the camera flash as the illumination source. (Left) Beamsplitter and cross-polarizer arrangements to illuminate the Bokode. (Right) Image of a passive Bokode captured with a camera and a light source behind a beamsplitter.*

angles, a condition under which the Bokode may not provide any data at all due to its limited angular range.

While tiled Data Matrix codes provide a generic Bokode pattern design for most applications, certain applications may benefit from specialized Bokode patterns. For applications that only need identification and no angular information, the same Data Matrix code can be repeated across the pattern, thus giving three bytes for identification. On the other hand, highly accurate angle information is read from a pattern that encodes a binary De Bruijn sequence [Zhang et al. 2002a]. A De Bruijn sequence $B(l, k)$ is a sequence with a symbol size l ($l = 2$ in our case since we use binary bits), for which every possible contiguous subsequence of length k or more appears exactly once. If we see k or more bits in the visible Bokode pattern, we can accurately estimate the angle. Figure 3(b) shows a part of a Bokode pattern containing a 1D De Bruijn sequence. Unlike the Data Matrix codes, the De Bruijn sequence does not offer robust error detection or identification.

5. Prototypes and Implementation

We explore several techniques to create and capture the Bokodes. Each design has advantages and limitations, and we believe that the best technique will depend on the exact application.

5.1. Generating Bokodes

Figure 4(a) shows an exploded view of our **active Bokode** prototype. We use a plano-convex lens with a $3mm$ diameter and $8mm$ focal length as the Bokode lenslet, and a battery powered LED for backlight. We print the Bokode pattern with a pixel size of $15\mu m$ on a transparency using a Heidelberg Herkules printer. The acrylic housing ensures that the distance between the lenslet and the pattern is exactly equal to the lenslet's focal length ($8mm$). The Bokode is prototyped for easy modification and a smaller housing is possible for manufacturing. Figure 4(b) shows a smaller assembled prototype that we used for some of our experiments.

We create completely **passive Bokodes** by replacing the LED with a *retroreflector* (Figure 5). We use the camera flash to illuminate the pattern behind the Bokode lenslet, and the retroreflector reflects the light back towards the camera lens. We place a polarizer in front of the camera lens, and another in front of the flash such that their polarization direction is perpendicular to one another. This eliminates the specular reflection from the Bokode lenslet and enhances the contrast of the Bokode pattern. We also experimented with a beamsplitter to position the flash at the center of projection of the camera. Figure 5 shows an image of the passive Bokode captured with the beamsplitter technique. We do not use the passive Bokode for any of the results in this paper due to low contrast with the current prototype (Section 7 discusses possible improvements).

Figure 6: *Capturing a 1D Bokode using camera translation. (Left) The camera translates perpendicular to the Bokode's optical axis during the length of the exposure. The camera has a small aperture size of $f/8$, the lens focuses at infinity, and the total translation is about $4cm$. (Right) Resulting photo as captured by the camera.*

We also explore novel optical designs that allow the camera to image the Bokode from extreme angles (Figure 4(c)). The design uses gradient index (GRIN) lenses, and is very similar to refractive compound superposition eye of the **Antarctic Krill** [Fernald 2006]. Please see the supplementary material for more information and preliminary results.

5.2. Capturing Bokodes

We used consumer cameras, Canon Digital Rebel XSi and the Canon 5D II, paired with reasonably large aperture lenses, EF $85mm$ $f/1.8$ and the EF $50mm$ $f/1.8$ for majority of the demonstrations. The lenses were used at their largest aperture setting, and manually focused to infinity.

For AR applications we capture two photos (one focused at the scene, and another at infinity) to simultaneously capture the scene and the Bokode information. We use a beamsplitter and two synchronized cameras that share the same center of projection. In the future this can be achieved with a camera changing aperture from narrowest to widest, with an auto-focus like mechanism for changing focus from Bokode distance to infinity in successive frames, or with an extra out of focus sensor for cameras with multiple CCDs.

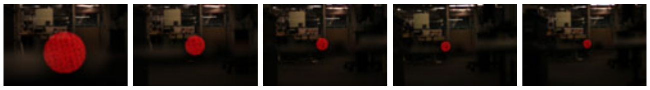
A camera with a much smaller aperture (such as a mobile phone camera) will see limited part of the Bokode pattern unless it is at close range. But, even from a large depth, we can make a larger part of the code viewable. We translate the camera within one exposure time with the lens focused at infinity (Figure 6). Different parts of the Bokode pattern are imaged on different parts of the sensor. This translation simulates a larger aperture in a single exposure, and is similar to X-ray laminography [Mohan et al. 2009].

5.3. Performance

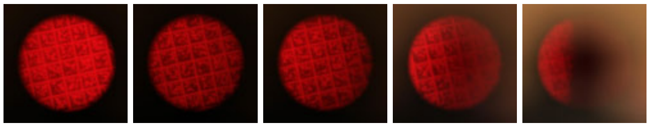
We analyze the properties of our Bokode prototype and evaluate its performance under several conditions. This analysis is specific to our current prototype that uses hand-assembled components and a pattern with a feature size of $p = 15\mu m$. A higher resolution pattern would greatly improve many of these characteristics.

Resolution limit: The printing resolution currently sets a limit on the information content of the Bokode pattern design. Additionally, diffraction due to the finite size of the Bokode lenslet sets a hard limit on the maximum resolving power of the Bokode. The angular resolution of the system is given by $\sin\theta = 1.22(\lambda/a_b)$, where λ is the wavelength of light, and a_b is the diameter of the lenslet. For

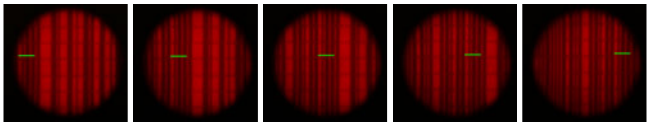
(a) Effect of ambient light on the Bokode pattern contrast. (Left to right) Increasing ambient light intensity by aiming multiple fluorescent lamps directly at the Bokode with a fixed exposure time and Bokode LED intensity. (Right most) Photo of a Bokode taken outdoors in sunlight.



(b) Photos with the camera at increasing distances away from the barcode ($25cm$ for the leftmost and $2m$ for the rightmost photo). The $15\mu m$ feature size Data Matrix codes are decoded reliably at distances up to $2m$ with a $85mm$ $f/1.8$ lens.



(c) Camera positioned at angles of $0°$, $10°$, $15°$, $20°$, and, $25°$ (left to right) with the Bokode's optical axis. Lens aberration degrades the image quality for angles greater than $20°$, making the Data Matrix codes harder to decode.



(d) Photos of a Bokode rotating by $1°$ in each successive frame. The green marker follows a single $k = 9$ De Bruijn sequence. As expected analytically, the green marker reliably shifts 10 positions with each degree.

Figure 7: *Performance of our Bokode prototype under changing illumination, distance, and angle.*

our prototype with a lenslet aperture of $3mm$ and a focal length of $8mm$, the resolution limit is approximately $1.8\mu m$ for visible light.

Ambient light: The contrast of the image of the Bokode pattern depends on the ambient light and texture around the Bokode lens. Figure 7(a) shows the image of a Bokode pattern captured under varying illumination conditions, including outdoors. High frequency texture surrounding the Bokode also reduces the contrast.

Working distance and angle range: Figure 7(b) shows photos of the Bokode captured at different Bokode-camera distances. Consistent with Equation 1, the maximum distance at which we see one complete Data Matrix is around $2m$. Reducing the pattern feature size from $15\mu m$ to $5\mu m$ increases this distance to $6m$ while still remaining within the diffraction limit. Figure 8 shows our initial results obtained for a 1D Bokode pattern with a $2.5\mu m$ feature size generated using electron beam lithography. We obtain photos with good contrast and detail from over $4m$ away from the Bokode. Figure 7(c) shows photos captured at different camera angles to the Bokode's optical axis. We obtain robustly decodable codes for a cone of approximately $20°$ around the optical axis. The angular range is limited by lens aberrations and vignetting due to directional nature of the LED. A higher quality lens and more diffuse LED may increase this range, but may result in an increased cost and reduced brightness.

Angle and depth estimation: We use a 1D De Bruijn sequence (Figure 3(b)) to test the angular resolution of our Bokode prototype. For our current prototype we have an angular resolution of approximately $\alpha = \arctan(p/f_b) \approx 0.1°$. Figure 7(d) shows the angular resolution where each successive image corresponds to a $1°$ angular shift. Figure 9 compares the angle estimation robustness of our Bokode prototype to ARToolKit [Kato and Billinghurst 1999] for small changes in angle, with the camera positioned along the optical axis for the Bokode, and exactly overhead for the planar tag. ARToolKit relies on changes in the shape of the black rectangle to estimate the angle. The changes in shape are significant when the tag is viewed from oblique angles, and provide reliable angle estimates. However, these changes are subtle when the camera is exactly overhead, resulting in jitter noise. Angle estimation from the Bokode is more robust in this case because it primarily relies on the digital information contained in the visible Data Matrix codes. We estimate the depth of the Bokode from the size of the circle of confusion it produces on the sensor (Equation 4). Like other depth from defocus systems, the depth resolution falls off inversely with distance [Pentland 1987].

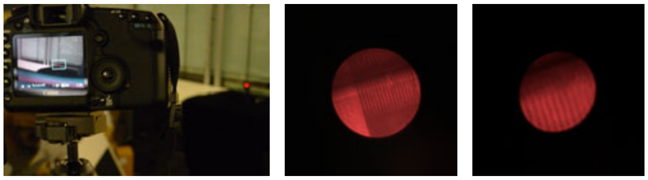


Figure 8: *Photos of a 1D Bokode pattern with $2\mu m$ feature size taken from $1m$ away (center), and over $4m$ away (right). The images are cropped from larger photos of the scene (see camera LCD).*

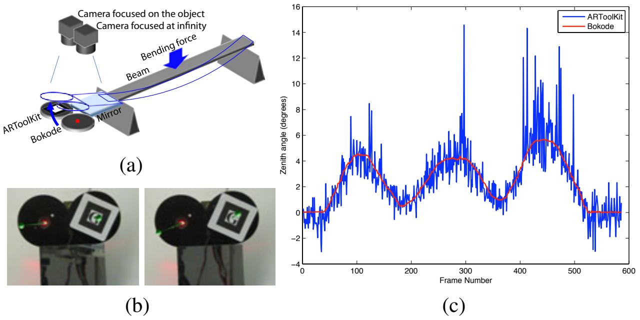


Figure 9: *Angle estimation robustness comparison showing best case of our Bokode prototype and worst case of ARToolKit. (a) We place a wooden beam between two wedges, and place a Bokode and an ARToolKit tag at one end of the beam. Two cameras observe the Bokode and tag from directly overhead; one focuses on the scene, and the other at infinity. (b) In focus photos with arrows depicting the estimated angles. (c) With the camera directly overhead, the estimated zenith angles using Bokodes are more robust than those obtained by ARToolKit. ARToolKit work best when the tag is viewed at an oblique angle (a weakness of the Bokode design); it is highly suboptimal to view the tag directly overhead (Bokode's strength).*

Overlap: Figure 10 shows photos in which multiple Bokode patterns overlap. We rely on the redundancy introduced by the tiling of Data Matrix codes, and the error correction offered by Reed Solomon coding to decode overlapping Bokodes. Unfortunately, Bokode pattern may be unrecoverable when the overlap is too great or if vital Data Matrix components such as the timing bits are lost.

6. Applications and Results

The most obvious application for the Bokode is for its use as a **identifier barcode** (see Figure 1). Unlike a traditional barcode, the Bokode is viewable from afar ($2m - 3m$) with a standard camera. Since the Bokode allows the barcode information to be directionally varying, a Bokode can send different information to cameras in different directions. Since the viewable Bokode pattern increases as

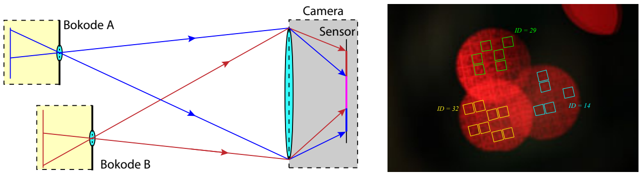


Figure 10: *Recovering data from overlapping Bokodes: (left) ray diagram demonstrating overlap of multiple Bokode patterns; (right) we recover all three IDs of overlapping Bokodes due to repetition coding offered by the tiled Data Matrix.*

the camera comes closer, the reader can learn more about a product as it moves closer. For a camera with a small aperture, we see a very large part of the Bokode pattern when the camera is held right next to the Bokode lenslet (Figure 12). This allows us to pack more information in a smaller physical region when the Bokode is used like a traditional barcode at close range.

The small physical size and the unobtrusive nature of Bokodes make them suitable for estimating camera pose and distance for certain **augmented reality** and **motion capture** applications (see Figure 11 and accompanying video). Unfortunately, the angular range of the current Bokode prototype is limited to approximately 20° as shown in Figure 7(c), and this may limit its applicability to general purpose AR applications. Combining a Bokode with existing planar fiducial based AR tags may provide reliable angular estimates for a wide range of angles. Unlike many other AR techniques, the Bokode itself occupies very few pixels in the focused photo, and is relatively unobtrusive to the scene. However, we require the use of a second camera focused at infinity to capture the angular information. For motion capture, we get identification in addition to position and orientation, and the system does not have to deal with marker swapping and marker reacquisition, even when the markers go outside the scene and come back.

7. Discussion

Limitations and Future Work: The Bokode design in its present form suffers from several limitations. The visible region of the Bokode pattern depends on the camera's aperture size, and this limits the class of cameras that the Bokode currently works with at a reasonable distance. Higher resolution patterns allows for the use of smaller aperture cameras, and our initial experiments with electron beam lithography suggest that we can easily get close to a micron feature size (Figure 8). The physical thickness of the Bokode is greater than traditional barcode or optical marker. It might be possible to reduce the depth with an origami lens [Tremblay et al. 2007], a Fresnel lenslet or a reflective transmission mode holographic optical element computed as a Fourier hologram or as a single hogel (holographic pixel). While the Bokode does need a relatively dark area with a low frequency texture around it, our experiments in Section 5.3 reveal that it is reasonably robust to ambient light. Auto-exposure and motion blur on some cameras may result in poor Bokode image quality, but a Bokode-mode on the camera may help. We can improve the current passive Bokode design by using better retroreflectors, or fluorescent reflectors coupled with a UV camera flash. The current Bokode pattern design is relatively simple and does not attempt to achieve an optimal performance. We are exploring the use of more advanced coding schemes to get higher bit rates, and to deal with overlapping Bokodes better. Finally, the current Bokode prototypes have a limited angular range of approximately 20°; we are investigating the use of multiple Bokodes, better quality lenslets, and the Krill eye design to increase this range.

Future Application Scenarios: While the Bokode may not completely replace competing techniques for any single application, we believe its versatility and flexibility will make it a useful tool in graphics, vision and human computer interaction research, and may even open up new application areas that need identification, tagging, or angle information at a distance. Applications discussed in Figure 12 include communication with camera-driven street-mapping services and rear-projected tabletop touch-based interfaces. With Bokodes in street billboards, the human eye will see the billboard information but an out of focus camera will capture the Data Matrix indicating a website link. Dynamic Bokodes and cameras embedded in cell phones might offer a solution for high speed near-field communication. Bokodes may also be used as multi-user interaction devices with large displays.

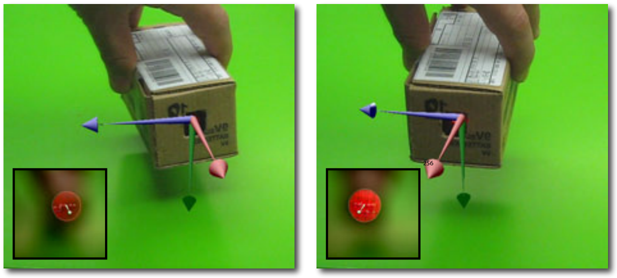


Figure 11: *Use of Bokodes to estimate angle for AR. A camera focused at infinity images the Bokode pattern (inset) and uses this information to estimate the relative camera pose. A second camera focused on the scene images the Bokode to a very small number of pixels. Augmented arrows show the axes relative to the object face.*

Conclusion: Conventional techniques encode information in physical dimensions of space, time and wavelength. The Bokode design presented in this paper encodes and decodes information in angular dimension. This allows standard cameras to see the world around us differently from how the human eye sees it, and allows a camera to detect identity and the relative angle to a small optical tag from a reasonably large distance. The Bokode design enhances the flexibility and usefulness of the classic barcode by allowing users to read and interpret them from large distances using equipment they may already have. We believe that the use of such personal barcode scanners opens up new avenues in the area of ubiquitous computing and human computer interaction. Additionally, we believe the Bokode design may benefit several other applications such as augmented reality, motion capture, computational probes in computational photography, and tools for camera calibration.

Acknowledgments

Thanks to Meng Heng Touch and Donald Eng for Bokode prototype construction; Paula Aguilera, Jonathan Williams and Evan Kutsko for the video; Eugene Wu for the sketches; Masahiko Inami, Frank Moss and Andy Lipman for useful discussions; the entire Camera Culture group for their unrelenting support; and the reviewers for their valuable feedback. This research was supported by research grants from Nokia Research and Samsung. Ramesh Raskar is supported by an Alfred P. Sloan Research Fellowship.

References

Abowd, G. D., and Mynatt, E. D. 2000. Charting past, present, and future research in ubiquitous computing. *ACM Trans. Comput.-Hum. Interact. 7*, 1, 29–58. 2

Azuma, R., Baillot, Y., Behringer, R., Feiner, S., Julier, S., and MacIntyre, B. 2001. Recent advances in augmented reality. *IEEE CG&A 21*, 6, 34–47. 2

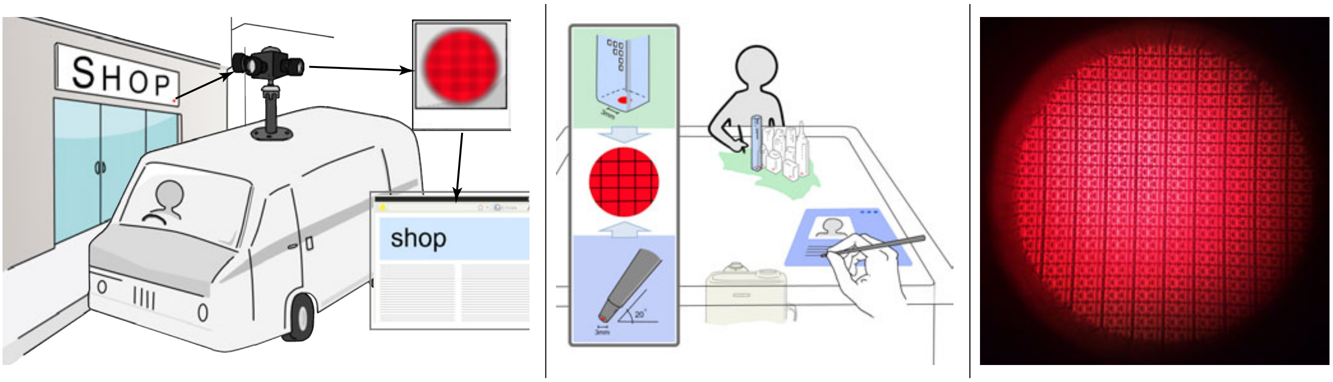


Figure 12: *Potential future application scenarios for Bokodes: (Left) Street Mapping Services: Shops may use Bokodes to provide useful information to the trucks as they drive down the street taking pictures. The trucks use a camera focused at infinity, and image the Bokode pattern using the translation technique discussed in Section 5.2. (Center) Rear projected Microsoft Surface: A camera focused at infinity can capture the information contained in tiny Bokodes attached under physical objects placed on the table surface. The system decodes position, identification, and the angle the object (such as a stylus pen) makes with the surface. (Right) Photo of a Bokode pattern (zoom in to see detail) taken by a cell phone camera right next to the Bokode lenslet gives significantly more bits than a traditional barcode of similar size.*

Claus, D., and Fitzgibbon, A. W. 2004. Reliable fiducial detection in natural scenes. In *ECCV 2004*, 469–480. 2

de Ipiña, D. L., Mendonça, P. R. S., and Hopper, A. 2002. TRIP: A low-cost vision-based location system for ubiquitous computing. *Personal Ubiquitous Computing 6*, 3, 206–219. 2

Fernald, R. D. 2006. Casting a genetic light on the evolution of eyes. *Science 313*, 5795, 1914–1918. 5

Fiala, M. 2005. ARTag, a fiducial marker system using digital techniques. In *IEEE CVPR*, vol. 2, 590–596. 2

ISO, 2006. Data Matrix bar code symbology specification. ISO/IEC 16022:2006. 2, 4

ISO, 2006. QR code 2005 bar code symbology specification. ISO/IEC 18004:2006. 2

Kato, H., and Billinghurst, M. 1999. Marker tracking and hmd calibration for a video-based augmented reality conferencing system. In *IWAR 99*, 85–94. 2, 6

Kato, H., and Tan, K. T. 2007. Pervasive 2D barcodes for camera phone applications. *IEEE Pervasive Computing 6*, 4, 76–85. 2

Kitamura, Y., Konishi, T., Yamamoto, S., and Kishino, F. 2001. Interactive stereoscopic display for three or more users. In *SIGGRAPH 2001*, ACM, 231–240. 2

Langlotz, T., and Bimber, O. 2007. Unsynchronized 4D barcodes. In *ISVC*, 363–374. 2

Levin, A., Fergus, R., Durand, F., and Freeman, W. T. 2007. Image and depth from a conventional camera with a coded aperture. In *SIGGRAPH 2007*, vol. 26. 2

Levoy, M., and Hanrahan, P. 1996. Light field rendering. In *SIGGRAPH 1996*, 31–42. 2

MacKay, D. J. C. 2003. *Information Theory, Inference, and Learning Algorithms*. Cambridge University Press. 2

Matsushita, N., Hihara, D., Ushiro, T., Yoshimura, S., Rekimoto, J., and Yamamoto, Y. 2003. ID CAM: A smart camera for scene capturing and id recognition. In *International Symposium on Mixed and Augmented Reality*, 227–236. 2

Mohan, A., Lanman, D., Hiura, S., and Raskar, R. 2009. Image Destabilization: Programmable defocus using lens and sensor motion. In *IEEE ICCP*. 5

Morton, A. Q. 1994. Packaging history: The emergence of the Uniform Product Code (UPC) in the United States. *History and Technology 11*, 1, 101–111. 2

Pavlidis, T., Swartz, J., and Wang, Y. P. 1990. Fundamentals of bar code information theory. *IEEE Computer 23*, 4, 74–86. 2

Pentland, A. P. 1987. A new sense for depth of field. *IEEE PAMI 9*, 4, 523–531. 6

Raskar, R., Beardsley, P., van Baar, J., Wang, Y., Dietz, P., Lee, J., Leigh, D., and Willwacher, T. 2004. RFIG Lamps: Interacting with a self-describing world via photosensing wireless tags and projectors. In *SIGGRAPH*, 406–415. 2

Sony, EyeToy. `http://www.eyetoy.com`. 2

Tateno, K., Kitahara, I., and Ohta, Y. 2007. A nested marker for augmented reality. In *IEEE VR*, 259–262. 2

Teller, S., Chen, K., and Balakrishnan, H. 2003. Pervasive pose-aware applications and infrastructure. *IEEE Computer Graphics and Applications* (July/August). 2

Tremblay, E. J., Stack, R. A., Morrison, R. L., and Ford, J. E. 2007. Ultrathin cameras using annular folded optics. *Applied Optics 46*, 4, 463–471. 7

Veeraraghavan, A., Raskar, R., Agrawal, A., Mohan, A., and Tumblin, J. 2007. Dappled Photography: Mask enhanced cameras for heterodyned light fields and coded aperture refocusing. In *SIGGRAPH 2007*, ACM, vol. 26, 69:1–69:12. 2

Want, R. 2003. RFID: A key to automating everything. *Scientific American*. 2

Weiser, M. 1993. Ubiquitous Computing. *IEEE Computer 26*, 10, 71–72. 2

Welch, G., Bishop, G., Vicci, L., Brumback, S., Keller, K., and Colucci, D. 1999. The HiBall Tracker: High-performance wide-area tracking for virtual and augmented environments. In *ACM VRST*. 2

Zhang, L., Curless, B., and Seitz, S. M. 2002. Rapid shape acquisition using color structured light and multi-pass dynamic programming. In *IEEE 3DPVT*, 24–36. 5

Zhang, X., Fronz, S., and Navab, N. 2002. Visual marker detection and decoding in AR systems: A comparative study. In *ISMAR 2002*, 97–106. 2

Zhang, L., Subramaniam, N., Lin, R., Raskar, R., and Nayar, S. 2008. Capturing images with sparse informational pixels using projected 3D tags. In *IEEE Virtual Reality*. 2