Discrete camera calibration from the information distance between pixel streams

Etienne Grossmann
Tyzx, Inc
etienne@tyzx.com

Francesco Orabona
LIRA-Lab, U. of Genoa
Francesco.Orabona@unige.it

José António Gaspar
ISR, Instituto Superior Técnico
jag@isr.ist.utl.pt

Abstract

We consider the problem of estimating the relative orientation of a number of individual photocells -or pixels- that hold fixed relative positions. The photocells measure the intensity of light traveling on a pencil of lines. We assume that the light-field thus sampled is changing, e.g. as the result of motion of the sensors and use the obtained measurements to estimate the orientations of the photocells.

We explore an information-theoretic and geometric approach: based on real-world data, we build a non-parametric functional relation linking the information distance between the data streams of two photocells, and the angular separation between the photocells. Then, given data streams produced by arrays of pixels in similar conditions, we use the functional relation to estimate the angles between pixels. Finally, we embed the estimated angles in the unit 3D sphere to obtain the estimated layout of the array.

1. Introduction

When we look around or read a conference paper, we only perceive the part of the visual stimulus that is relevant to us. Our human visual system performs an abstraction of the underlying optical and biological processes, of which we are mostly unaware.

Computer vision also strives to only "perceive" the relevant part, and performs an abstraction of the underlying optical, electronic and computational processes. A successful example is camera calibration, that allows to transform distorted images into canonical perspective (or other) images. These images are better fit for higher-level processing stages, by making abstraction of lens distortion, raw image size and imager geometry. In a calibrated image, each pixel (u, v) is mapped to a known 3D direction (x, y, z), and determining that mapping is the task of calibration.

Calibration is typically bootstrapped by using some basic properties of the imaging system. In particular, in all the cases that we are aware of, the local topology is known: pixel (u, v) is a neighbor of $(u + 1, v)$; the triplet (u, v), $(u + 1, v)$, $(u + 2, v)$ is approximately aligned. Locally, and for many practical purposes, an uncalibrated image looks like a calibrated image. In particular, it is possible to perform processing operations such as edge, extrema, corner and feature localization on uncalibrated images, just as well as on calibrated images.

This paper departs from traditional computer vision on this point: we do not assume that pixels (u, v) and $(u + 1, v)$ are neighbors. In fact, we are given pixels indexed by a single index n. Since such a camera is peculiar by the absence of a-priori known topological information, it is only natural to call it a *discontinuous*, or *discrete camera*.

The problem that we address may seem too general to pertain to computer vision. Indeed, digital cameras are unlikely to ever produce discontinuous images[1]. Yet we must also consider the computation that occur in biological systems, where it is unclear how much of the geometry of the sensor is known beforehand, and how much is determined by processing visual stimuli. We must also consider robots equipped with arrays of photocells, and new visual sensors [16] that may have variable geometry.

Our goal here is thus to determine the 3D direction pointed by each pixel, under the assumption that pixels sample light that travels along rays that intersect in a unique *center of projection*. In the impossibility of performing local image processing, traditional calibration techniques [20, 8] are out of the question.

Less traditional non-parametric methods that assume a smooth image mapping and smooth motion [11, 7] can obviously not be applied either. By using controlled-light stimuli or known calibration, matches could be obtained, allowing to use match-based non-parametric techniques [14]. In [4], a non-central projection sensor samples the light-field in a discontinuous way: the response at each pixel is the convolution of the lightfield with a possibly multimodal function; this "point-spread function" is only sufficiently estimated to reconstruct images on a fixed plane, and calibration is done with controlled stimuli. In this study however, we wish to exclude known calibration objects and other controlled stimuli.

It thus seems that we are left with statistical approaches. The input data will consist in a stream of values $x(i, t)$, where the integer index i denotes the photocell or pixel, and t, the time. Statistical properties of this stream will be exploited to determine the geometry of the sensor.

[0]This work was partially supported by TYZX, Inc, by the Portuguese FCT POS_C program that includes FEDER funds, and by the EU-project URUS FP6-EU-IST-045062

[1]Except in rare cases such as a SLR camera with a stereo adapter or split mirror.

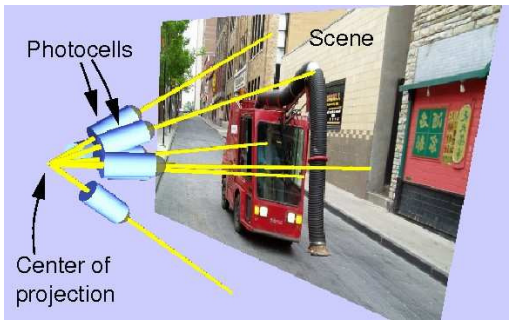


Figure 1. A discrete camera consists of a number of photocells (pixels) that measure the light traveling along pencil of lines. More generally, one could consider that each photocell is characterized by a point spread function defined on the 3D sphere.

1.1. Related work

Our approach to calibration is closely related to the work of Pierce and Kuipers [12], who measure the dissimilarity, or distance, between sensor elements that are not necessarily light sensors. The elements are then embedded in a metric space using metric scaling [9], which also determines the dimension of the space. A relaxation method then improves this embedding, so that the Euclidean distance between sensor elements better matches the dissimilarity between the sensor inputs. Getting very close to the problem addressed in the present paper, the authors use this method to reconstitute the geometry of a rectangular array of visual sensors that scans a fronto-parallel image.

Going further, Olsson et al. [10] use the information metric of [1] as a more appropriate method to measure the distance between visual or other sensor elements. They also show how visual sensors -the pixels of the camera of a mobile robot- can be mapped to a plane, either using the method of [12], or their own, that embeds sensor elements specifically in a square grid.

The works of Olsson et al. and of Pierce and Kuipers are very interesting to computer vision researchers, but they only give a glimpse of what sensor input encoding could do for computer vision. In particular, the geometry of the embedding space is either abstract or fixed to a grid, and lacks a general connection to the geometry of the sensor. Filling this gap is one of the motivations for the present work.

Since the present work exploits statistical properties of the light-field of the world surrounding a light sensor, this article is also related to research on the statistical properties of real-world images. In that area, a model of image formation is used, but images, rather than sequences, are studied. That research has put in evidence fundamental properties, in terms of local, global and spectral statistics, of real-world images, and found ways to exploit these properties for computer vision tasks, such as classification [19], image restoration [5] and 3D inference [13].

Although these results are of great interest, they are not directly applicable in our case, first because we lack images and because that research does not use temporal information that is crucial to our work.

Moreover, these statistics are about images formed on a planar image plane, which is hindrance in our case: first, we do not want to exclude the case of visual sensor elements that are separated by more than 180 degrees, such as the increasingly popular omnidirectional cameras. Also, the local statistical properties of perspective images depend of the orientation of the image plane with respect to the scene, except in special constrained cases such as the fronto-parallel "leaf world" of Wu et al. [21]. Defining images on the unit sphere thus appears as a natural way to render image statistics independent of the sensor orientation, at least with proper assumptions on the surrounding world and/or the motion of the sensor.

1.2. Proposed approach

In the present work, we assume that the statistical properties of the data streams produced by pairs of sensor elements depends only the angular separation between the elements. This assumption, when one considers the images as observations of a random field defined on the sphere, is equivalent [15] to assuming that the random field is homogeneous - we could say "anisotropic." This assumption clearly does not hold in an anisotropic world, unless the orientation of the sensor is uniformly distributed amongst all unitary transformations of the sphere. In terms of computer vision and robotics, our assumption amounts to saying that the sensor is randomly oriented, so that each photocell is just as likely to sample the light-field in any direction.

The great practical utility of this assumption is that, as a consequence, the statistical properties of a pair of data streams generated by two photocells depend only on the angle separating them. In this situation, one can envision estimating the angular separation from a measure of dissimilarity (information distance, Sec. 2.1) between the streams, and this is precisely what we do (Sec. 2.5). In order to estimate angles from dissimilarity measures, we empirically observe (Sec. 2) the relation between these two quantities. We may then (Sec. 3.1) embed these angular estimates using simple techniques from distance geometry [2]. The whole process is outlined in Algorithm 1.

Finally, Section 4 presents some conclusions and a long list of questions to be addressed by future research.

2. The relation between angular distance and information distance

In this section, we recall the definition of information distance [1]. We then study how the information distance between two pixel signals is related to the angle formed by the pixels, assuming a discrete camera moving indoors.

Algorithm 1 Procedure for estimating the geometry of a discrete camera.

Input: Signals produced by N rigidly connected light sensors at T time instants: $x(i,t)$, $1 \leq i \leq N$, $1 \leq t \leq T$. The light sensors are rigidly connected and point towards unknown 3D directions $X_1, \ldots, X_N$, $X_i \in \mathbb{R}^3$, $X^\top X = 1$.

Output: Estimates of the directions $X_1, \ldots, X_N$.

Algorithm:

1. Estimate the information distance d_{ij} between the temporal signals $x(i,.)$ and $x(j,.)$, for each $1 \leq i, j \leq N$.
2. Estimate the angular separations $\theta_{ij} = \cos\left(X_i^\top X_j\right)$.
3. Embed the angular separation in the 3D sphere: find $X_1, \ldots, X_N$ s.t. $X_i^\top X_i = 1$, $X_i^\top X_j = \cos(\theta_{ij})$, $1 \leq i, j \leq N$.

2.1. Information distance

Given two random variables x and y (in our case, the values produced by individual pixels of a discrete camera) taking values in a discrete set $\{1, \ldots, Q\}$, the *normalized information distance* between x and y is [1]:

$$d(x,y) = 2 - (H(y) + H(x)) / H(x,y), \qquad (1)$$

where $H(x,y)$ is the Shannon entropy of the paired random variable (x,y), and $H(x)$ and $H(y)$ are the entropies of x and y, respectively. It is easy to show that Eq. (1) defines a distance over random variables and that it takes values in $[0,1]$.[2]

We now explain how we study the relation between information distance and angular separation. For this purpose, we consider a discrete camera with pixels separated by angles ranging from 0.5 to 180 degrees.

2.2. Image sensor

It is convenient to take a discrete camera consisting of pixels located on a plane, separated by geometrically increasing angles, spanning half a great circle (180°) of the 3D sphere (Fig. 2, top). The smallest gap is half a degree.

We simulate a discrete camera with known Euclidean geometry by sampling a calibrated panoramic image with unique projection center at fixed locations. Images are acquired by a VStone catadiopric camera consisting of a perspective camera fitted to a hyperbolic mirror. This system is modeled as single projection center camera [6] with a $360° \times 210°$ field of view with a $\sim 45°$ blind spot at the south pole. The mirror occupies a 453×453 pixel region of the image. Image values at non-integer pixel locations are obtained by bilinear interpolation. The angular separation between neighboring pixels in the panoramic is usually slightly smaller than 0.5°, so that, in the "tighter" part of the discrete camera layout, there exists a slight linear dependence between the values of consecutive pixels. Also, some mild vignetting occurs[3], but, except for these minor inconveniences, simulating a discrete camera by an omnidirectional camera presents many advantages: no other specialized hardware is needed and each omnidirectional image can be used to simulate many discrete camera "images", as in Fig. 2, bottom. With respect to perspective cameras, the available field of view allows to study very-wide-angle discrete cameras.

2.3. Data acquisition

The camera is hand-held and undergoes "random" general rotation and translation, according to the author´s whim, while remaining near the middle of the room, at 1.0 to 1.8 meters from the ground. We acquired three sequences consecutively, in very similar conditions and joined them in a single sequence totaling 1359 images, i.e. approximately 5 minutes of video at ˜4.5 frames per second.

To simulate the discrete camera, we randomly choose an orientation (i.e. half a great circle) such that all pixels of the discrete camera fall in the field of view of the panoramic camera. Figure 2 shows two such choices of orientations. For each choice of orientation, we produce a sequence of samples $x(i,t)$, $1 \leq i \leq 31$, $1 \leq t \leq 1359$, where each $x(i,t) \in \{0, \ldots, 255\}$. Choosing 100 different orientations, we obtain as 100 discrete sensors and 100 arrays of data $x_n(i,t)$, $1 \leq n \leq 100$.

It is convenient to re-quantize the signals from 255 gray-levels to a more parsimonious representation with $Q < 255$ bins. As in [10], we choose bins that maximize the entropy, i.e. bins that contain equal numbers of values $x_n(i,t)$, $1 \leq i \leq 31$, $1 \leq t \leq 1359$ and $1 \leq n \leq 100$.

Considering any pair of pixel (indices) $1 \leq i, j \leq 31$, the angular separation $\theta_{i,j}$ is the main object to estimate by our method, and it is known in our test setup. We show in Sec. 2.4 how to compute the information distance $d_n(i,j)$ between signals $x_n(i,t)$ and $x_n(j,t)$, $1 \leq t \leq 1359$ and, as a result, get a dataset of angle-distance pairs $(\theta_{i,j}, d_n(i,j))$, $1 \leq n \leq 100$, $1 \leq i, j \leq 31$, which we call $\mathcal{D}$:

$$\mathcal{D} = \{(\theta_{ij}, d_n(i,j)) \mid 1 \leq i, j \leq 31,\ 1 \leq n \leq 100\}. \quad (2)$$

We will then use this dataset in Section 2.5 to build a

[2] The normalized distance is more convenient than the information distance $d(x,y) = 2H(x,y) - H(y) - H(x)$, which is bounded by $\log_2 Q$.

[3] One may verify that the average and standard-deviation of pixels are not exactly uniform in the image.

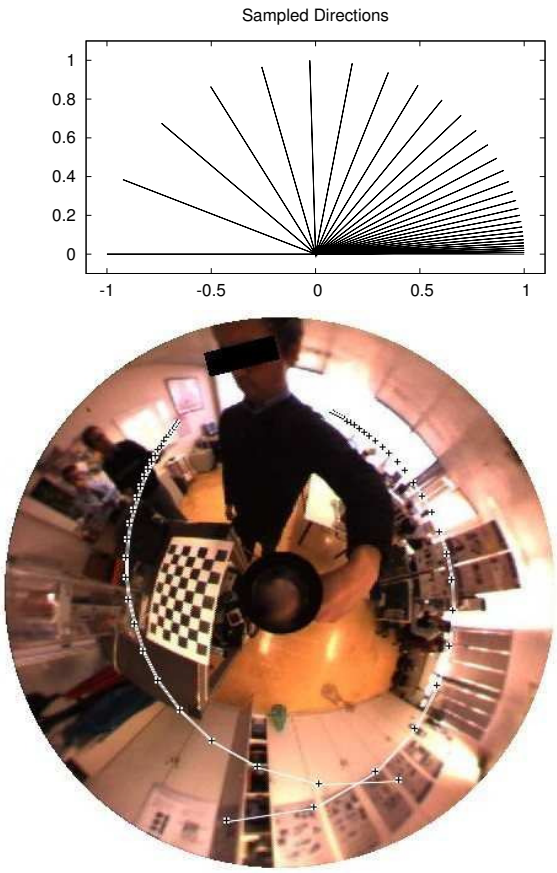


Figure 2. **Top:** Geometry of a discrete camera consisting of a planar array of thirty one (31) pixels, spanning 180° in the plane. The first two pixels are separated by 0.5°, the separation between consecutive photocells increases geometrically (ratio $\simeq 1.14$), so that the 31st photocell is antipodal with respect to the first.
Bottom: Two instances of the linear discrete camera, inserted in an omnidirectional image. Pixels locations are indicated by small crosses connected by white lines.

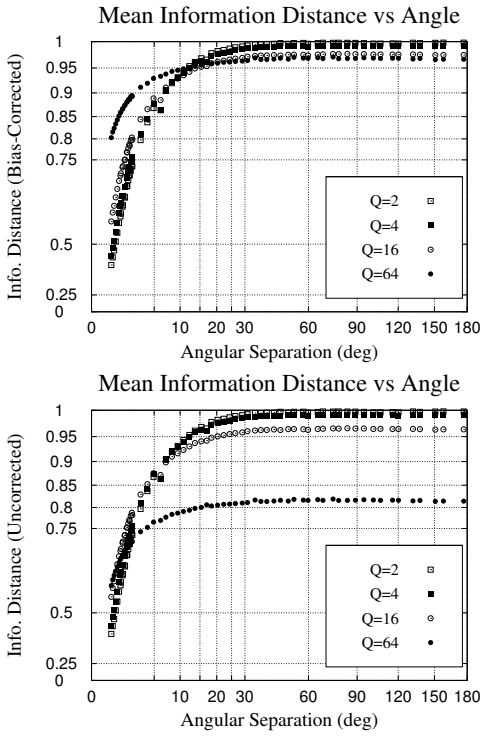


Figure 3. Dependency of the information distance $d(i,j)$ on the angular separation θ_{ij}. The height of each plotted point is the average of 100 estimates of the information distance between signals $x(i, 1 \leq t \leq 1359)$ and $x(j, 1 \leq t \leq 1359)$. The abscissa is the angular separations between pixels i and j of the discrete camera. Each curve corresponds to a different number Q of quantization bins. The quantization bins are all taken to have weight $1/Q$.
Top: Information distances computed with the first-order bias correction term of Eq. (3) removed. **Bottom:** Distances computed without correction.

functional relation that allows to estimate the angular separation θ_{ij} from the information distance $d(i,j)$.

2.4. Estimating the information distance

Caution should be taken when estimating the information distance (1) from finite samples $x(t)$, $y(t)$, $1 \leq t \leq T$: it is relatively common knowledge that replacing unknown probabilities $p_x(q)$ by sample frequencies $\hat{p_x}(q) = |\{t|x(t) = q\}|/Q$[4] in the expressions of the entropy results in biased estimates $\hat{H}(x)$, with expectancy

$$E\left\{\hat{H}\right\} = H - \frac{Q-1}{2T} + \frac{1 - \sum_q \frac{1}{p_x(q)}}{12T^2} + O\left(\frac{1}{T^3}\right). \quad (3)$$

[4] $|.|$ denotes the set cardinal.

This bias in turn causes a bias in the information distance estimates in Eq. (1). While it is easy to correct for the first bias term, $(Q-1)/2T$, correcting for the other terms is more delicate [17]. In the present work, we only correct the first bias term and we show below the benefits of doing so.

It should also be noted that the second term in Eq. (3) has a variance that increases with the number Q of quantization bins, so that it is also important to choose Q appropriately.

Figure 3 shows the average information distance Eq. (1) between the outputs of sensors with known angular separation. Each plot holds four curves, corresponding to $Q = 2$, 4, 16 and 64. The top plot uses bias correction, whereas the bottom plot does not.

The curves in the top plot are better grouped, showing that, when bias reduction is applied, the information distance estimates depend less on Q. This is important if one is interested in estimating properties of the information distance that depend on angular separation and on the ambient light field, rather than on Q and T.

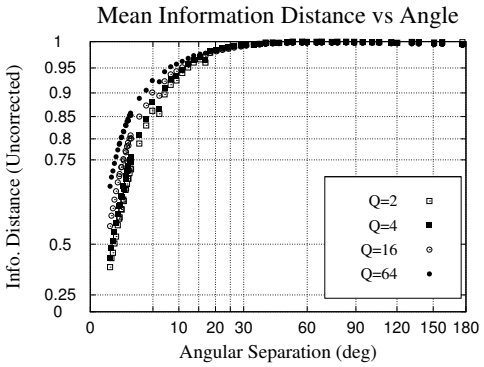


Figure 4. When the signal sample is longer as here, the information distance estimates are better behaved: here, a signal of length $T = 135900$ frames is obtained by concatenating the signals used in the previous experiment, and the information distance is computed between pairs of signals. Here, the information distance is computed without bias reduction. The values obtained with bias reduction are very similar.

Another way to reduce the importance of Q, when estimating information distance values, is to use longer signals. The concatenation of the 100 signal arrays, where all sensors have the same geometry, can be seen as a pseudo-signal that is 100 times longer, corresponding to approximately 90 minutes of data.

Figure 4 shows that, for long enough signals, the distance estimates are much better behaved, as they vary less on Q. However, one also sees that the estimates obtained for large Q require still longer signals. For angles smaller than $20°$, the number of quantization bins, Q still influences to a large extent the distance estimates.

The conclusion of these experiments is that adding the first correction term in Eq. (3) to entropy estimates in Eq. (1) is justified and beneficial, especially when only short sequences are available. One should note that correcting for bias in the entropy estimates is certainly less important in the situation of Olsson et al. [10], where the distance measurements are embedded directly in an abstract space. Obtaining accurate information distance estimates is important in our approach because our goal is to map information distance to angles. In the experiments below, we apply the bias-correction term to entropy estimates, and use $Q = 4$.

2.5. Estimating angular separation from information distance

Figures 3 and 4 also show one of the major issues that appear when one wishes to estimate the angular separation between pixels from the information distance estimated from their signals: for large angular separation, the information distance reaches a value close to one (corresponding to nearly independent signals) and remains nearly constant. In this section, we study the statistical properties of angle-distance pairs, with the goal of estimating the angle from the information distance.

In the absence of a model linking the information distance to the angular separation, we take a nonparametric approach. We estimate the probability density function of the angular separation knowing the information distance, $P(\theta \mid d)$ by applying standard Bayesian inference methods to our dataset of angle-distance pairs $\mathcal{D}$ of Eq. (2).

We cover the space of angles $[0, \pi]$ with 35 disjoint intervals $I_1, \ldots, I_{35}$ defined by angles[5] $0 < \theta_1 < \ldots < \theta_{34} < \pi$, and similarly cover the space of information distances $[0, 1]$ with 50 intervals $J_1, \ldots, J_{50}$, defined by uniformly spaced values $0.2 = d_1 < \ldots < d_{50} = 1$. We then call $C(p, q)$ the number of instances of angle-distance pairs in each rectangle: for any $1 \leq p \leq 35$ and $1 \leq q \leq 50$:

$$C(p, q) = |\{(n, i, j) \mid \theta_{ij} \in I_p \,,\; d_n(i, j) \in J_q\}| \,. \quad (4)$$

The expected value of $C(p, q)$ is approximately

$$E(C(p, q)) \simeq P(d \in J_q \mid \theta \in I_p) \, C_\theta(p) \,,$$

where $C_\theta(p) = \sum_q C(p, q)$ is the count of data points corresponding to $\theta_{ij} \in I_p$, and is known, since it is defined by the geometry of our probe (shown in Fig. 2, top). Knowing $C(p, q)$ thus provides an estimate of $P(d \in J_q \mid \theta \in I_p)$.

All that is now needed, to get an estimate of $P(\theta \in I_p \mid d \in J_q)$, is a prior for $P(\theta)$, which would be chosen in accordance to the knowledge about the angles that will be estimated. In our case, we simply take $P(\theta \in I_p) \propto C_\theta(p)$, which is appropriate to estimate angles between pixels of a sensor similar to that of Fig. 2. We thus use the estimate

$$P(\theta \in I_p \mid d \in J_q) \simeq \frac{C(p, q)}{C_d(q)}, \quad (5)$$

where $C_d(q) = \sum_p C(p, q)$.

Figure 5 shows the estimates of $P(\theta \in I_p \mid d \in J_q)$ graphically, as a graylevel image. The black curve shows the expected value of this conditional p.d.f., which we use as the estimate of θ knowing d.

In practice, for an information distance of $(d_{i-1} + d_i)/2$, the estimated angle is taken to be

$$\hat{\theta}((d_{q-1} + d_q)/2) = E(\theta \mid d \in J_q) \,,$$

and the estimated angle at intermediate values is obtained by affine interpolation from the nearest values.

Experimental validation

We now assess the quality of this estimate, by applying it to a set of information distances and comparing the estimated and true angles. In this experiment, the signals are obtained in the same conditions -same image sequence, same sensor shape, but different sensor orientations- as the dataset $\mathcal{D}$.

[5] Again, we choose angles with geometrically increasing spacing.

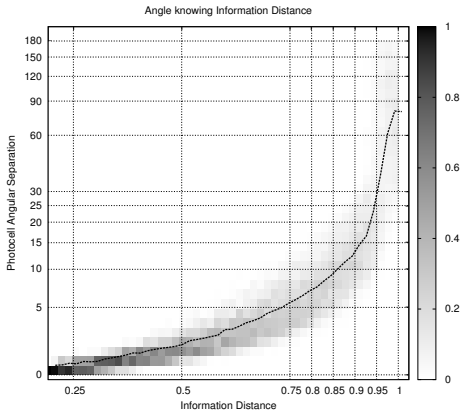


Figure 5. Probability density of the angular separation, when the information distance is known. Each column of pixels represents a conditional pdf $P(\theta \in I_p \mid d \in J_q)$ of the angle θ, knowing the information distance d. The curve superposed to the image is the expected value of the angle knowing the distance.

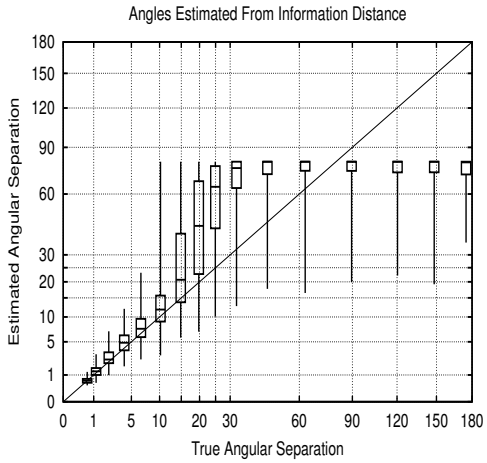


Figure 6. Statistical characterization of angles estimated from the (estimated) information distance between the signals produced by the pixels of a discrete camera. The angles are estimated by linear interpolation of the values of the non-parametric model described above. The first and third quartile define the top of each box. The line in each box is located at the median estimated angle, and the whiskers indicate the extrema.

Figure 6 shows that the estimated angles are fairly accurate for angular separations smaller than $5°$, but degrades sharply for greater values: around $10°$, a positive bias and an increase in variability appears; starting at $30°$, estimates "saturate" at the highest value in the non-parametric model.

3. Calibrating a discrete camera

Having seen how to estimate the angle between two pixels from the information distance between their respective data streams, we now show how useful this method is for estimating angles coming from different sensor shapes, and subsequently for estimating the whole sensor geometry.

3.1. Embedding points in the sphere

We now consider the problem:

Problem 1) Spherical embedding problem: Given angle estimates θ_{ij}, $1 \leq i, j \leq N$, find points X_i on the unit sphere, separated by angles approximately equal to θ_{ij}, i.e. $X_i^\top X_j \simeq \cos\theta_{ij}$, for all i, j.

This problem can be reduced to the classical problem of distance geometry [2]:

Problem 2) Euclidean embedding problem: Given distance estimates D_{ij}, $1 \leq i, j \leq N$, find points Y_i in a metric vector space, such that, for all i, j, $\|Y_i - Y_j\| \simeq D_{ij}$

Indeed, by defining an extra point $Y_0 = (0, 0, 0)$, and distances $D_{ij} = \sqrt{2 - 2\cos\theta_{ij}}$, the mapping of the first problem to the second is immediate. Moreover, Schoenberg´s classical algorithm for solving Problem 2, when applied to Problem 1, simplifies to the rank-3 approximation of the matrix C with terms $C_{ij} = \cos\theta_{ij}$. In order to solve Problem 1, we thus use the algorithm:

1. Compute the matrix C with terms $C_{ij} = \cos\theta_{ij}$, $1 \leq i, j \leq N$.
2. Compute, using the SVD decomposition, the rank-3 approximation $\tilde{C} = UU^\top$ of C, where U is $N \times 3$.
3. Define $X_i = (U_{i1}, U_{i2}, U_{i3}) \,/\, \|(U_{i1}, U_{i2}, U_{i3})\|$.

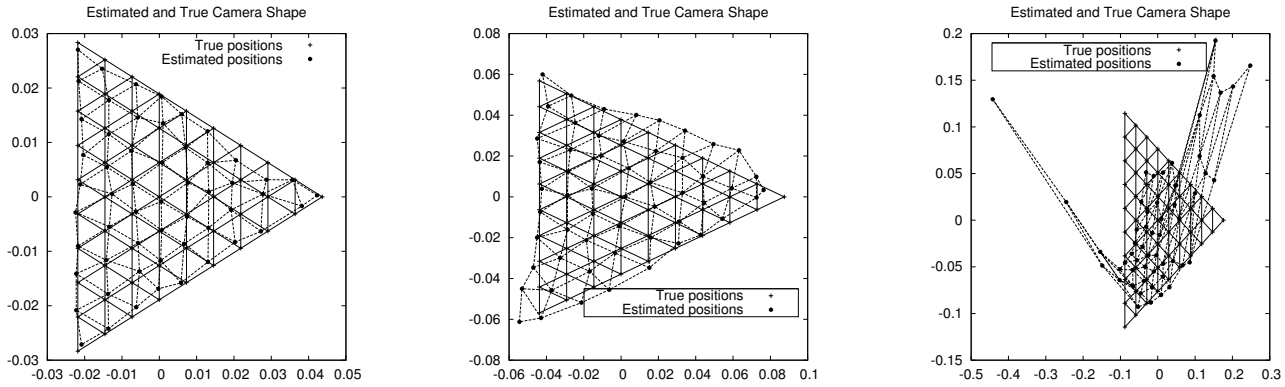
One should note that this very simple algorithm is not optimal in many ways. In particular, it does not take into account that the error in the angles θ_{ij} is greater in some cases than in others.

3.2. Sensor calibration

We now evaluate the results of this very simple algorithm on data produced by the angle-estimating method of Sec. 2.5. For this purpose, we produce sequences of pixel signals in the same conditions as in Sec. 2.3, except that the sensor shape is different. The information distance between pixels is then estimated from these signals, the angular separation between the pixels is estimated, and the embedding method of Sec. 3.1 is applied to these angular estimates.

Figure 7 shows the results of the reconstruction, when the pixels lie on a triangular grid of the image plane, scaled so that the angular diameter of the whole discrete camera is $5°$ (left), $10°$ (middle) or $20°$ (right). The lines connecting photocells have no purpose other than illustration.

These results are typical results of what researchers reproducing our method may encounter. It is easy to see that



(Sensor diameter: 5°) ((Sensor diameter: 10°) ((Sensor diameter: 20°)

Figure 7. True and estimated pixel layouts of a discrete camera. To estimate the layout, each pixel was sampled 1359 times. The information distance between pairs of pixels was estimated from these values, and converted to angular estimates using our non-parametric model. The angular distances were then embedded in the sphere. For visualization, they are aligned by the usual procrustes method, mapped to the plane by projective mapping with unit focal length, and line segments indicate the original pixel neighborhood relations.

for small-diameter discrete cameras, the estimated shape is pretty well determined by the information distance between the pixel streams, whereas, when the angular diameter increases, the noise in the estimated angles gradually overwhelms the embedding algorithm.

Figure 8 shows the results of the reconstruction, when the pixels lie on a 10×10 square grid of the image plane, scaled so that the angular diameter of the whole discrete camera is 5° (left) or 10° (right). This last grid is less good than typical, the usual outcome having an aspect closer to that of Fig. 7 left and middle.

These preliminary results -we have not performed extensive benchmarking of the complete calibration methodpoint towards the necessity of a more performant spherical embedding procedure, in order to better estimate the geometry of wide-angle nonparametric cameras. It is likely that an iterative refinement procedure would improve the final results. In particular, an embedding method is needed that can cope with unequal errors in the input. Such methods exist [18, 3] for Euclidean embedding, and could be adapted to the case of spherical embedding.

4. Discussion

We have presented the proof-of-concept for the estimation of the geometry of a discrete array of light-sensitive elements. The presented method is based on the exploitation of properties of the ambient light-field in a man-made environment. We raised some of the issues that are bound to be met by similar methods, in particular:

1. considering images defined on the sphere rather than on the plane,

2. estimating information distance from finite samples,

3. building a functional relationship that links information distance to angular separation, and

4. embedding angular estimates in a sphere.

On each of these points, some questions remain open for further work: first, what is the effect of individual photocells not having the same sensitivity, as is likely to happen in practice? what are the effects of simulating a sensor with an omnidirectional camera?

Concerning the estimation of information distance, it may be useful to estimate directly the bias in information distance estimates, rather than simply applying bias-correction to entropy estimates. In addition, our method assumes that consecutive samples are i.i.d.; it would probably be beneficial to acknowledge the temporal correlation of the signals.

Concerning image statistics, what are useful statistics of images defined on the sphere? Modeling images as observations of homogeneous random processes, while the world is not isotropic, is only correct if the orientation of the sensor follows a uniform distribution. A more appropriate model would probably acknowledge the existence of a dominant vertical (and perhaps forward) direction.

Concerning embedding, what is a practical way of embedding in the sphere angular estimates with greatly varying noise characteristics? Will one be able to adapt existing methods, or will new methods be necessary?

The empirical observation of the strong relation between angular separation and information distance is crucial to the presented algorithm. Can one build a physical basis for this relation, or at least link it to empirically known spectral properties of real-world images?

Can the present approach be generalized to a non-central camera? This would require determining statistical properties of the lightfield as a five-dimensional random field, and

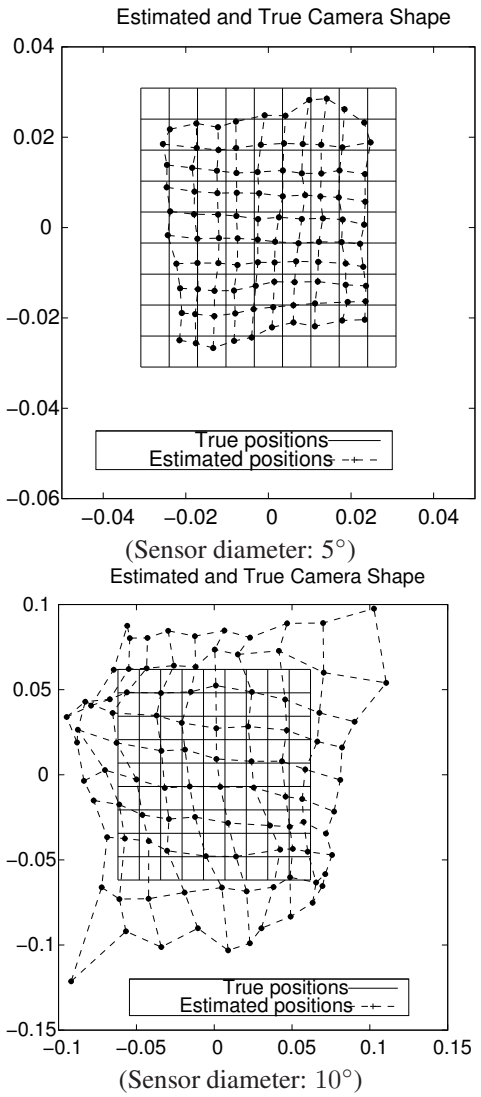


Figure 8. True and estimated pixel layouts of a discrete camera consisting of photocells lying on a rectangular grid. Apart from sensor shape, the experimental conditions are the same as in Fig. 7.

solving an altogether different embedding problem.

These questions are so basic that one can only guess at the amount of work required to produce a mature method for calibrating discrete cameras.

References

[1] J. P. Crutchfield. Information and its metric. In L. Lam and H. C. Morris, editors, *Nonlinear Structures in Physical Systems–Pattern Formation, Chaos and Waves*, pages 119–130. Springer-Verlag, 1990.

[2] J. Dattorro. *Convex Optimization & Euclidean Distance Geometry*. Meboo Publishing, 2005.

[3] P. Drineas, A. Javed, M. Magdon-Ismail, G. Pandurangan, R. Virrankoski, and A. Savvides. Distance matrix reconstruction from incomplete distance information for sensor network localization. In *Third Annual IEEE Communications Society Conference on Sensor Mesh, and Ad Hoc Communications and Networks (SECON)*, 2006.

[4] R. Fergus, A. Torralba, and W. T. Freeman. Random lens imaging. Technical Report MIT CSAIL TR 2006-058, Massachusetts Institute of Technology, 2006.

[5] W. T. Freeman, E. C. Pasztor, and O. T. Carmichael. Learning low-level vision. *International Journal of Computer Vision*, 40(1):25–47, 2000.

[6] C. Geyer and K. Daniilidis. A unifying theory for central panoramic systems and practical applications. In *proc. ECCV*, volume II, pages 445–461. Springer-Verlag, 2000.

[7] E. Grossmann, E.-J. Lee, P. Hislop, D. Nistér, and H. Stewénius. Are two rotational flows sufficient to calibrate a smooth non-parametric sensor? In *proc. IEEE CVPR*, 2006.

[8] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2000.

[9] W. J. Krzanowski. *Principles of Multivariate Analysis: A User's Perspective*. Clarendon Press, Statistical Science Series, 1988.

[10] D. P. Lars Olsson, Chrystopher L. Nehaniv. Sensory channel grouping and structure from uninterpreted sensor data. In *NASA/NoD Conference on Evolvable Hardware*, 2004.

[11] D. Nistér, H. Stewenius, and E. Grossmann. Non-parametric self-calibration. In *proc. ICCV*, 2005.

[12] D. Pierce and B. Kuipers. Map learning with uninterpreted sensors and effectors. *Artificial Intelligence Journal*, 92(169-229), 1997.

[13] B. Potetz and T. S. Lee. Scaling laws in natural scenes and the inference of 3d shape. In *NIPS – Advances in Neural Information Processing Systems*, pages 1089–1096. MIT Press, 2006.

[14] S. Ramalingam, P. Sturm, and S. Lodha. Towards complete generic camera calibration. In *Proc. CVPR*, volume 1, pages 1093–1098, 2005.

[15] R. Roy. Spectral analysis for a random process on the sphere. *Annals of the institute of statistical mathematics*, 28(1), 1976.

[16] G. Schmidt and D. T. Moore. Tapered gradient index microlenses for compound lens arrays. In *Proc. SPIE Vol. 6342, International Optical Design Conference*, 2006.

[17] T. Schuermann. Bias analysis in entropy estimation. *J. Phys. A: Math. Gen*, 37:L295–L301, 2004. arXiv:cond-mat/0403192v3.

[18] Y. Shang, W. Ruml, Y. Zhang, and M. P. J. Fromherz. Localization from mere connectivity. In *MobiHoc '03: Proc. ACM Intl. Symp. on Mobile Ad Hoc Networking & Computing*, pages 201–212. ACM Press, 2003.

[19] A. Torralba and A. Oliva. Statistics of natural image categories. *Network: Computation in Neural Systems*, 14:391–412, 2003.

[20] R. Tsai. An efficient and accurate camera calibration technique for 3D machine vision. In *IEEE Conf. on Computer Vision and Pattern Recognition*, 1986.

[21] N. W. Y, S.-C. Zhu, and C.-. Guo. From information scaling of natural images to regimes of statistical models. Technical Report 2004010111, Department of Statistics, UCLA, 2004.