

## Spectral Solution of Large-Scale Extrinsic Camera Calibration as a Graph Embedding Problem

Matthew Brand      Matthew Antone\*      Seth Teller†

TR-2004-100    May 2004

### Abstract

Extrinsic calibration of large-scale ad hoc networks of cameras is posed as the following problem: Calculate the locations of  $N$  mobile, rotationally aligned cameras distributed over an urban region, subsets of which view some common environmental features. We show that this leads to a novel class of graph embedding problems that admit closed-form solutions in linear time via partial spectral decomposition of a quadratic form. The minimum squared error (MSE) solution determines locations of cameras and/or features in any number of dimensions. The spectrum also indicates insufficiently constrained problems, which can be decomposed into well-constrained rigid subproblems and analyzed to determine useful new views for missing constraints. We demonstrate the method with large networks of mobile cameras distributed over an urban environment, using directional constraints that have been extracted automatically from commonly viewed features. Spectral solutions yield layouts that are consistent in some cases to a fraction of a millimeter, substantially improving the state of the art. Global layout of large camera networks can be computed in a fraction of a second.

This work may not be copied or reproduced in whole or in part for any commercial purpose. Permission to copy in whole or in part without payment of fee is granted for nonprofit educational and research purposes provided that all such whole or partial copies include the following: a notice that such copying is by permission of Mitsubishi Electric Research Laboratories, Inc.; an acknowledgment of the authors and individual contributions to the work; and all applicable portions of the copyright notice. Copying, reproduction, or republishing for any other purpose shall require a license with payment of fee to Mitsubishi Electric Research Laboratories, Inc. All rights reserved.

Copyright © Mitsubishi Electric Research Laboratories, Inc., 2004  
201 Broadway, Cambridge, Massachusetts 02139

European Conference on Computer Vision (ECCV)



# Spectral solution of large-scale extrinsic camera calibration as a graph embedding problem

Matthew Brand<sup>1</sup>, Matthew Antone<sup>2</sup>, and Seth Teller<sup>3</sup>

<sup>1</sup> Mitsubishi Electric Research Labs, 201 Broadway, Cambridge MA 02139 USA,  
<http://www.merl.com/people/brand/>

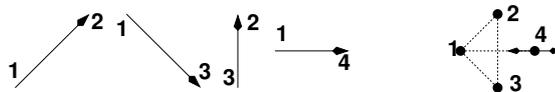
<sup>2</sup> AlphaTech, 6 New England Executive Park, Burlington MA 01803 USA

<sup>3</sup> MIT, 77 Massachusetts Avenue, Cambridge MA 02139 USA

**Abstract.** Extrinsic calibration of large-scale ad hoc networks of cameras is posed as the following problem: Calculate the locations of  $N$  mobile, rotationally aligned cameras distributed over an urban region, subsets of which view some common environmental features. We show that this leads to a novel class of graph embedding problems that admit closed-form solutions in linear time via partial spectral decomposition of a quadratic form. The minimum squared error (MSE) solution determines locations of cameras and/or features in any number of dimensions. The spectrum also indicates insufficiently constrained problems, which can be decomposed into well-constrained rigid subproblems and analyzed to determine useful new views for missing constraints. We demonstrate the method with large networks of mobile cameras distributed over an urban environment, using directional constraints that have been extracted automatically from commonly viewed features. Spectral solutions yield layouts that are consistent in some cases to a fraction of a millimeter, substantially improving the state of the art. Global layout of large camera networks can be computed in a fraction of a second.

## 1 Introduction

Consider a set of images taken from a large number of viewpoints distributed over a broad area such as a city. The source might be a network of security cameras, one or more tourists with cameras, or the collected uploads of camera-enabled mobile phones in a neighborhood. Knowledge of camera positions will be useful for 3D-reconstruction of the environment, tracing the tourist's path, and offering location-aware services to the



**Fig. 1.** A toy example of the embedding problem. The set of directional constraints at left must be assembled into a maximally consistent graph. The solution (at right) may have degrees of freedom. E.g., node 4 is only partially constrained. In 2D the problem is trivial. In higher dimensions the constraint set may be simultaneously inconsistent, overconstrained, and underconstrained. Our method characterizes and calculates the space of all solutions.

---

phone users. We seek to infer these viewpoints from the image set. This paper considers the sparse case, where most viewpoint pairs have nothing in common, but small subsets of cameras view some common environmental features. E.g., in a city, buildings and other large occluders ensure that most images contain local features only.

Antone and Teller [1] demonstrated that it is possible to discover a sparse set of feature correspondences in image sets of this nature. The correspondences can be found from random search or from rough indicators of co-location, e.g., image histogram matches, assignments to wireless communication cells, or global positioning system (GPS) data. Given two internally calibrated cameras that view sufficiently many common scene features, it is possible to determine the *relative* geometry of the cameras, up to scale and orientation [2,3]. Antone and Teller further demonstrated that the *global* orientation of a set of partially overlapped views can be determined from an analysis of feature correspondences and vanishing points in the images [1]. Therefore we have directional information about some of the vectors connecting viewpoints to features (or viewpoints to nearby viewpoints), but no information about distances or locations. This paper introduces a fast spectral method for making a global assignment of viewpoints and features to 3D locations from these sparse directional constraints.

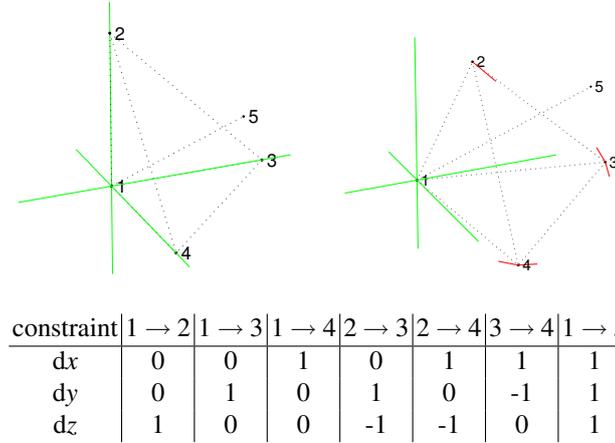
### 1.1 Related work in computer vision

The subject of extrinsic camera calibration has been treated broadly in the computer vision literature [4,2,3]. We review here methods that have decoupled rotational and translational degrees of freedom (DOFs), or have been demonstrated for large or uncertain inputs.

Several researchers have factored the 6-DOF extrinsic calibration problem in order to reduce the number of parameters to be simultaneously estimated. Both interactive [5,6,7] and automated [8,9] methods exist, and have been demonstrated for relatively small numbers of images. Interactive methods do not scale effectively, and are vulnerable to operator error and numerical instability. Projective techniques [10,11] recover structure and pose only up to an arbitrary projective transformation. Other researchers have described structure-from-motion methods using singular value decomposition [12] or random search [13]. Most of these methods contemplate a much richer set of correspondences than available in our problem. Antone and Teller proposed an iterative algorithm for extrinsic calibration of networks of omni-directional images, using iterative least-squares [1]; each iteration takes time linear in the number of constraints. Section 4 benchmarks our method against theirs.

### 1.2 Related work in graph embeddings

The problem we treat is a particularly well-specified graph embedding problem, and as such bears relation to barycentric embedding proofs of Tutte [14,15] and their modern-day descendent, the locally linear embedding algorithm of Roweis and Saul [16]. These methods constrain node locations to be linear mixtures of their neighbor's locations, with known mixture weights; our method constrains node locations to be linear mixtures of rays emanating from their neighbors, with mixture weights unknown. As it turns out, our solution method has a novel algebraic structure; in contrast to all prior spectral embedding methods, the solution is specified by a single eigenvector, regardless of the dimensionality of the embedding.



**Fig. 2.** A simple embedding problem in  $\mathbb{R}^3$ . The table of constraints yields a perfect embedding, shown at left with node 1 at the origin, nodes 2-4 on each axis, and node 5 free to slide along a line passing through the middle of the face  $\triangle 234$  and the origin. If the constraints are made inconsistent (e.g., by changing the first directional constraint  $1 \rightarrow 2$  to  $[1, 1, 2]^\top$ ) the least-squares embedding (at right) spreads error evenly over the inconsistent constraints. Node 5 is still free to slide. The small quivers emanating from nodes 2,3,4 show the distances from these nodes to the ray constraints that involve them; the minimal (nontranslational) eigenvalue of  $\mathbf{H}_{\mathcal{E}}$  sums those squared distances.

## 2 Directionally constrained embeddings

The directionally constrained embedding problem is formally posed as follows: We are given a set of  $N$  nodes  $\mathbf{x}_i \in \mathbb{R}^d$  and an incomplete specification of node-to-node directions  $\mathbf{d}_{ij} \in \mathbb{R}^d$  for some  $i, j \in [1, N] \subset \mathbb{N}, i \neq j$ . The signs and lengths of the true node-to-node displacements are unknown. What is the set of consistent embeddings in  $\mathbb{R}^d$ ? Is the embedding uniquely determined (up to scale and translation)?

We develop a spectral solution based on the thin eigenvalue decomposition (EVD). Let embedding matrix  $\mathbf{X} \doteq [\mathbf{x}_1, \dots, \mathbf{x}_N] \in \mathbb{R}^{d \times N}$  contain the location  $\mathbf{x}_i \in \mathbb{R}^d$  of the  $i$ th node in column  $i$ . We seek an embedding where node-to-node displacements  $(\mathbf{x}_j - \mathbf{x}_i)$  are maximally consistent with the constraint directions  $\mathbf{d}_{ij}$ .

### 2.1 Maximum covariance spectral solution

To start simply, first consider maximizing the squared length of the projections of the displacements onto the constraints:

$$\mathcal{C}(\mathbf{X}) \doteq \sum_{ij \in \text{constrained nodes}} \|(\mathbf{x}_i - \mathbf{x}_j)^\top \mathbf{d}_{ij}\|^2, \quad (1)$$

where  $\|\mathbf{X}\|$  denotes Frobenius (Euclidean) norm. Clearly  $\mathcal{C}$  is a quadratic form and therefore the problem is convex. To maximize  $\mathcal{C}$ , consider the vertical concatenation of

$N$  embedding vectors  $\mathbf{y} \doteq \text{vec } \mathbf{X} = [\mathbf{x}_1^\top, \mathbf{x}_2^\top, \dots, \mathbf{x}_N^\top]^\top \in \mathbb{R}^{dN \times 1}$  and the symmetric matrix

$$\mathbf{H}_C \doteq \left[ \bowtie_i \sum_j (\mathbf{d}_{ij} \mathbf{d}_{ij}^\top + \mathbf{d}_{ji} \mathbf{d}_{ji}^\top) \right] - \left[ \Downarrow_i \left[ \Rightarrow_j (\mathbf{d}_{ij} \mathbf{d}_{ij}^\top + \mathbf{d}_{ji} \mathbf{d}_{ji}^\top) \right] \right] \in \mathbb{R}^{dN \times dN} \quad (2)$$

where  $\Downarrow, \Rightarrow, \bowtie$  denote vertical, horizontal, and diagonal concatenation, respectively.

**Proposition 1.** *The maximizing eigenvector of  $\mathbf{H}_C$  determines the embedding  $\mathbf{y}$  that maximizes  $C$  up to sign and scale.*

*Proof.* By construction,

$$\begin{aligned} C &= \sum_{ij} (\mathbf{x}_i - \mathbf{x}_j)^\top \mathbf{d}_{ij} \mathbf{d}_{ij}^\top (\mathbf{x}_i - \mathbf{x}_j) \\ &= \sum_{ij} \mathbf{x}_i^\top (\mathbf{d}_{ij} \mathbf{d}_{ij}^\top) \mathbf{x}_i + \mathbf{x}_j^\top (\mathbf{d}_{ij} \mathbf{d}_{ij}^\top) \mathbf{x}_j - \mathbf{x}_i^\top (\mathbf{d}_{ij} \mathbf{d}_{ij}^\top) \mathbf{x}_j - \mathbf{x}_j^\top (\mathbf{d}_{ij} \mathbf{d}_{ij}^\top) \mathbf{x}_i \\ &= \mathbf{y}^\top \mathbf{H}_C \mathbf{y}. \end{aligned}$$

By the Schmidt-Eckart-Young theorem, the maximum of quadratic form  $(\mathbf{y}^\top \mathbf{H}_C \mathbf{y}) / (\mathbf{y}^\top \mathbf{y})$  is the largest eigenvalue  $\lambda_{\max}(\mathbf{H}_C)$ , attained at the corresponding eigenvector  $\mathbf{y} = \mathbf{v}_{\max}(\mathbf{H}_C)$ .

The optimal embedding is therefore  $\mathbf{X} = \mathbf{y}^{(d)} = \mathbf{v}_{\max}^{(d)}$ , an order- $d$  vector-transpose that reshapes vector  $\mathbf{y} \in \mathbb{R}^{dN}$  into matrix  $\mathbf{X} \in \mathbb{R}^{d \times N}$ .  $\blacksquare$

*Remark 1.* The norm of any directional vector  $\|\mathbf{d}_{ij}\|$  determines how strongly it constrains the final solution; if  $\|\mathbf{d}_{ij}\| = 0$  the constraint is removed from the problem. Thus we may entertain sparse, weighted constraints.

## 2.2 Minimum squared-error (MSE) spectral solution

Maximum covariance problems tend to favor solutions in which large displacements are directionally accurate, sometimes at the expense of directionally inaccurate short displacements. A minimum squared-error framework is preferable, because it spreads error evenly over the solution, and guarantees an exact solution when allowed by the constraints. To obtain the MSE solution, we minimize the components of the displacements that are orthogonal to the desired directions. The error function is

$$\mathcal{E}(\mathbf{X}) \doteq \sum_{ij} (\|(\mathbf{x}_i - \mathbf{x}_j)^\top \mathbf{d}_{ij}^\perp\| \cdot \|\mathbf{d}_{ij}\|)^2, \quad (3)$$

where  $\mathbf{d}_{ij}^\perp$  is an orthonormal basis of the null-space of  $\mathbf{d}_{ij}$ . One may visualize each constraint  $\mathbf{d}_{ij}$  as a ray emanating from the node  $i$  or  $j$ ;  $\mathcal{E}$  sums, over all nodes, the squared distance from a node to each ray on which it should lie (scaled by the length of the ray). The constraints are all weighted equally when all  $\mathbf{d}_{ij}$  have the same norm. Of course, if the constraints admit an errorless embedding, it will be invariant to any nonzero rescaling of the constraints. To minimize  $\mathcal{E}$ , let  $\mathbf{H}_{\mathcal{E}}$  be constructed as  $\mathbf{H}_C$ , except that  $(\mathbf{d}_{ij} \mathbf{d}_{ij}^\top)$  is replaced everywhere in equation 2 with

$$(\mathbf{d}_{ij}^\top \mathbf{d}_{ij}) \cdot \mathbf{I} - \mathbf{d}_{ij} \mathbf{d}_{ij}^\top. \quad (4)$$

**Proposition 2.** *The minimizing eigenvector of  $\mathbf{H}_{\mathcal{E}}$  in the space orthogonal to  $\mathbf{1}_{[N \times 1]} \otimes \mathbf{I}_{[d \times d]}$  determines the nondegenerate embedding  $\mathbf{y}$  that minimizes  $\mathcal{E}$  up to sign and scale.*

*Proof.* Following the previous proof,  $\mathcal{E}(\mathbf{y}) = \mathbf{y}^\top \mathbf{H}_{\mathcal{E}} \mathbf{y}$  and is minimized by eigenpair  $\{\mathcal{E}(\mathbf{y}^{(d)}) = \lambda_{\min}(\mathbf{H}_{\mathcal{E}}), \mathbf{y} = \mathbf{v}_{\min}(\mathbf{H}_{\mathcal{E}})\}$  because

$$(\mathbf{d}_{ij}^\top \mathbf{d}_{ij}) \cdot \mathbf{I} - \mathbf{d}_{ij} \mathbf{d}_{ij}^\top = (\mathbf{d}_{ij}^\top \mathbf{d}_{ij}) \cdot (\mathbf{I} - \mathbf{d}_{ij} \mathbf{d}_{ij}^\top / (\mathbf{d}_{ij}^\top \mathbf{d}_{ij})) = (\mathbf{d}_{ij}^\perp) \|\mathbf{d}_{ij}\|^2 (\mathbf{d}_{ij}^\perp)^\top,$$

a scaled orthogonal projector that isolates the component of  $\mathbf{x}_i - \mathbf{x}_j$  that is orthogonal to  $\mathbf{d}_{ij}$  and scales it by  $\|\mathbf{d}_{ij}\|^2$ . The  $\mathbf{1} \otimes \mathbf{I}$  constraint arises because the directional constraints are trivially satisfied by mapping all nodes to a single point in  $\mathbb{R}^d$ . This implies that the nullspace of  $\mathbf{H}_{\mathcal{E}}$  contains  $d$  nuisance eigenvectors that give an orthogonal basis for locating the point anywhere in  $\mathbb{R}^d$ . Algebraically, that basis is spanned by  $\mathbf{1}_{[N \times 1]} \otimes \mathbf{I}_{[d \times d]}$ , because  $\mathbf{H}_{\mathcal{E}}(\mathbf{1} \otimes \mathbf{I}) = \mathbf{0}_{[Nd \times d]}$ . (This also gives an orthogonal basis for translating the solution in the embedding space). Therefore the nondegenerate embedding must be in the (possibly approximate) nullspace of  $\mathbf{H}_{\mathcal{E}}$  that is orthogonal to  $\mathbf{1} \otimes \mathbf{I}$ . ■

Appendix A outlines fast and stable methods for computing  $\mathbf{y}$  in linear time.

*Remark 2.* Solutions are determined up to sign and scale ( $\mathcal{E}(k\mathbf{X}) = k^2 \mathcal{E}(-\mathbf{X})$  for  $k \in \mathbb{R}^+$ ). This compares favorably to current methods of spectral graph embeddings, where a  $d$ -dimensional embedding requires  $d$  eigenvectors and is determined only up to arbitrary affine or orthogonal transforms in  $\mathbb{R}^d$ .

*Remark 3. (Uncertain constraints)* If a direction is uncertain, one may replace vector  $\mathbf{d}_{ij}$  with a matrix  $\mathbf{D}_{ij}$  whose orthogonal columns are each scaled by the certainty that the constraint lies in that direction. The outer product  $\Sigma_{ij} \doteq \mathbf{D}_{ij} \mathbf{D}_{ij}^\top$  is effectively the covariance of a Gaussian distribution over possible directions. The associated scaled orthogonal projector is  $\|\Sigma_{ij}\|_2 \cdot \mathbf{I} - \Sigma_{ij}$ . If the columns of  $\mathbf{D}_{ij}$  are unscaled (unit norm), the directional constraint is simply weakened to a subspace constraint, i.e., that  $\mathbf{x}_j - \mathbf{x}_i$  lie as close as possible to the subspace spanned by  $\mathbf{D}_{ij}$ .

### 3 Problem pathologies

Although the spectral solution is MSE-optimal w.r.t. the constraints, in vision problems the constraints themselves are derived from data and thus may be problematic. Therefore additional tools are needed to detect and resolve ill-posed problems where the constraint data is insufficient or inconsistent.

#### 3.1 Underconstrained problems

A problem is underconstrained when (1) the connectivity graph is disconnected, allowing two partial embeddings that can rigidly transform in each other's coordinate frame, or when (2) all constraints on a node are collinear, allowing it to slide along any one directional constraint. Both cases will manifest themselves as multiple (near-) zero eigenvalues in the spectrum of  $\mathbf{H}_{\mathcal{E}}$ ; rotating any two eigenvectors in this (approximate) nullspace will animate one such undesired degree of freedom (DOF), giving an *orbit* of solutions. In this way eigenvalue multiplicity diagnoses the dimensionality of the subspace of solutions. However, multiplicity is not 1-to-1 with excess DOFs; an orbit may be redundantly expressed in  $d$  eigenvectors, each giving different dynamics for varying

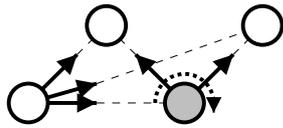
the same positional DOF. Some further analysis is needed to characterize the intrinsic DOFs in the problem specification.

When a solution has many DOFs, it is useful to cluster nodes that articulate together. Intuitively, two nodes articulate together if they both have nonzero entries in an eigenvector associated with a zero eigenvalue. Let us call the collection of all such eigenvectors the *DOF matrix*. The optimal clustering is given by the low-dimensional orthogonal *binary* (0/1-valued) basis that best approximates the DOF matrix. Finding such a basis is known to be NP-hard, so we take recourse in continuous relaxations of this problem, e.g., a spectral clustering of the row-vectors of the DOF matrix, or an independent components analysis of its columns. The former seeks groupings whose motions are most decorrelated; the latter seeks full statistical independence, a stronger condition. An example is given below in figure 5. By identifying nonrigidities in the solution, this analysis provides a useful basis for deciding which nodes need additional constraints, e.g., in the form of additional nearby views and associated baseline directions.

### 3.2 Problems admitting “negative lengths”

Because  $\mathcal{E}(-\mathbf{X}) = \mathcal{E}(\mathbf{X})$ , the projection of a node-to-node displacement onto the desired direction,  $(\mathbf{x}_j - \mathbf{x}_i)^\top \mathbf{d}_{ij}$ , can be positive *or* negative. In general, perfectly constrained problems will admit solutions only where projections are either all positive or all negative. But problems that are under constrained or that have inconsistent constraints may have MSE solutions with some projections of varied sign. An all-positive solution can be found via quadratic programming (QP). The QP problem has an elegant statement using the eigenvalue decomposition  $\mathbf{H}_{\mathcal{E}} \rightarrow \mathbf{V} \text{diag}(\lambda) \mathbf{V}^\top$ : We seek a nonzero vector  $\mathbf{m} \neq \mathbf{0}$  that mixes the eigenvectors  $\mathbf{y} = \mathbf{V}\mathbf{m}$  to incur minimal squared error  $\mathcal{E}_{\text{QP}} \doteq \mathbf{m}^\top \text{diag}(\lambda) \mathbf{m} = \mathcal{E}(\mathbf{X})$  while keeping all lengths positive. Formally, the QP problem is stated: Minimize  $\mathcal{E}_{\text{QP}}$  subject to  $\mathbf{m}^\top \mathbf{V}^\top [\Rightarrow_{ij} \text{vec} \mathbf{K}_{ij}] \geq \mathbf{1}$ , where each  $\mathbf{K}_{ij} \in \mathbb{R}^{d \times N}$  is all zeros except for its  $i$ th and  $j$ th columns, which are  $\mp \mathbf{d}_{ij}$ , respectively. Scaling the resulting  $\mathbf{m}$  to unit norm is equivalent to performing constrained optimization on a hypersphere. Since eigenvectors paired to large eigenvalues are unlikely to be used, we restrict the problem to consider only the smallest eigenvalue/vector pairs by truncating  $\mathbf{V}$  and  $\lambda$ . This yields a very small quadratic programming problem.

### 3.3 Problems having misaligned (rotated) nodes



**Fig. 3.** A set of directional constraints (solid arrows) that has a well-constrained embedding for every rotation of the shaded node. The two upper nodes simply slide along the dashed lines.

In the *alignment problem* the  $i$ th node’s constraints  $\{\mathbf{d}_{ij}\}_j$  are perturbed by a rotation  $\mathbf{R}_i^\top$  which we want to identify and remove prior to embedding. Assuming that the directional constraints are consistent with observation data, such perturbations will only be detectable as sources of error in the embedding. Sadly there exist some well-posed embedding problems where rotations induce no errors; see figure 3. Even when rotations do produce error, the alignment problem is

almost certainly not convex because the error function is quartic in the rotation parameters (ignoring orthonormality constraints). However, if the perturbations are small

and error-producing, it is reasonable to expect that the error function is approximately convex in the neighborhood of the optimal solution, and that the MSE embedding is only mildly distorted in this neighborhood. Thus an alternating least-squares procedure that computes embeddings from rotated constraints and rotations from embeddings will iterate to a set of constraints that admits a lower-error embedding. Strictly speaking, this is guaranteed only if one rotation is updated per iteration, but for small perturbations we find that the error declines monotonically when all rotations are updated *en masse*. Solution for the optimal rotation is a Procrustes alignment problem: Collect normalized directional constraints in matrix  $\mathbf{A}_i \doteq [\Rightarrow_j \mathbf{d}_{ij} / \|\mathbf{d}_{ij}\|]$  and normalized embedding displacements in matrix  $\mathbf{B}_i \doteq [\Rightarrow_j (\mathbf{x}_j - \mathbf{x}_i) / \|\mathbf{x}_j - \mathbf{x}_i\|]$ . The optimal aligning rotation  $\mathbf{R}_i = \arg \min_{\mathbf{R} \in \mathbb{R}^{d \times d} | \mathbf{R}^\top \mathbf{R} = \mathbf{I}} \|\mathbf{R} \mathbf{A}_i - \mathbf{B}_i\|_F$  is  $\mathbf{R}_i = \mathbf{V} \mathbf{U}^\top$  from the singular value decomposition  $\mathbf{U} \text{diag}(\mathbf{s}) \mathbf{V}^\top = \mathbf{A}_i \mathbf{B}_i^\top$ . Normalization prevents potential errors due to incorrect displacement lengths.

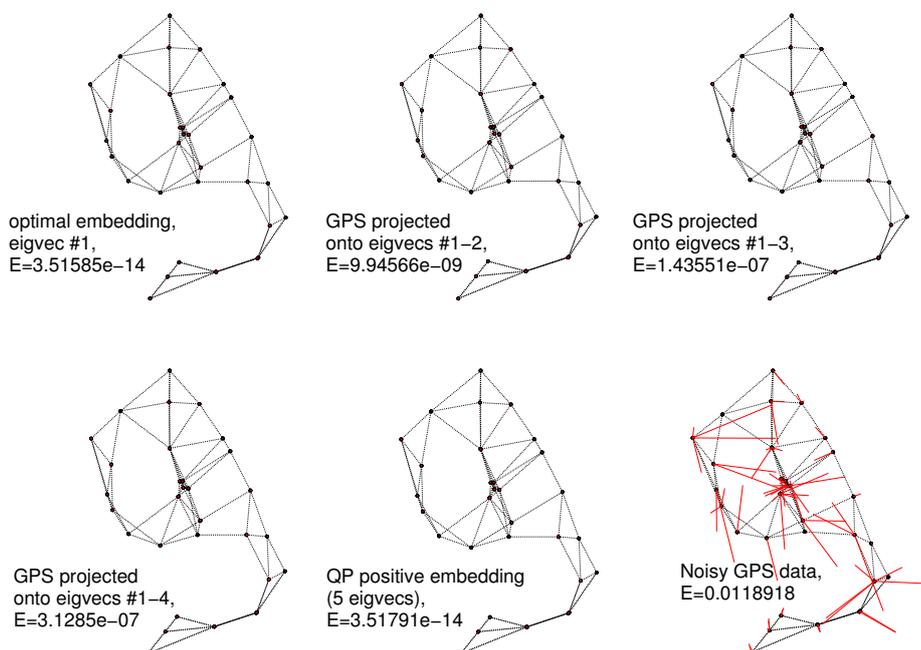
## 4 Example calibrations of camera networks

The least-squares spectral solution can be implemented in less than fifteen lines of Matlab code. We first illustrate its properties with a simple problem in  $\mathbb{R}^3$ . Figure 2 shows a perfectly embeddable problem and how the embedding changes when the constraints are made inconsistent. Both problems have a single underconstrained node, so  $\lambda_{\min}(\mathbf{H}_{\mathcal{E}})$  has multiplicity two (ignoring translational DOFs). Rotating the associated eigenvectors causes this node to slide back and forth on its constraint ray while the rest of the solution changes scale to accommodate the algebraic constraint  $\|\mathbf{y}\| = 1$ .

To assess this method’s usefulness on real data with objective performance metrics, we obtained data derived from hundreds of cameras scattered over a university campus [1] and posted to <http://city.lcs.mit.edu>. The datasets consist of 3D directional constraints that were obtained by triangulating pairs of cameras against commonly viewed features of buildings and an analysis of vanishing points in images. The triangulations alone do not give a complete or consistent calibration. The cameras also have rough estimates of ground truth from 2.5-meter-accurate GPS sensors. Other than the noisy GPS data, there is no ground truth for this data. Therefore, we use the self-consistency measures proposed by Antone and Teller to assess the quality of the spectral embeddings as extrinsic camera registrations. Following [1], we report the 3D error of node positions with respect to the directional constraints that apply to them (distance to constraint rays, in millimeters), and the 2D orientation error of node-to-node displacements (angle between displacement and constraint ray, in degrees).

### 4.1 Green court

The Green Court dataset consists of 32 nodes and 80 directional constraints spanning an area of roughly 80 by 115 meters. The algorithm in [1] recovered global position consistent on average to within 45 millimeters. The maximum position error for any node was 81mm. Reported CPU time of a C implementation was roughly one hour, of which the final layout phase took “a matter of seconds” (personal communication).

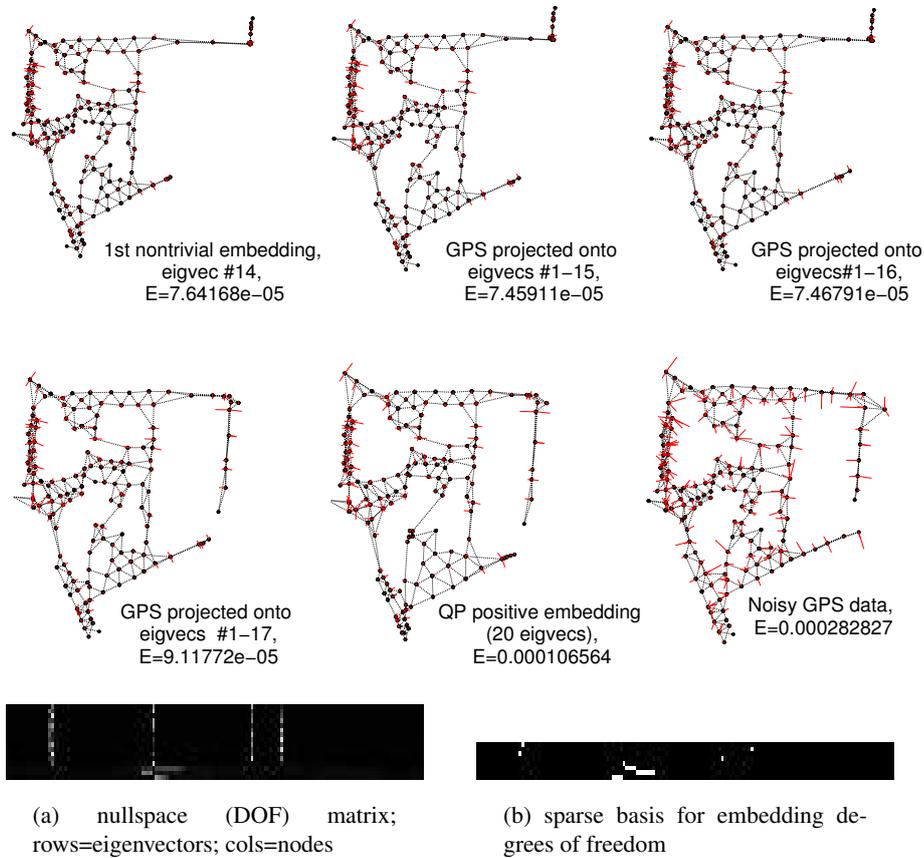


**Fig. 4.** Spectral embedding of the Green Court data, viewed from above. Dotted graph edges indicate problem constraints; each **quiver** represents the unsatisfied component of a constraint, magnified  $10\times$ . The rightmost bottom graph shows that GPS errors on the scale of a few meters are being corrected by visually determined constraints. The large holes are building footprints. In this case, the raw spectral solution has strictly positive edge lengths and is thus identical to the QP solution. Note that the spectral embedding has much lower residual ( $\mathcal{E}$ ) than the GPS data. Projecting the GPS data onto low-error embedding subspaces is analogous to data denoising via PCA.

Our Matlab implementation computed the optimal (minimum sum-squared error) solution in roughly 1/4 second, reducing consistency errors by roughly four orders of magnitude (see table 1). Figure 4 shows the optimal embedding using the minimizing eigenvector and several embeddings obtained by projecting the GPS data onto low-order eigenvectors of  $\mathbf{H}$ , a form of data denoising analogous to that offered by principal components analysis.

## 4.2 Ames Court

The Ames Court dataset consists of 158 nodes and 355 directional constraints spanning roughly 315 by 380 meters. Antone and Teller report solving a 100-node subset of this problem in which all nodes are properly constrained [1]. They recovered global position consistent on average to within 57mm. The maximum pose inconsistency was 88mm. Reported total CPU time was roughly four hours.



**Fig. 5.** Top: Spectral embedding of the Ames Court data, illustrated as in figure 4. Note how the QP solution fixes degrees of freedom in the spectral solution. Subfigure (a) depicts the DOF matrix, transposed such that each row grayscale-codes an eigenvector giving a zero- or low-cost degree of freedom in the embedding. Intensities indicate motility of nodes. Subfigure (b) shows a sparse binary basis for these degrees of freedom obtained by thresholding an independent components analysis of the DOF matrix. Each row indicates a node or group of nodes that can be moved without cost; physically, these are “armatures” of the embedding, often nodes chained by collinear constraints.

The spectral embedding of the Ames dataset took roughly 3 seconds to compute. Again, the consistency errors are reduced but the results are not directly comparable with those in [1] because the problem is rather different than the version reported in [1]—we have many more nodes and constraints, some of which are inconsistent. For example, the Ames dataset contained one node whose constraints were rotated  $43^\circ$  out of alignment with the rest of the data. More notably, our dataset is also underconstrained.

Green dataset spectral embedding	mean	max	std
positional error (mm)	$3.541 \times 10^{-3}$	$1.197 \times 10^{-2}$	$2.390 \times 10^{-3}$
orientation error (degrees)	$1.823 \times 10^{-5}$	$1.639 \times 10^{-4}$	$2.182 \times 10^{-5}$
Ames dataset spectral embedding	mean	max	std
positional error (mm)	$5.696 \times 10^2$	$2.511 \times 10^3$	$3.704 \times 10^2$
orientation error (degrees)	$2.280 \times 10^0$	$4.273 \times 10^1$	$3.035 \times 10^0$

**Table 1.** Consistency errors of the spectral embeddings of Green and Ames datasets. World scale was estimated from GPS baselines.

This affords an opportunity to illustrate how the spectral analysis identifies degrees of freedom in the solution.

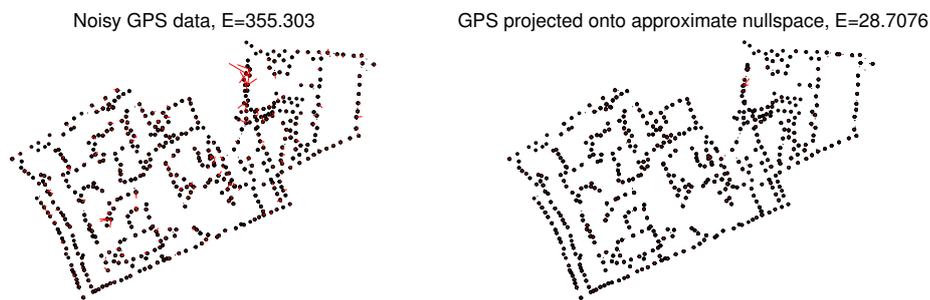
The problem is underconstrained in that several nodes lack more than one linearly independent constraint, and thus can slide freely. It is also overconstrained in that many of the multiply constrained nodes have no error-free embedding; the constraints are slightly inconsistent. Consequently the first 13 eigenvectors give  $\mathcal{E} = 0$  embeddings in which the inconsistently constrained nodes are collapsed into point clusters while the rest—mainly underconstrained nodes—are distributed through space. These “degenerate” solutions turn out to be degrees-of-freedom (DOFs) of the nontrivial solution, eigenvector  $\mathbf{v}_{14}$ , which is the first eigenvector with nonzero error  $\mathcal{E}(\mathbf{v}_{14}^{(3)}) = \lambda_{14} \approx 7.3 \times 10^{-5}$ . It specifies an embedding that distributes all nodes through space in a reasonable but imperfect reconstruction of the true scene geometry—reflecting constraints that are not consistent with the true geometry of the scene. Figure 5 shows that the GPS data is well reconstructed by projection to and back-projection from a low-error embedding subspace, comprising the zero-error DOFs (eigenvectors #1-13, which articulate individual nodes), the base solution (eigenvector #14), and a few small-error DOFs (eigenvectors #15-17, which articulate groups of densely connected nodes forming the “arms” of the embedding). Enforcing positive lengths via a small quadratic programming problem automatically finds the correct embedding as a mixture of the 20 lowest-error eigenvectors. This QP embedding has roughly 1/3 the error of the GPS data, and remains stable when more eigenvectors are considered in the QP problem. Rotational alignment further reduces the error  $\mathcal{E}$  by roughly an order of magnitude and corrects the  $43^\circ$  misaligned node, among others. The rotationally aligned embedding is not diagrammed because it is visually indistinguishable from the QP solution.

### 4.3 Campus

The Campus dataset consists of 566 nodes and 866 directional constraints spanning roughly a square kilometer. This is far too few to specify an embedding ( $\dim(\text{null}(\mathbf{H}_{\mathcal{E}})) \approx 370$ ), and indeed this dataset has never been processed as a whole. We used the 400 low-order eigenvectors of  $\mathbf{H}_{\mathcal{E}}$  to denoise this data; figure 6 shows that consistency error declines by an order of magnitude.

## 5 Discussion

The spectral method offers a linear-time optimal solution to graph embedding from directional constraints. These solutions have maximal fidelity to the constraints and



**Fig. 6.** Denoising of MIT GPS dataset via approximate nullspace projection. The group of nodes showing large inconsistencies in the original data (left) are straightened out in the denoised data (right), consistent with the true scene geometry (a street).

provide a basis for identifying flaws in the constraints. Some flaws can be automatically detected and corrected, yielding high-quality embeddings of ill-constrained problems such as the Ames Court dataset even when the constraints have substantial systematic errors. However in general there are classes of ill-posed problems that can be resolved only via additional observations. In the case of the Ames court dataset, where we have a nonrigid embedding that is partitioned into rigid subgraphs and a preferred layout based on positivity constraints, one would propose new views where two subgraphs come close to each other. In that light, one particularly attractive property of spectral schemes is that a solution can be updated in near-linear time when new data arrives, since new nodes and new constraints can be expressed as a series of low-rank modifications to an expanded  $\mathbf{H}_E$  matrix. Therefore new viewpoints and environmental features can be added incrementally and efficiently.

## References

1. Antone, M., Teller, S.: Scalable extrinsic calibration of omni-directional image networks. *IJCV* **49** (2002) 143–174
2. Hartley, R., Zisserman, A.: *Multiple View Geometry in Computer Vision*. Cambridge University Press, Cambridge (2000)
3. Faugeras, O., Luong, Q.T., Papadopoulos, T.: *The Geometry of Multiple Images*. MIT Press, Cambridge, MA (2001)
4. Horn, B.K.P.: *Robot Vision*. MIT Press, Cambridge, MA (1986)
5. Taylor, C.J., Kriegman, D.J.: Structure and motion from line segments in multiple images. In: *Proc. IEEE International Conference on Robotics and Automation*. (1992) 1615–1620
6. Becker, S., Bove, V.M.: Semiautomatic 3-D model extraction from uncalibrated 2-D camera views. In: *Proc. SPIE Image Synthesis*. Volume 2410. (1995) 447–461
7. Debevec, P.E., Taylor, C.J., Malik, J.: Modeling and rendering architecture from photographs: A hybrid geometry- and image-based approach. In: *Proc. SIGGRAPH*. (1996) 11–20
8. Shigang, L., Tsuji, S., Imai, M.: Determining of camera rotation from vanishing points of lines on horizontal planes. In: *Proc. ICCV*. (1990) 499–502

9. Leung, J.C.H., McLean, G.F.: Vanishing point matching. In: Proc. ICIP. Volume 2. (1996) 305–308
10. Mundy, J.L., Zisserman, A., eds.: Geometric Invariance in Computer Vision. MIT Press, Cambridge, MA (1992)
11. Luong, Q.T., Faugeras, O.: Camera calibration, scene motion, and structure recovery from point correspondences and fundamental matrices. IJCV **22** (1997) 261–289
12. Poelman, C.J., Kanade, T.: A paraperspective factorization method for shape and recovery. In: Proc. ECCV. (1994) 97–108
13. Adam, A., Rivlin, E., Shimshoni, I.: ROR: Rejection of outliers by rotations in stereo matching. In: Proc. CVPR. (2000) 2–9
14. Tutte, W.: Convex representations of graphs. Proc. London Mathematical Society **10** (1960) 304–320
15. Tutte, W.: How to draw a graph. Proc. London Mathematical Society **13** (1963) 743–768
16. Roweis, S.T., Saul, L.K.: Nonlinear dimensionality reduction by locally linear embedding. Science **290** (2000) 2323–2326

## A Computational considerations

Present-day numerical eigensolvers may not separate the nullspace of  $\mathbf{H}_{\mathcal{E}}$  into translational and embedding eigenvectors, and in general are prone to numerical error separating the nullspace and near-nullspace eigenvectors. Explicitly suppressing the translational eigenvectors usually improves the numerical stability of the problem. To do so, project  $\mathbf{H}_{\mathcal{E}}$  onto the orthogonal basis  $\mathbf{Q} \in \mathbb{R}^{Nd \times (N-1)d}$  of the null-space of the translation basis (i.e.,  $\mathbf{Q}^{\top} \mathbf{Q} = \mathbf{I}$  and  $\mathbf{Q}^{\top} (\mathbf{1} \otimes \mathbf{I}) = \mathbf{0}$ ), eigen-decompose the reduced problem there, then back-project the eigenvectors:

$$\mathbf{V}' \Lambda \mathbf{V}'^{\top} \stackrel{\text{EVD}}{\leftarrow} \mathbf{Q}^{\top} \mathbf{H}_{\mathcal{E}} \mathbf{Q} \quad (5)$$

$$\mathbf{V} \leftarrow \mathbf{Q} \mathbf{V}. \quad (6)$$

The quadratic form  $\mathbf{H}_{\mathcal{E}}$  is sparse and the null-space basis  $\mathbf{Q}$  has a very simple structure, suggesting special computational strategies to defray the cost of computing a very large EVD. In fact, neither matrix need be computed explicitly to obtain the desired eigenvector. First we observe that one can use equation (2) to compute  $\mathbf{y}^{\top} (\mathbf{I} - \mathbf{H}_{\mathcal{E}})$  directly from  $\mathbf{y}$  and the directional constraints, thereby yielding a power method for computing the eigenpair  $\{\lambda_{\max}(\mathbf{I} - \mathbf{H}_{\mathcal{E}}), \mathbf{v}_{\max}(\mathbf{I} - \mathbf{H}_{\mathcal{E}})\} = \{1 - \lambda_{\min}(\mathbf{H}_{\mathcal{E}}), \mathbf{v}_{\min}(\mathbf{H}_{\mathcal{E}})\}$  without forming  $\mathbf{H}_{\mathcal{E}}$ . Second, note that  $\mathbf{Q}$  is a centering matrix:  $\mathbf{Q} = \text{null}((\mathbf{1} \otimes \mathbf{I})^{\top}) = \text{null}(\mathbf{1}^{\top}) \otimes \mathbf{I}$ . The effect of  $\mathbf{Q}$  in equations (5–6) is to force the solution to be centered on the origin by ensuring that all rows of  $\mathbf{X} = \mathbf{y}^{(d)}$  sum to zero. Equations (5–6) may be dispensed with by modifying the power method to recenter  $\mathbf{y}$  on each iteration. This results in an  $O(dc)$  time algorithm for  $d$  dimensions and  $c > N$  constraints. For sparsely constrained problems, complexity is mildly supralinear in the number of nodes ( $O(dc) \approx O(dN)$ ); in a densely constrained problem the complexity will approach  $O(dN^2)$  from below.