Robust Point Correspondence by Concave Minimization*

João Maciel and João Costeira

Instituto de Sistemas e Robótica – Instituto Superior Técnico
Av. Rovisco Pais, 1049-001 Lisboa Codex, PORTUGAL
http://www.isr.ist.utl.pt/vislab/
email:{maciel,jpc}@isr.ist.utl.pt
phone:+351-21-8418083 fax:+351-21-8418291
Internal Report: VISLAB-TR 11/99

Abstract. We propose a new methodology for reliably solving the correspondence problem between points of two or more images. This is a key step in most problems of Computer Vision and, so far, no general reliable method exists to solve it.
Most methods admit domain specific assumptions that dictate the whole formulation of the problem. Furthermore, features are frequently missing or added through a sequence of images so the existence of outliers must be considered.
Our methodology is able to handle most of the commonly used assumptions in a unique formulation, independent of the domain of application and type of features. It performs correspondence and outlier rejection in a single step, and achieves global optimality with feasible computation. As far as we know, none of the existing methods meets these three requirements simultaneously.
Feature selection and correspondence are represented with a generalization of partial permutation matrices on an integer optimization problem. This is a blunt formulation, which considers the whole combinatorial space of possible point selections and correspondences. We find its global optimal solution by finding an equivalent concave objective function and relaxing the search domain into its compact convex-hull. The special structure of this extended problem assures its equivalence to the original one, but it can be optimally solved by efficient algorithms that avoid combinatorial search — e.g. *simplex*.
We exemplify the use of the methodology with pixel patch correlation and epipolar search in a trinocular system. Experiments in real images demonstrate the performance of our implementations. Finally we outline the use of the rigidity assumption under scaled orthographic projection, in a fully uncalibrated framework.

Keywords: correspondence problem, feature matching, constrained optimization, stereo vision.

* The authors would like to thank João Xavier and César Silva for the most valuable discussions. Special thanks to Prof. José Moura for his support and comments.

Table of Contents

1 Introduction

Estimating feature correspondences between two or more images is a long standing fundamental problem in Computer Vision. Most methods for 3D reconstruction, image alignment, object recognition or classification and camera self-calibration start by assuming that image feature points were extracted and put to correspondence. This is, therefore, a key step in most problems of Computer Vision and, so far, no general reliable method exists to solve it.

We propose a new methodology for reliably solving the point correspondence problem. Usually the correspondence problem is formulated as an optimization problem with a discrete domain and its solution falls under some search procedure. There are three main difficulties associated with the problem. First, it is ill posed in the sense that there are no general constraints to reduce its intrinsic ambiguity. Second, it suffers from high complexity due to the huge dimensionality of the combinatorial search space. Finally the existence of outliers must be considered, since features can be missing or added through a sequence of images, due to occlusions and errors of the feature extraction procedure.

1.1 Previous Work

In order to deal with the ambiguity of the correspondence problem, all methods must impose domain-specific assumptions. These assumptions are reflected on the three components of the optimization problem, namely the objective function — criterion — the constraints and the optimization algorithm. Next, we present examples of correspondence methods, organized according to these three aspects. Comprehensive surveys can be found in [6, 22].

The distinction between criteria and constraints is some times difficult to make, because algorithms hide the structure of the inherent optimization problem. We consider all assumptions that can be partially violated as criteria. The most commonly used criterion is image correlation [19, 4, 11], which reflects the assumption of image-brightness similarity. Another usual choice is proximity assumption [21, 11]. The most general assumption is probably 3D scene rigidity [1, 2, 16, 17]. The smoothness of the disparity field is also frequently used to disambiguate between multiple solutions [17, 13].

Assumptions linked to constraints are those that must be strictly followed. Examples are the order constraint [13, 15], the epipolar constraint [13, 15], and unicity of correspondence [5].

Finally, the large diversity of current methods comes mainly from the many different optimization algorithms. Dynamic programming [13], exhaustive search, analytical solution and convex minimization [11] are some of the few that guaranty optimality. Non-optimal approaches include greedy algorithms, simulated annealing, relaxation [5], alternation of optimization and projection on the constraints [2], graph search [15] and randomized search [17, 19].

Most methods use domain specific formulations that set bounds on their range of applications and, some times, create implicit unwanted constraints. For instance, the method of [13] does not apply when the order constraint is violated, because dynamic programming intrinsically assumes it. Furthermore, algorithms that do not consider the existence of outliers have poor performances when, for example, occlusions occur. Finally, most methods fail to guarantee global optimality of the obtained solutions, or have to explicitly include additional assumptions. For example, [19] deals with outliers but depends on an iterative scheme, which is not proved to converge to the global solution.

2 Contribution

Our methodology is generic, in the sense that it is able to handle most of the commonly used assumptions in a unique formulation. It performs correspondence and outlier rejection in a single step, achieving global optimality with feasible computation. As far as we know, no other method in the literature meets these three requirements simultaneously.

We use a single universal representation of the correspondence problem and a single procedure for its solution, independently of the domain of application and type of features. Both problems of feature selection and correspondence are formulated as single integer optimization problem that considers the whole space of possible point selections and correspondences. We find its global solution avoiding combinatorial search without having to impose additional assumptions. We do so by relaxing the discrete search domain into its compact convex-hull. The special structure of the constraints and objective function assure that this extension generates an equivalent problem that can be optimally solved by efficient algorithms — e.g. *simplex*. This technique is commonly used to solve combinatorial problems that emerge in different contexts [9].

In this paper we show the use of the methodology with some of the most frequent assumptions, and their performance is demonstrated by experiments in real images. Finally we outline the use of rigidity — under scaled orthographic projection — as unique assumption in a fully uncalibrated framework.

3 Methodology

For the sake of simplicity, we will consider the two image correspondence problem. The extension to sequences of many images is straightforward. In section 4.2 we show a three-image situation.

3.1 Feature Representation

Consider two sets of feature-points observed either by a stereo pair or a single moving camera. Consider that at least some of these pairs are the projections of the same 3D points. We collect the features $\mathbf{x}_p$ and $\mathbf{y}_p$, respectively from the first and second images, in two stacks of row vectors

$$\mathbf{X} = \begin{bmatrix} x_{1,1} & \cdots & x_{1,N} \\ \vdots & & \vdots \\ x_{p_1,1} & \cdots & x_{p_1,N} \end{bmatrix} \quad , \quad \mathbf{Y} = \begin{bmatrix} y_{1,1} & \cdots & y_{1,N} \\ \vdots & & \vdots \\ y_{p_2,1} & \cdots & y_{p_2,N} \end{bmatrix} \tag{1}$$

Each row represents a feature. Features can represent the image coordinates of feature-points ($N = 2$) or any image-related quantity like local intensity moments or even a local neighborhood of intensities. The type of information conveyed by the features has to be coherent with the criterion, but does not affect formulation of the correspondence problem itself.

3.2 Correspondence as an optimization problem

Using the previous definitions we can formulate the correspondence problem as an integer minimization problem $\mathbf{P}^* = \arg\min_{\mathbf{P}} \; J(\mathbf{X}, \mathbf{Y}, \mathbf{P})$, where $\mathbf{P}$ is a zero-one variable that represents the correspondence and outlier rejection.

This formulation unveils the difficulties inherent to the correspondence problem. First, to guaranty robustness in the presence of outliers, $\mathbf{P}$ must allow some features not to be corresponded, so it cannot be a simple permutation. Second, to avoid combinatorial explosion, we must be able to extend the zero-one domain and solve the problem using guided search.

Partial Permutation Matrices. The selection and correspondence between two vectors of features are represented by a $p_1 \times p_2$ indicator matrix $\mathbf{P} \in \mathcal{P}_{\mathrm{p}}(p_1, p_2)$, the set of $p_1 \times p_2$ *partial permutation matrices* ($\mathrm{p_p}$-matrices). We adopt Definition 1, generalizing the usual definition [8] to non-square matrices.

Definition 1 *A $p_1 \times p_2$ real matrix $\mathbf{P}$ is a $\mathrm{p_p}$-matrix* iff *it complies with the following conditions:*

$$\begin{aligned} (i) \;& \mathbf{P}_{i,j} \in \{0, 1\} \;,\; \forall i = 1 \ldots p_1 \;,\; \forall j = 1 \ldots p_2 \\ (ii) \;& \textstyle\sum_{i=1}^{p_1} \mathbf{P}_{i,j} \leq 1 \;,\; \forall j = 1 \ldots p_2 \\ (iii) \;& \textstyle\sum_{j=1}^{p_2} \mathbf{P}_{i,j} \leq 1 \;,\; \forall i = 1 \ldots p_1 \end{aligned}$$

A $\mathrm{p_p}$-matrix is a permutation matrix to which some columns and rows of zeros were added. Each entry $\mathbf{P}_{i,j}$ set to 1 indicates that features $\mathbf{X}_{i\cdot}$ (row i of $\mathbf{X}$) and $\mathbf{Y}_{j\cdot}$ (row j of $\mathbf{Y}$) are put to correspondence. $\mathrm{p_p}$-matrices represent, at most, one correspondence for each feature, and allow some features not to be matched. If row $\mathbf{P}_{i\cdot}$ is a row of zeros then feature $\mathbf{X}_{i\cdot}$ is not matched. If column $\mathbf{P}_{\cdot j}$ is a column of zeros then feature $\mathbf{Y}_{j\cdot}$ is not matched. This results on the following minimization problem

Problem 1
$$\begin{aligned} \mathbf{P}^* = \arg\min_{\mathbf{P}} \quad & J(\mathbf{X}, \mathbf{Y}, \mathbf{P}) \\ \text{s.t.} \quad & \mathbf{P} \in \mathcal{P}_{\mathrm{p}}(p_1, p_2) \end{aligned}$$

where J is a scalar objective function. Both correspondence and outlier rejection are intrinsic to this formulation because each element of $\mathcal{P}_{\mathrm{p}}(p_1, p_2)$ permutes only a sample of the features. In [19] the solution is computed based on a few random samples of the data, while Problem 1 chooses the best among all possible combinations of samples and permutations.

Constraining the number of correspondences. To avoid the trivial solution $\mathbf{P}^* = \mathbf{0}$, we establish a fixed number of correspondences $p_\mathrm{t} \leq \min(p_1, p_2)$ by considering a slightly different set of matrices $\mathcal{P}_\mathrm{p}^{p_\mathrm{t}}(p_1, p_2)$. We call these *rank-$p_\mathrm{t}$ partial permutation matrices* (rank-p_t $\mathrm{p_p}$-matrices).

Definition 2 *A $p_1 \times p_2$ real matrix $\mathbf{P}$ is a rank-p_t $\mathrm{p_p}$-matrix* iff *it complies with conditions (i), (ii), (iii) and*

$$(iv)\ \sum_{i=1}^{p_1}\sum_{j=1}^{p_2} \mathbf{P}_{i,j} = p_\mathrm{t}$$

Constraining the optimization problem to $\mathcal{P}_\mathrm{p}^{p_\mathrm{t}}(p_1, p_2)$ leads to a process of picking up just the best p_t correspondences, without having to impose thresholds to the criterion function. Like in most robust methods [19], p_t should be kept near the minimum number of features required by the assumed model or lower than the estimated number of inliers.

The case $p_\mathrm{t} = p_2 \geq p_1$ yields a very simple formulation, which is particularly useful when very few reliable features are extracted from the first image, while the second image is densely sampled. The simplicity comes from the fact that the conditions (*iii*) and (*iv*) can be changed to a single condition (*iii'*). We refer to the resulting set of matrices by $\mathcal{P}_\mathrm{p}^\mathrm{c}(p_1, p_2)$, the set of *columnwise partial permutation matrices* (columnwise $\mathrm{p_p}$-matrices).

Definition 3 *A $p_1 \times p_2$ real matrix $\mathbf{P}$ is a columnwise $\mathrm{p_p}$-matrix* iff *it complies with conditions (i), (ii) and*

$$(iii')\ \sum_{j=1}^{p_2} \mathbf{P}_{i,j} = 1\ ,\ \forall i = 1 \ldots p_1$$

3.3 Reformulation with a compact convex domain

Problem 1 is a bluntly posed — brute force — integer minimization problem. Traditional methods use greedy optimization, simulated-annealing or other methods that do not guarantee optimality.

We solve it by creating an equivalent problem — having the same solution — in a real compact convex domain, for which exact efficient algorithms exist. This is a common solution for some combinatorial optimization problems [9]. We consider the following equivalent *sub-stochastic problem* with concave objective

Problem 2
$$\mathbf{P}^* = \arg\min_{\mathbf{Q}} \ J_\epsilon(\mathbf{X}, \mathbf{Q}\mathbf{Y}) \quad \text{s.t.} \quad \mathbf{Q} \in \mathcal{DS}_\mathrm{s}(p_1, p_2)$$

where J_ϵ is a concave version of J — Equation 2 — and $\mathcal{DS}_\mathrm{s}(p_1, p_2)$ is the set of $p_1 \times p_2$ *doubly sub-stochastic matrices* — Definition 5.

The remaining of this section contains the proof for the equivalence of Problems 1 and 2 and the extension for the inclusion of rank-fixing constraints.

Equivalence of the Problems 1 and 2. Theorem 1 states the fundamental reason for the equivalence. Its proof can be found in [10].

Theorem 1 *A strictly concave function $J : \mathcal{C} \rightarrow \mathbb{R}$ attains its global minimum over a compact convex set $\mathcal{C} \subset \mathbb{R}^n$ at an extreme point of $\mathcal{C}$.*

The constraining set of a minimization problem with concave objective function can be changed to its convex-hull, provided that all the points in the original set are extreme points of the new compact set.

The problem now is how to find a concave function $J_\epsilon : \mathcal{DS}_\mathrm{s}(p_1,p_2) \rightarrow \mathbb{R}$ having the same values as J at every point of $\mathcal{P}_\mathrm{p}(p_1,p_2)$. Furthermore, we must be sure that the convex-hull of $\mathcal{P}_\mathrm{p}(p_1,p_2)$ is $\mathcal{DS}_\mathrm{s}(p_1,p_2)$, and that all $\mathrm{p_p}$-matrices are vertices of $\mathcal{DS}_\mathrm{s}(p_1,p_2)$, even in the presence of the rank-fixing constraint.

Concave equivalents to minimization problems. Consider Problem 1, where $J(\mathbf{q})$ is a class C^2 scalar function. Each entry of its Hessian matrix $\mathbf{H}(\mathbf{q})$ is a continuous function $\mathbf{H}_{ij}(\mathbf{q})$. J can be changed to its concave version J_ϵ by

$$J_\epsilon(\mathbf{q}) = J(\mathbf{q}) + \sum_{i=1}^{n} \epsilon_i q_i^2 - \sum_{i=1}^{n} \epsilon_i q_i \tag{2}$$

Remember that the constraints of Problem 1 include $q_i \in \{0,1\}$, $\forall i$, so $J_\epsilon(\mathbf{q}) = J(\mathbf{q})$, $\forall \mathbf{q} \in \mathcal{P}_\mathrm{p}(p_1,p_2)$. On the other hand $\mathcal{P}_\mathrm{p}(p_1,p_2)$ is bounded by a hypercube $\mathcal{B} = \{\mathbf{q} \in \mathbb{R}^n : 0 \le q_i \le 1,\ i = 1 \ldots n\}$. All $\mathbf{H}_{ij}(\mathbf{q})$ are continuous functions so they are bounded for $\mathbf{q} \in \mathcal{B}$ — Weierstrass' theorem. This means that we can always choose a set of finite values ϵ_r, defined by

$$\epsilon_r \le -\frac{1}{2}\left(\max_{\mathbf{q}} \sum_{s=1, s\neq r}^{n} \left| \frac{\partial^2 J(q)}{\partial q_r \partial q_s} \right| - \min_{\mathbf{q}} \frac{\partial^2 J}{\partial q_r^2} \right) \tag{3}$$

which impose the Hessian of J_ϵ to be *diagonally dominant* with negative diagonal.

Definition 4 *A $n \times n$ real matrix $\mathbf{H}$ is strictly diagonally dominant* iff

$$\mathbf{H}_{i,i} > \sum_{j=1, j\neq i}^{n} |\mathbf{H}_{i,j}| \ ,\ \forall i = 1, \ldots, n \tag{4}$$

A strictly diagonally dominant matrix having only negative elements on its diagonal is strictly negative definite [7], so these values of ϵ_r will guarantee that $J_\epsilon(\mathbf{q})$ is concave for $\mathbf{q} \in \mathcal{B}$ and, therefore, also for $\mathbf{q} \in \mathcal{DS}_\mathrm{s}(p_1,p_2)$.

Compact convex constraints for $\mathcal{P}_{\mathrm{p}}(p_1, p_2)$ problems. Problem 2 is constrained to the set of $\mathrm{ds_s}$-matrices. We extend the usual definition of $\mathrm{ds_s}$-matrices to the non-square case.

Definition 5 *A $p_1 \times p_2$ real matrix $\mathbf{Q}$ is a ds_{s}-matrix* iff *it complies with conditions (ii), (iii) and*

$$(v)\ \mathbf{Q}_{i,j} \geq 0\ ,\ \forall i = 1 \dots p_1\ ,\ \forall j = 1 \dots p_2$$

This set has the structure of a compact convex set in $\mathrm{I\!R}^{p_1 \times p_2}$. Its extreme points are the elements of $\mathcal{P}_{\mathrm{p}}(p_1, p_2)$. Theorem 2 states this in a different way. This a generalization of Birkhoff's theorem, and our proof can be found in C.

Theorem 2 *A matrix $\mathbf{Q}$ is a ds_{s}-matrix* iff *there is a finite number N of matrices $\mathbf{P}_1, \dots, \mathbf{P}_N \in \mathcal{P}_{\mathrm{p}}(p_1, p_2)$ and nonnegative scalars $\alpha_1, \dots, \alpha_N$ such that $\sum_{k=1}^{N} \alpha_k = 1$ and $\mathbf{Q} = \sum_{k=1}^{N} \alpha_k \mathbf{P}_k$.*

This fact together with Theorem 1 proves that the continuous Problem 2 is equivalent to the original discrete Problem 1, since we're assuming that J_ϵ was conveniently made concave. This way we can indirectly solve Problem 1 using guided search algorithms. Figure 1 summarizes the whole process.

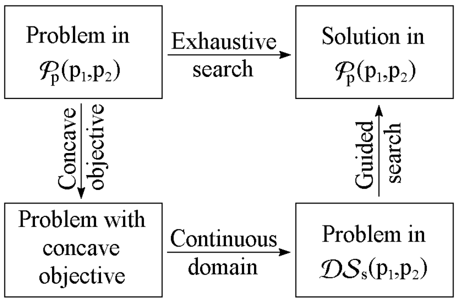

Fig. 1. Efficient solution of the combinatorial problem.

If we use the $\mathcal{P}_{\mathrm{p}}^{p_{\mathrm{t}}}(p_1, p_2)$ set instead then its compact extension is the set of *rank-p_{t} doubly sub-stochastic matrices* (rank-p_{t} $\mathrm{ds_s}$-matrices).

Definition 6 *A $p_1 \times p_2$ real matrix $\mathbf{Q}$ is a rank-p_{t} ds_{s}-matrix* iff *it complies with conditions (ii), (iii), (iv) and (v).*

$\mathcal{DS}_{\mathrm{s}}^{p_{\mathrm{t}}}(p_1, p_2)$ is still a compact convex set whose extreme points are the elements of $\mathcal{P}_{\mathrm{p}}^{p_{\mathrm{t}}}(p_1, p_2)$, as stated in Theorem 3.

Theorem 3 *A matrix $\mathbf{Q}$ is a rank-p_{t} ds_{s}-matrix* iff *there is a finite number N of matrices $\mathbf{P}_1, \dots, \mathbf{P}_N \in \mathcal{P}_{\mathrm{p}}^{p_{\mathrm{t}}}(p_1, p_2)$ and nonnegative scalars $\alpha_1, \dots, \alpha_N$ such that $\sum_{k=1}^{N} \alpha_k = 1$ and $\mathbf{Q} = \sum_{k=1}^{N} \alpha_k \mathbf{P}_k$.*

Our proof of Theorem 3 can be found in D. Finally, we refer to the convex hull of $\mathcal{P}_{\mathrm{p}}^{\mathrm{c}}(p_1, p_2)$ as $\mathcal{S}_{\mathrm{s}}^{\mathrm{c}}(p_1, p_2)$, the set of *columnwise sub-stochastic matrices* (columnwise $\mathrm{s_s}$-matrices). It is a compact convex set whose extreme points are the elements of $\mathcal{P}_{\mathrm{p}}^{\mathrm{c}}(p_1, p_2)$ — see D.

3.4 Constraints in canonical form

Most concave and linear programming algorithms assume that the problems have their constraints in canonical form. We now show how to put the constraints that define $\mathcal{DS}_{\mathrm{s}}^{p_{\mathrm{t}}}(p_1, p_2)$, $\mathcal{DS}_{\mathrm{s}}(p_1, p_2)$ and $\mathcal{S}_{\mathrm{s}}^{\mathrm{c}}(p_1, p_2)$ in canonical form, that is, how to state Problem 2 as

Problem 3

$$\begin{aligned} \mathbf{P}^* \;=\; \arg \min_{\mathbf{q}} \quad & J_\epsilon(\mathbf{X}, \mathbf{Y}, \mathbf{q}) \\ \text{s.t.} \quad & \mathbf{A}\mathbf{q} \leq \mathbf{b} \\ & \mathbf{q} \geq \mathbf{0} \end{aligned}$$

where $\mathbf{A}_{[m \times n]}$ and $\mathbf{b}_{[m \times 1]}$ define the intersection of m left half-planes in $\mathbb{R}^n$.

The natural layout for our variables is a matrix $\mathbf{Q}$, so we use $\mathbf{q} = \mathrm{vec}(\mathbf{Q})$, where vec() stacks the columns of its operand into a column vector. Condition (*ii*) is equivalent to $\mathbf{Q}.\mathbf{1}_{[p_2 \times 1]} \leq \mathbf{1}_{[p_1 \times 1]}$. Applying the vec operator [12] to both sides of this inequality we obtain $\left(\mathbf{1}_{[1 \times p_2]}^\top \otimes \mathbf{I}_{[p_1]}\right) \mathbf{q} \leq \mathbf{1}_{[p_1 \times 1]}$, where $\otimes$ is the Kronecker product, so set

$$\mathbf{A}_1 = \mathbf{1}_{[1 \times p_2]}^\top \otimes \mathbf{I}_{[p_1]} \quad ; \quad \mathbf{b}_1 = \mathbf{1}_{[p_1 \times 1]} \tag{5}$$

By the same token we express condition (*iii*) as

$$\mathbf{A}_2 = \mathbf{I}_{[p_2]} \otimes \mathbf{1}_{[1 \times p_1]}^\top \quad ; \quad \mathbf{b}_2 = \mathbf{1}_{[p_2 \times 1]} \tag{6}$$

Condition (*iv*) is equivalent to $\mathbf{1}_{[1 \times p_1 p_2]}^\top \mathbf{q} = p_{\mathrm{t}}$, so

$$\mathbf{A}_3 = \begin{bmatrix} \mathbf{1}_{[1 \times p_1 p_2]}^\top \\ -\mathbf{1}_{[1 \times p_1 p_2]}^\top \end{bmatrix} \quad ; \quad \mathbf{b}_3 = \begin{bmatrix} p_{\mathrm{t}} \\ -p_{\mathrm{t}} \end{bmatrix} \tag{7}$$

The intersection of conditions (*ii*), (*iii*) and (*iv*) results on the constraints of Problem 3 with

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_1 \\ \mathbf{A}_2 \\ \mathbf{A}_3 \end{bmatrix} \quad ; \quad \mathbf{b} = \begin{bmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \\ \mathbf{b}_3 \end{bmatrix} \tag{8}$$

3.5 Minimizing a linearly constrained concave function

To solve linear problems we use the *simplex* algorithm, and for concave problems we use an extension of the exact method of [3]. Other exact methods provide more efficient procedures [10, 9, 14] but this one is simple and easy to implement. Like the *simplex* algorithm, worst case complexity is factorial, but typically it visits only a small fraction of the vertices of the constraints. It performs surprisingly well in concave quadratic problems.

The method is based on a very simple iterative scheme. In each iteration the *next best* solution of a linear program is computed [20]. This can be accomplished by a few *simplex* pivoting steps. As iterations run, a sequence of ever-improving vertices of the constraining polytope is return, as well as a sequence of tighter and tighter bounds on the global minimum. Global optimality is tested by checking for coherence between the current best solution and the bounds.

3.6 Outline of the methodology

1. Extract points of interest, e.g. edges or corners (if applicable).
2. Convert the representation of these points into the appropriate sets of features $\mathbf{X}$ and $\mathbf{Y}$ — Equation 1.
3. Use $\mathbf{X}$ and $\mathbf{Y}$ to build the objective function J. Examples are in Sections 4.1, 4.2 and 5 — Equations 9, 12 and 13.
4. Use the procedure in Section 3.3 to produce an equivalent concave objective function J_ϵ — Equations 2 and 3.
5. Convert the desired convex constraints — Definitions 5 and 6 — into the canonical form using Equations 5, 6 and 7 of Section 3.4.
6. Solve the problem using a linear or concave programming algorithm.

4 Experiments

In this section we will consider two of the most frequently used assumptions and insert them in the described framework. The resulting methods are tested in real images and their robustness is compared with benchmark algorithms.

4.1 Image correlation criterion

Correlation of image patches is a popular criterion for correspondence because of its simplicity and robustness under photometric distortions. Most methods that use this criterion owe their complexity to the solution of emerging ambiguities and outlier rejection. Our formulation solves this in a natural way.

To use this criterion, features consist of image patches with N pixels centered around the previously segmented points of interest. Row i of $\mathbf{X}$ (and $\mathbf{Y}$) of Equation 1 is the row vectorization of a patch around the ith feature-point of the first (and second) image. The sum of the correlation coefficients of the rows of $\mathbf{X}$ and $\mathbf{Y}$ is given by the matrix inner product of $\hat{\mathbf{X}}$ and $\hat{\mathbf{Y}}$, which are normalized to have zero mean and unit norm rows. So, the objective function is $J_1(\mathbf{Q}) = -\mathrm{tr}\left(\mathbf{Q}\hat{\mathbf{Y}}\hat{\mathbf{X}}^\top\right)$. Using algebraic properties of the trace operator [12]

$$\begin{aligned} J_1(\mathbf{q}) &= -\mathbf{c}_1^\top \mathbf{q} \\ \mathbf{c}_1 &= \mathrm{vec}\left(\hat{\mathbf{X}}\hat{\mathbf{Y}}^\top\right) \end{aligned} \tag{9}$$

which is linear in $\mathbf{q} = \mathrm{vec}(\mathbf{Q})$. Problem 3 can be solved by *simplex* algorithm.

Results We applied the described method to the image pair of Figure 2, taken from the *Kitchen* sequence[2]. Feature selection is not a requirement of the method-

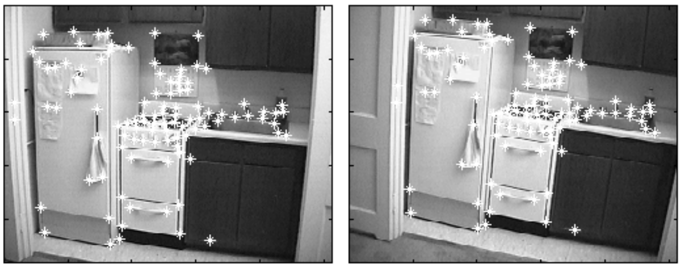

Fig. 2. Two images — 480×512 pixels and 256 gray levels — from the *Kitchen* sequence with 102 segmented feature points.

ology, but a practical way of reducing the dimensionality of the problem. The extreme case would be to consider image patches around each and every pixel.

To test the performance of the method we compared its results with those of two benchmark algorithms. The first algorithm solves the same problem with the same constraints with a greedy — suboptimal — approach. It builds a list with all possible correspondences and sequentially chooses the best ones, removing from the list all the pairs that become incompatible — in terms of the constraints. Both this algorithm and ours used 11 pixels wide correlation windows.

We found no other method that is able to reject outliers without additional assumptions so we decided to use a validation algorithm described in [19], that uses an extra rigidity assumption and is known to achieve very reliable results. This algorithm works on the feature pairs returned by the greedy algorithm. It applies a random sampling procedure to use only a few of the pairs in the estimation of the Fundamental matrix, and keeps the solution with smallest median of the feature to epipolar distances. It selects correspondences consistent with this Fundamental matrix. The number of iterations was set to 200.

[2] Data was provided by the Modeling by Videotaping group in the Robotics Institute, Carnegie Mellon University.

We measured the number of incorrect matches returned by the algorithms in experiments performed on three sets of corrupted data. The first set consisted of images with added zero-mean gaussian noise. In the second set, images were corrupted with salt-and-pepper noise. In the last set, outlier features — with randomly generated coordinates — were added to the list, and all the images were corrupted with zero mean gaussian noise with 50% of the images standard deviation. The results of the experiments are summarized on Figure 3.

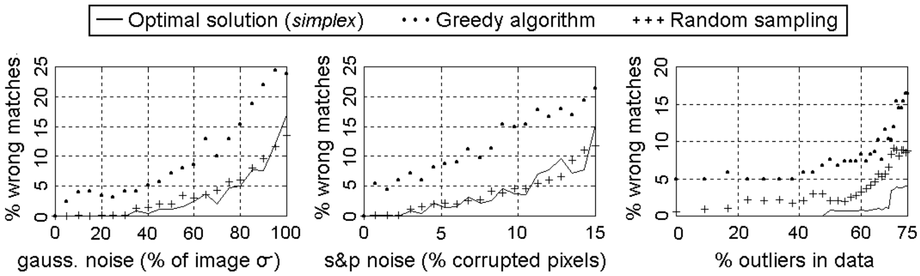

Fig. 3. Average number of incorrect matches found in 40 trials for each noise level.

Discussion The greedy solution consistently produced higher number of mismatches, so we conclude that assuring optimality is a key factor on the reduction of the number of mismatches. In the case of 40 features plus 40 outliers, the cardinality of $\mathcal{P}_{\mathrm{p}}(p_1, p_2)$ is roughly 10^{70} — see F. Exhaustive search would be impractical, while the *simplex* algorithm visits less than 300 solutions.

The reliability of the random sampling method is close to optimal when no outliers are present. Remember that it uses an extra rigidity assumption, so it is not applicable to nonrigid scenes where our method still performs well.

When more than 50% outliers are present, the reliability of the validation method decreases abruptly, since it relies on a median estimator. Even for a moderate number outliers the random sampling algorithm returns a higher percentage of mismatches than our method. This is mainly because the total number of returned correspondences is low — many good matches are rejected — so a few mismatches represent a considerable percentage. This effect would be reduced if the correspondences were revised according to the computed Fundamental matrix — epipolar search. [19] proposes a revision procedure that can be iterated with the validation phase, but no guarantee of convergence is given.

We conclude that simultaneous rejection and correspondence of features is a reliable strategy, even in the presence of as much as 70% of outliers in data. This performance was achieved with a feasible increment of computation. The linear problems were solved by the *lpSolve*[3] implementation of the *simplex* algorithm — running on a 166MHz Pentium processor — in a fraction of a second.

[3] written by M. Berkelaar and J. Dirks at Eindhoven U. of Technology

4.2 Epipolar search

We will now describe how to formulate an epipolar search method with our methodology. Consider a trinocular system in generic configuration — focal points are not collinear — from which we known all Fundamental Matrices. Figure 4 shows the notation. Matrices $\mathbf{P}_{1,2}$ and $\mathbf{P}_{1,3}$ are the partial-permutation

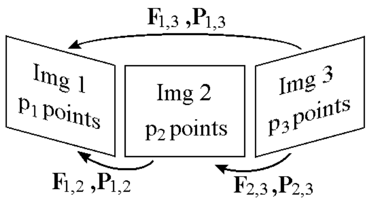

Fig. 4. Notation for a trinocular system.

variables of our problem. Each known Fundamental matrix $\mathbf{F}_{k,l}$ defines p_l epipolar lines $\mathcal{L}_{k,l}^{m}$, $m = 1, \ldots, p_l$ on image k. A point on image k corresponding to the m-th point on image l must lie *close*[4] to $\mathcal{L}_{k,l}^{m}$. This defines a constraint that we represent by an indicator matrix $\mathbf{S}_{k,l}$. Entry (i, j) of $\mathbf{S}_{k,l}$ is set to 1 if the i-th point on image k is close to $\mathcal{L}_{k,l}^{j}$. This means that entry (i, j) of $\mathbf{P}_{k,l}$ is a variable. On the other hand, if entry (i, j) of $\mathbf{S}_{k,l}$ is set to 0, then entry (i, j) of $\mathbf{P}_{k,l}$ permanently set to 0.

We represent these constraints implicitly with a *squeezed* set of variables $\mathbf{p}_{k,l}^{\mathsf{c}}$ of dimension $n_{k,l}$. These do not include the entries of $\mathbf{P}_{k,l}$ fixed to 0. We recover the full correspondence matrices through $\mathrm{vec}(\mathbf{P}_{k,l}) = \mathbf{B}_{k,l}\,\mathbf{p}_{k,l}^{\mathsf{c}}$, so the sub-stochastic constraints become

$$\begin{bmatrix} \mathbf{1}_{[1\times p_k]}^{\top} \otimes \mathbf{I}_{[p_l]} \\ \mathbf{I}_{[p_k]} \otimes \mathbf{1}_{[1\times p_l]}^{\top} \end{bmatrix} \mathbf{B}_{k,l}\,\mathbf{p}_{k,l}^{\mathsf{c}} \leq \mathbf{1}_{[1\times n_{k,l}]} \tag{10}$$

We close the loop by estimating the *compound correspondence* $\hat{\mathbf{P}}_{2,3} = \mathbf{P}_{1,2}^{\top}\mathbf{P}_{1,3}$ and checking its coherence with $\mathbf{S}_{2,3}$. The objective function is the sum of all point-to-epipolar distances

$$J_2\left(\mathbf{P}_{1,2}, \mathbf{P}_{1,3}, \hat{\mathbf{P}}_{2,3}\right) = \sum_{i=1}^{p_1}\sum_{j=1}^{p_2}\left(\mathbf{P}_{1,2} \odot \mathbf{D}_{1,2}\right)_{i,j} + \sum_{i=1}^{p_1}\sum_{j=1}^{p_3}\left(\mathbf{P}_{1,3} \odot \mathbf{D}_{1,3}\right)_{i,j} + \sum_{i=1}^{p_2}\sum_{j=1}^{p_3}\left(\hat{\mathbf{P}}_{2,3} \odot \mathbf{D}_{2,3}\right)_{i,j} + \sum_{i=1}^{p_2}\sum_{j=1}^{p_3}\left(\hat{\mathbf{P}}_{2,3} \odot \bar{\mathbf{S}}_{2,3}\right)_{i,j} \tag{11}$$

where $\odot$ is the elementwise product, $\bar{\mathbf{S}}_{2,3} = \left(\mathbf{1}_{[p_2\times p_3]} - \mathbf{S}_{2,3}\right)$, and $\mathbf{D}_{k,l}(i, j)$ is a matrix with the distances between the points $i = 1, \ldots, p_k$ of image k and

[4] define a distance threshold or choose a few from the nearest

the epipolar lines $\mathcal{L}^{j}_{k,l}$. The lack of coherence between $\hat{\mathbf{P}}_{2,3}$ and $\mathbf{S}_{2,3}$ is penalized by the last term. The other terms are used to disambiguate between different compatible solutions. By algebraic manipulation, we get the objective function

$$J_2(\mathbf{p}) = \mathbf{p}^\top \mathbf{J}_2 \mathbf{p} + \mathbf{c}_2^\top \mathbf{p} \tag{12}$$

written in a complete vector of variables $\mathbf{p} = \begin{bmatrix} \mathbf{p}^{\mathrm{c}}_{1,2} \\ \mathbf{p}^{\mathrm{c}}_{1,3} \end{bmatrix}$, and with

$$\begin{aligned} \mathbf{J}_2 &= \mathbf{B}_{1,2}^\top \left[\left(\mathbf{D}_{2,3} + \bar{\mathbf{S}}_{2,3} \right) \otimes \mathbf{I}_{[p_1]} \right] \mathbf{B}_{1,3} \\ \mathbf{c}_2 &= \begin{bmatrix} \mathbf{B}_{1,2}^\top \operatorname{vec}\left(\mathbf{D}_{1,2}\right) \\ \mathbf{B}_{1,3}^\top \operatorname{vec}\left(\mathbf{D}_{1,3}\right) \end{bmatrix} \end{aligned}$$

Results We applied the described method to the images of Figure 5, which are details taken from the *Castle* sequence[5]. The white lines are epipolar lines

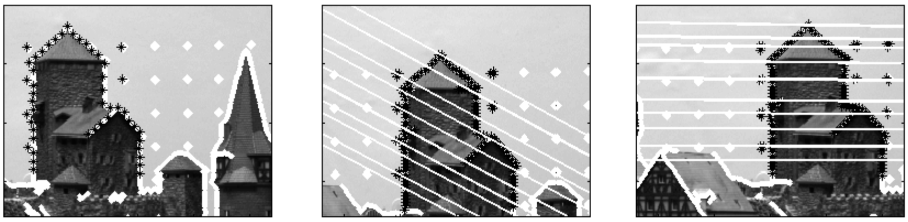

Fig. 5. Three images from the *Castle* sequence, and some of the epipolar lines.

corresponding to a few points of the first image. Note the amount of ambiguities due to multiple crossings with edges.

The white dots are edge points from the Canny edge extractor. All the used feature points — the crosses in black — come from inside a manually defined rectangular region of interest. They were automatically chosen by a bucketing procedure to guaranty a minimum distance between them. This subsample was performed in order to reduce the dimensionality of the problem. At the end we obtained 50 points from the first image and 130 points from each of the other images. Note that the second and third images contain, at least, 80 outliers, so the problem is solved in the presence of roughly 60% of outliers in the data. We set $p_{\mathrm{t}} = 30$, and obtained the correspondences in Figure 6.

Discussion We visually detected 3 mismatches in $\mathbf{P}_{1,2}$ and 2 mismatches in $\mathbf{P}_{1,3}$, corresponding to 8% errors. We conclude that the epipolar constraints of a

[5] Data was provided by the Calibrated Imaging Laboratory at Carnegie Mellon University, supported by ARPA, NSF, and NASA.

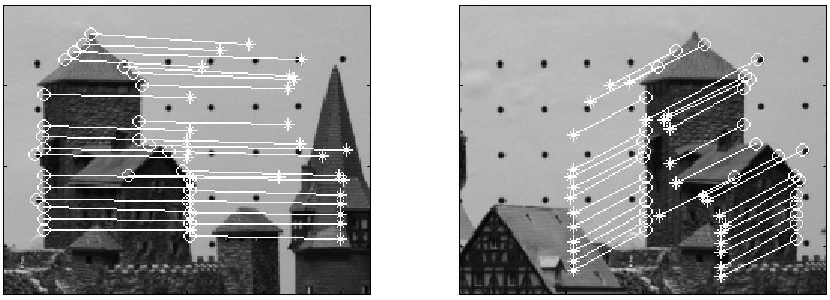

Fig. 6. Graphical representation of the obtained $\mathbf{P}_{1,2}$ and $\mathbf{P}_{1,3}$.

calibrated trinocular system are enough to build a correspondence method that is reliable in situations with many ambiguities.

The algorithm stopped after less than 100 iterations, when the bounds — Section 3.5 — were closer than a fixed threshold. We set this threshold to admit only solutions without any violation of the epipolar constraints. The depicted solution is not the optimal, though its mismatches occur only when more than one point lie inside the lozenge defined by the crossing of two epipolar lines.

Any set of compatible points inside these lozenges corresponds to a valid solution in terms of the epipolar constraints. The first three terms of the objective function — Equation 11 — introduce small differences between the objective value of these solutions. They successfully eliminate most of the degeneracies of the problem, but the introduction of more assumptions would probably result on a better performance.

5 Further Developments

One step forward would be to impose one single general assumption about the scene — rigidity — and no assumptions whatsoever about scene projections. Our first approach considers the use of a fully uncalibrated scaled-orthographic camera pair.

Suppose we segment p_1 feature-points from the first image and p_2 from the second, then measured their image coordinates u_p^1 and v_p^1 and arranged these values in two matrices $\mathbf{X}$ and $\mathbf{Y}$ like in Equation 1. We will try to determine a matrix $\mathbf{P}^* \in \mathcal{P}_{\mathrm{p}}^{\mathrm{c}}(p_1, p_2)$, such that $\mathbf{X}$ and $\mathbf{P}^*\mathbf{Y}$ will have the corresponding features placed at the same rows. To do so we use $\mathbf{C} = \mathbf{I}_{[p_1]} - \frac{1}{p_1}\mathbf{1}_{[p_1 \times p_1]}$ to normalize the rows of the observation matrices for zero mean and group them in $\mathbf{O}_{[p_1 \times 2F]} = [\mathbf{CX} \mid \mathbf{CPY}]$. Theorem 4 presents a very useful property of this matrix. Its proof can be found in G.

Theorem 4 *When* $\mathbf{P} = \mathbf{P}^*$ *is the correct* $\mathcal{P}_{\mathrm{p}}^{\mathrm{c}}(p_1, p_2)$ *matrix, and the measurements are taken without noise, the matrix of grouped centered scaled-orthographic observations* $\mathbf{O}$ *is, at most, of rank three.*

This is a generalization of the Rank Theorem of [18] for scaled-orthographic cameras. Nondegenerate — fully 3D — objects provide rank-3 observation matrices when $\mathbf{P} = \mathbf{P}^*$. In conclusion, in the absence noise, our goal would be to find the $\mathbf{P}^* \in \mathcal{P}_{\mathrm{p}}(p_1, p_2)$ matrix that minimizes rank $(\mathbf{O})$.

5.1 Approximate rank

Suppose now that the measurements are affected by noise, so that we observe $\mathbf{O}' = \left[\mathbf{C}\left(\mathbf{X} + \mathbf{E}^X\right) \mid \mathbf{CP}\left(\mathbf{Y} + \mathbf{E}^Y\right)\right]$. In general, $\mathbf{O}'$ has rank 4 for every $\mathbf{P}$, but we can measure the error of its optimal LS rank-3 approximation in λ_4, the smallest singular value of $\mathbf{O}'$. When noise levels are small enough, λ_4 is minimized when $\mathbf{P} = \mathbf{P}^*$. This can be cast into a polynomial objective function that generates Problem 4.

Problem 4

$$\mathbf{P}^* = \arg\min_{\mathbf{P}} \quad J_3(\mathbf{P}) \qquad \text{s.t.} \quad \mathbf{P} \in \mathcal{P}_{\mathrm{p}}(p_1, p_2)$$

with

$$J_3(\mathbf{P}) = \boldsymbol{\omega}^\top \mathbf{q}^{[4]} \tag{13}$$

where $\boldsymbol{\omega}$ *is a vector independent of* $\mathbf{Q}$ *and* $\mathbf{q}^{[4]} = \mathbf{q} \otimes \mathbf{q} \otimes \mathbf{q} \otimes \mathbf{q}$.

Theorem 5 states the conditions for which Problem 4 is equivalent to our problem. Its proof, and the validation of the λ_4 criterion, can be found in H.

Theorem 5 *It is always possible to find a scalar* $\epsilon > 0$ *such that if* $\left\|\mathbf{E}^X\right\| < \epsilon$ *and* $\left\|\mathbf{E}^Y\right\| < \epsilon$ *then the correspondence problem for points of a nondegenerate rigid object observed by a scaled orthographic camera is equivalent to Problem 4.*

A detailed expression for $\boldsymbol{\omega}$ is explicitly computed in G, so we can use the procedure of Section 3.3 to find a sub-stochastic extension of Problem 4. Ongoing work is being conducted on the implementation of an efficient algorithm to solve fourth-order concave programming problems, since the algorithm we used in the experiments of Section 4.2 is not suitable for $J_3(\mathbf{P})$.

6 Conclusion

We have shown a methodology to solve the correspondence problem, which avoids unwanted assumptions by requiring their explicit statement. Furthermore it reliably handles outliers, even in situations where other robust methods fail. Many assumptions can be formulated as linear objective functions. In this case the problem is solved in a highly efficient way by the *simplex* algorithm.

The most important limitation of the methodology is the dimensionality of the optimization problems. This poses some efficiency problems when the objective functions are high-order polynomials. A practical way of minimizing this is the selection of a small number of reliable features in one of the images.

Ongoing work is being conducted on the implementation of an efficient algorithm for high-order polynomial problems, and dealing with the assumption of rigidity under perspective projection. We plan to extend the methodology to dealing with extended sequences of images. Also we are working on reducing the dimensionality of the problem by integration of more than one assumption.

References

1. R. Basri. On uniqueness of correspondence under orthographic and perspective projections. Technical Report AIM 1333, MIT, AI Lab, 1991.
2. R. Berthilsson and K. Astrom. Reconstruction of 3D-curves from 2D-images using affine methods for curves. In *Proc. CVPR*, 1997.
3. A. Cabot and R. Francis. Solving certain nonconvex quadratic minimiation problems by ranking the extreme points. *Operations Research*, 18(1):82–86, Feb 1970.
4. R. Deriche, Z. Zhang, Q. Luong, and O. Faugeras. Robust recovery of the epipolar geometry for an uncalibrated stereo rig. In *Proc. ECCV*, pages 567–576, 1994.
5. S. Gold, A. Rangarajan, C. Lu, S. Pappu, and E. Mjolsness. New algorithms for 2D and 3D point matching. *Pattern Recognition*, 31(8):1019–1031, 1998.
6. R. Haralick and L. Shapiro. *Computer and Robot Vision*. Addison-Wesley, 1993.
7. R. Horn and C. Johnson. *Matrix Analysis*. Cambridge U. Press, 1985.
8. R. Horn and C. Johnson. *Topics in Matrix Analysis*. Cambridge U. Press, 1991.
9. R. Horst and P. Pardalos, editors. *Handbook of Global Optimization*. Kluwer, 1995.
10. R. Horst, P. Pardalos, and N. Thoai. *Introduction to Global Optimization*, volume 3 of *Nonconvex Optimization and Its Appl.* Wiley, 1990.
11. B. Lucas and T. Kanade. An iterative image registration technique with an application to stereo vision. In *Proc. 7th International Joint Conference on AI*, 1981.
12. H. Lutkepohl. *Handbook of Matrices*. Wiley, 1996.
13. Y. Ohta and T. Kanade. Stereo by intra- and inter-scanline search using dynamic programming. *IEEE Trans. PAMI*, 7(2):139–154, March 1985.
14. P. Pardalos and J. Rosen. Methods for global concave minimization: A bibliographic survey. *SIAM Review*, 28(3):367–379, Sep 1986.
15. S. Roy and I. Cox. A maximum-flow formulation of the n-camera stereo correspondence problem. In *Proc. ICCV*, 1997.
16. A. Shashua. Correspondence and affine shape from two orthographic views. Technical Report AIM 1327, MIT, AI Lab, 1991.
17. G. Sudhir, S. Banerjee, and A. Zisserman. Finding point correspondences in motion sequences preserving affine structure. *CVIU*, 68(2):237–246, November 1997.
18. C. Tomasi and T. Kanade. Shape from motion from image streams under orthography: a factorization method. *IJCV*, 9(2):137–154, November 1992.
19. P. Torr. *Motion Segmentation and Outlier Detection*. PhD thesis, Dept. Engineering Science, U. Oxford, 1995.
20. S. Wallace. Pivoting rules and redundancy schemes. *BIT*, 25:274–280, 1985.
21. Z. Zhang. Iterative point matching for registration of free-form curves. Technical Report RR-1658, INRIA, 1992.
22. Z. Zhang. Le problème de la mise en correspondance: L'état de l'art. Technical Report RR-2146, INRIA, 1993.

A Notation

This list briefly explains the notation used throughout the text.

$\otimes$		Kronecker product
$\odot$		Hadamard (or Schur) element-wise product
$\mathbf{0}_{[m\times n]}$		$m \times n$ matrix of zeros
$\mathbf{1}_{[m\times n]}$		$m \times n$ matrix of ones
α^f		Scale factor for the scaled-orthographic camera in frame f
α_k		Weights of a convex combination
$\boldsymbol{\Delta}$	$= \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \otimes \begin{bmatrix} 0 & \frac{1}{2} \\ -\frac{1}{2} & 0 \end{bmatrix}$	$\mathbf{A}_{[2\times 2]} \Rightarrow \det(\mathbf{A}) = \mathrm{vec}(\mathbf{A})^\top \boldsymbol{\Delta} \mathrm{vec}(\mathbf{A})$
$\boldsymbol{\Lambda}$	$= \begin{bmatrix} \alpha^1 & \cdots & \alpha^{p_1} \end{bmatrix} \otimes \mathbf{I}_{[2]}$	Matrix of the camera scale factors
$\boldsymbol{\omega}$		A vector that defines a fourth order polynomial $\boldsymbol{\omega}^\top \mathbf{q}^{[4]}$
$\boldsymbol{\Pi}_n^k$		A fixed $n^k \times n^k$ permutation such that $\forall \mathbf{A}_{[n\times n]}, \forall k \in \mathbb{N} \Rightarrow \mathrm{vec}\left(\mathbf{A}^{[k]}\right) = \boldsymbol{\Pi}_n^k \mathrm{vec}(\mathbf{A})^{[k]}$
$\boldsymbol{\Xi}$	$= \mathbf{C}^\top \left[\mathbf{I}_{[p_1]} - \mathbf{CX}(\mathbf{CX})^+\right] \mathbf{C}$	Intermediate result
$\mathbf{A}$		Matrix of the linear constraints Also used to represent a generic matrix
$\mathbf{A}_i$		i-th row or sub-matrix of $\mathbf{A}$
$\mathbf{A}^+$		Moore-Penrose inverse of $\mathbf{A}$
$\mathbf{A}^{[k]}$	$= \mathbf{A} \otimes \cdots \otimes \mathbf{A}$	Kronecker power
$\mathbf{b}$		Right-hand side of the constraints
$\mathcal{B}$	$= \{\mathbf{q} \in \mathbb{R}^n : 0 \le q_i \le 1,\ \forall i\}$	Unit hypercube in $\mathbb{R}^n$
$\mathbf{C}$	$= \mathbf{I}_{[p_1]} - \frac{1}{p_1}\mathbf{1}_{[p_1\times p_1]}$	Centralizing matrix
$\mathcal{C}$		A compact convex subset of $\mathbb{R}^n$
$\mathrm{ds_s}$		Short for *doubly sub-stochastic*
$\mathcal{DS}_\mathrm{s}(p_1, p_2)$		Set of $p_1 \times p_2$ doubly sub-stochastic matrices
$\mathcal{S}_\mathrm{s}^\mathrm{c}(p_1, p_2)$		Set of $p_1 \times p_2$ columnwise sub-stochastic matrices
$\mathcal{DS}_\mathrm{s}^{p_\mathrm{t}}(p_1, p_2)$		Set of $p_1 \times p_2$ doubly sub-stochastic matrices of rank p_t
F		Number of frames in the image sequence
$\mathbf{H}$	$= \left[\frac{\partial^2 J}{\partial x_i \partial x_j}\right]$	Hessian matrix of function J
$\mathbf{i}^f, \mathbf{j}^f$		Orthonormal image coordinate system for frame f, written in world coordinates
$\mathbf{I}_{[n]}$		n-dimensional identity matrix
$J, J_1, J_2, \ldots$		Scalar cost functions

$\mathbf{J}$		Matrix of the quadratical part of cost function J
$\mathbf{K}_{m,l}$		Commutation matrix (fixed $ml \times ml$ permutation)
$\mathbf{L}, \mathbf{M}, \mathbf{N}$		Generic matrices
N		Dimensionality of the features
$\mathbf{O}$		Collection of centered observations
p		Feature number
$p_1, p_2, \ldots$		Number of features on the successive images
p_t		Rank of the correspondence (the number of points put in correspondence)
$\mathrm{p_p}$		Short for *partial permutation*
$\mathbf{P}$		A permutation or partial permutation matrix
$\mathcal{P}_\mathrm{p}(p_1, p_2)$		Set of $p_1 \times p_2$ partial permutation matrices
$\mathcal{P}_\mathrm{p}^\mathrm{c}(p_1, p_2)$		Set of $p_1 \times p_2$ columnwise partial permutation matrices
$\mathcal{P}_\mathrm{p}^{p_\mathrm{t}}(p_1, p_2)$		Set of $p_1 \times p_2$ partial permutation matrices of rank p_t
$\mathbf{q}$	$= \mathrm{vec}(\mathbf{Q})$	Vectorized version of $\mathbf{Q}$ (the variable)
$\mathbf{Q}$		A (doubly) stochastic or sub-stochastic matrix
$\mathbf{R}$	$= \left[\mathbf{i}^1\, \mathbf{j}^1\ \cdots\ \mathbf{i}^F\, \mathbf{j}^F\right]$	Camera rotations
$\mathbf{s}_p$		3D coordinated of the p-th point (row-vector)
$\mathrm{s_s}$		Short for *sub-stochastic*
$\mathbf{S}$	$= \begin{bmatrix} \mathbf{s}_1 \\ \vdots \\ \mathbf{s}_{p_1} \end{bmatrix}$	Matrix of shape (3D coordinates)
$\mathcal{S}_\mathrm{s}^\mathrm{c}(p_1, p_2)$		Set of $p_1 \times p_2$ columnwise sub-stochastic matrices
u_p^f, v_p^f		Image coordinates of the p-th feature point on the f-th frame
$\tilde{u}_p^f\ \ (\tilde{v}_p^f)$	$= u_p^f - \frac{1}{p_1}\sum_{p=1}^{p_1} u_p^f$	Centered observations
$\mathrm{vec}(\cdot)$		Vectorization operator
$x_{p,k}, y_{p,k}$		k-th coordinate of feature p, respectively on the first and second images
$\mathbf{X}, \mathbf{Y}$		Matrix gatherings of the features
$\mathbf{X}_i, \mathbf{Y}_i$		Full representations of individual features (i-th row of $\mathbf{X}$, j-th row of $\mathbf{Y}$)
$\hat{\mathbf{X}}, \hat{\mathbf{Y}}$		Row-normalized (zero mean, unit norm) versions of $\mathbf{X}, \mathbf{Y}$
$\tilde{\mathbf{X}}, \tilde{\mathbf{Y}}$		Matrices of the centered observations

B Reference guide of algebraic results

This section lists some results that are frequently used throughout the text. Omitted demonstrations and related results can be found in [7, 8, 12].

B.1 Results on the vectorization operator

- For any two matrices $\mathbf{A}_{[l \times m]}$ and $\mathbf{B}_{[m \times n]}$

$$\mathrm{vec}(\mathbf{AB}) = \left(\mathbf{B}^\top \otimes \mathbf{I}_{[m]}\right) \mathrm{vec}(\mathbf{A}) \tag{14}$$

- For any matrix $\mathbf{A}$

$$\mathrm{vec}(\mathbf{A}) = \mathbf{K}_{m,l}\ \mathrm{vec}\left(\mathbf{A}^\top\right) \tag{15}$$

  where $\mathbf{K}_{m,l}$ is a commutation matrix, which is a fixed $ml \times ml$ permutation matrix

B.2 Results on the trace of a matrix

- For any two matrices $\mathbf{A}$ and $\mathbf{B}$ with compatible dimensions

$$\mathrm{tr}(\mathbf{AB}) = \mathrm{vec}\left(\mathbf{B}^\top\right)^\top \mathrm{vec}(\mathbf{A}) \tag{16}$$

- For any four matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ and $\mathbf{D}$ with compatible dimensions

$$\mathrm{tr}\,(\mathbf{ABCD}) = \mathrm{vec}\left(\mathbf{C}^\top\right)^\top \left(\mathbf{B}^\top \otimes \mathbf{D}\right) \mathrm{vec}(\mathbf{A}) \tag{17}$$

$$= \mathrm{vec}\left(\mathbf{A}^\top\right)^\top \left(\mathbf{D}^\top \otimes \mathbf{B}\right) \mathrm{vec}(\mathbf{C}) \tag{18}$$

- For any three matrices $\mathbf{A}_{[l \times m]}$, $\mathbf{B}_{[m \times n]}$ and $\mathbf{C}_{[n \times r]}$

$$\mathrm{tr}\,(\mathbf{ABC}) = \mathrm{vec}\left(\mathbf{B}^\top\right)^\top \left(\mathbf{A}^\top \otimes \mathbf{C}\right) \mathrm{vec}(\mathbf{I}_{[r]}) \tag{19}$$

B.3 Results on the determinant

- For any two matrices $\mathbf{M}_{[l \times m]}$ and $\mathbf{N}_{[l \times n]}$, if $\mathbf{L}_{[l \times (m+n)]} = [\mathbf{M} \mid \mathbf{N}]$ then

$$\det\left(\mathbf{L}^\top \mathbf{L}\right) = \det\left(\mathbf{M}^\top \mathbf{M}\right) \det\left[\mathbf{N}^\top \left(\mathbf{I}_{[m]} - \mathbf{N}\mathbf{N}^+\right) \mathbf{N}\right] \tag{20}$$

  where $\mathbf{N}^+$ is the Moore-Penrose inverse of $\mathbf{N}$.
- For any 2×2 matrix $\mathbf{A}$

$$\det(\mathbf{A}) = \mathrm{vec}(\mathbf{A})^\top \boldsymbol{\Delta}\ \mathrm{vec}(\mathbf{A}) \tag{21}$$

  with

$$\boldsymbol{\Delta} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \otimes \begin{bmatrix} 0 & \frac{1}{2} \\ -\frac{1}{2} & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & \frac{1}{2} \\ 0 & 0 & -\frac{1}{2} & 0 \\ 0 & -\frac{1}{2} & 0 & 0 \\ \frac{1}{2} & 0 & 0 & 0 \end{bmatrix} \tag{22}$$

  This handy result does not come in the references, but its demonstration is trivial.

B.4 Results on Kronecker products

- For any four matrices **A**, **B**, **C** and **D** with compatible dimensions

$$\mathbf{AB} \otimes \mathbf{CD} = (\mathbf{A} \otimes \mathbf{C})(\mathbf{B} \otimes \mathbf{D}) \tag{23}$$

- For any integers n and k there is a fixed permutation matrix $\boldsymbol{\Pi}_n^k$ such that, for any $n \times n$ matrix **A**

$$\mathrm{vec}\left(\mathbf{A}^{[k]}\right) = \boldsymbol{\Pi}_n^k \ \mathrm{vec}(\mathbf{A})^{[k]} \tag{24}$$

- For any two matrices **A** and **B** with compatible dimensions and for any integer k

$$(\mathbf{A}\ \mathbf{B})^{[k]} = \mathbf{A}^{[k]}\mathbf{B}^{[k]} \tag{25}$$

Proof. This result can be proven by induction. For $k = 1$ the result is evident. We now have to prove it for $k' = k + 1$ assuming it is true for k. To do so we rely on the fact that, for any matrix **A**

$$\mathbf{A}^{[k]} \otimes \mathbf{A} = \mathbf{A} \otimes \cdots \otimes \mathbf{A} = \mathbf{A}^{[k+1]} \tag{26}$$

which comes from the associativity of the Kronecker product. Using this, we can say that

$$(\mathbf{A}\ \mathbf{B})^{[k+1]} = (\mathbf{A}\ \mathbf{B})^{[k]} \otimes (\mathbf{A}\ \mathbf{B}) \tag{27}$$

$$= \left(\mathbf{A}^{[k]}\mathbf{B}^{[k]}\right) \otimes (\mathbf{A}\ \mathbf{B}) \tag{28}$$

The last step comes from the assumption that the result holds for k. Finally, using property 23

$$(\mathbf{A}\ \mathbf{B})^{[k+1]} = \left(\mathbf{A}^{[k]} \otimes \mathbf{A}\right)\left(\mathbf{B}^{[k]} \otimes \mathbf{B}\right) \tag{29}$$

$$= \mathbf{A}^{[k+1]}\mathbf{B}^{[k+1]} \tag{30}$$

□

C $\mathcal{P}_{\mathrm{p}}(p_1, p_2)$ are the vertices of $\mathcal{DS}_{\mathrm{s}}(p_1, p_2)$

We're about to prove that a matrix $\mathbf{Q}$ is a $\mathrm{ds_s}$-matrix *iff* there is a finite number N of matrices $\mathbf{P}_1, \ldots, \mathbf{P}_N \in \mathcal{P}_{\mathrm{p}}(p_1, p_2)$ and nonnegative scalars $\alpha_1, \ldots, \alpha_N$ such that $\sum_{k=1}^{N} \alpha_k = 1$ and $\mathbf{Q} = \sum_{k=1}^{N} \alpha_k \mathbf{P}_k$.

This is a generalization of Birkhoff's theorem, which states that the set of $n \times n$ doubly stochastic matrices is a compact convex set whose extreme points are the $n \times n$ permutation matrices. Our proof for works much in the same way as those presented in [7] for Birkhoff's theorem. It is divided in three main steps. Consider $\mathbf{P} \in \mathcal{P}_{\mathrm{p}}(p_1, p_2)$ and $\mathbf{Q} \in \mathcal{DS}_{\mathrm{s}}(p_1, p_2)$. In Step 1 we show the sufficiency of the condition, that is, that any convex combination of $\mathrm{p_p}$-matrices gives a $\mathrm{ds_s}$-matrix. Steps 2 and 3 show the necessity of the condition. In Step 2 we show that every $p_1 \times p_2$ $\mathrm{p_p}$-matrix is an extreme point of $\mathcal{DS}_{\mathrm{s}}(p_1, p_2)$. Finally, in Step 3 we show that there is no extreme point of $\mathcal{DS}_{\mathrm{s}}(p_1, p_2)$ lying outside $\mathcal{P}_{\mathrm{p}}(p_1, p_2)$.

Step 1 Each entry $\mathbf{Q}_{i,j}$ of $\mathbf{Q}$ is a linear combination of the correspondent entries of the N matrices $\mathbf{P}_k$

$$\mathbf{Q}_{i,j} = \sum_{k=1}^{N} \alpha_k \mathbf{P}_k(i,j) \tag{31}$$

Since $\alpha_k \geq 0$, $\forall k = 1 \ldots N$ and $\sum_{k=1}^{N} \alpha_k = 1$, then the minimum and maximum possible values of an entry $\mathbf{Q}_{i,j}$ of $\mathbf{Q}$ are, respectively

$$\min(\mathbf{Q}_{i,j}) = \sum_{k=1}^{N} \alpha_k \min[\mathbf{P}_k(i,j)] \;=\; \sum_{k=1}^{N} \alpha_k.0 \;=\; 0 \tag{32}$$

$$\max(\mathbf{Q}_{i,j}) = \sum_{k=1}^{N} \alpha_k \max[\mathbf{P}_k(i,j)] \;=\; \sum_{k=1}^{N} \alpha_k.1 \;=\; 1 \tag{33}$$

where, $\min[\mathbf{P}_k(i,j)]$ and $\max[\mathbf{P}_k(i,j)]$ stand for the minimum and maximum possible values of each entry of matrix $\mathbf{P}_k$.

This shows that condition (v) holds true for all matrices like our $\mathbf{Q}$. Furthermore, by linearity, and using condition (ii)

$$\begin{aligned}
\sum_{i=1}^{n} \mathbf{Q}_{i,j} &= \sum_{i=1}^{n} \sum_{k=1}^{N} \alpha_k \mathbf{P}_k(i,j) \\
&= \sum_{k=1}^{N} \sum_{i=1}^{n} \alpha_k \mathbf{P}_k(i,j) \\
&= \sum_{k=1}^{N} \left[\alpha_k \sum_{i=1}^{n} \mathbf{P}_k(i,j) \right] \\
&\leq \sum_{k=1}^{N} \alpha_k.1 = 1 \; , \; \forall j = 1
\end{aligned} \tag{34}$$

The same reasoning could be done on the rows, so conditions (ii) and (iii) are also always verified. In conclusion, our $\mathbf{Q}$ matrix is always a $\mathrm{ds_s}$-matrix.

Step 2 Consider that $\mathbf{P}$ is an $p_1 \times p_2$ $\mathrm{p_p}$-matrix. It will then have, at most, one nonzero entry (set to 1) in each row and in each column.

Suppose that this matrix is not an extreme point of $\mathcal{DS}_\mathrm{s}(p_1, p_2)$. Then it will be possible to find two matrices $\mathbf{A}, \mathbf{B} \in \mathcal{DS}_\mathrm{s}(p_1, p_2)$, $\mathbf{A} \neq \mathbf{B} \neq \mathbf{P}$ and two positive scalars α_1, α_2 such that $\alpha_1 + \alpha_2 = 1$ and $\mathbf{P} = \alpha_1 \mathbf{A} + \alpha_2 \mathbf{B}$. Then all entries of $\mathbf{A}$ and $\mathbf{B}$ that correspond to zero entries of $\mathbf{P}$ will be zeroed, because of the non-negativity of both the scalars and the entries of the matrices in $\mathcal{DS}_\mathrm{s}(p_1, p_2)$. Therefore $\mathbf{A}$ and $\mathbf{B}$ will have, at most, one nonzero entry per row and per column. Since they belong to $\mathcal{DS}_\mathrm{s}(p_1, p_2)$ their row and column sums are equal or less than 1, so all those nonzero entries must be equal or lower than 1, so $\mathbf{A} = \mathbf{B} = \mathbf{P}$, which contradicts the initial supposition. In conclusion, all elements of $\mathcal{P}_\mathrm{p}(p_1, p_2)$ are extreme points of $\mathcal{DS}_\mathrm{s}(p_1, p_2)$.

Step 3 Let's now suppose that an extreme point $\mathbf{Q} \in \mathcal{DS}_\mathrm{s}(p_1, p_2)$ is not a $\mathrm{p_p}$-matrix. We will make a distinction between two possible cases

a) All rows and columns sum either 0 or 1 (at least one of them sums 1, otherwise $\mathbf{Q} = \mathbf{0}$, which is a $\mathrm{p_p}$-matrix). In this case there will be at least one row (say row i_1) with at least 2 nonzero entries. In that row, choose an entry $0 < \mathbf{Q}_{i_1,j_1} < 1$, in column j_1. Place $+1$ at the corresponding entry (i_1, j_1) of an $n \times n$ matrix of zeros $\mathbf{A}$. Since $\sum_{i=1}^{n} \mathbf{Q}(i, j_1) = 1$ (cannot be 0 because $\mathbf{Q}_{i_1,j_1} > 0$), then there must be some other nonzero entry in that column, other than $\mathbf{Q}_{i_1,j_1}$. Choose $0 < \mathbf{Q}_{i_2,j_1} < 1$, and place -1 at the corresponding entry (i_1, j_1) of $\mathbf{A}$. Continue the process of choosing the successive entries $\mathbf{Q}_{i,j}$ (each successive pair of which is alternately in the same row or column) and placing alternately $+1$ and -1 in the corresponding entries of $\mathbf{A}$ until an entry is chosen that has previously been chosen (a cycle is closed). Let $\mathbf{Q}_{i',j'}$ and $\mathbf{Q}_{i'',j''}$ be respectively the smallest and the largest entries in the cycle.
Now define $\mathbf{Q}_+ = \mathbf{Q} + \epsilon\mathbf{A}$ and $\mathbf{Q}_- = \mathbf{Q} - \epsilon\mathbf{A}$ with $0 < \epsilon < \min(\mathbf{Q}_{i',j'}, 1 - \mathbf{Q}_{i'',j''})$. It is clear that this ϵ will make both $\mathbf{Q}_+$ and $\mathbf{Q}_-$ obey to condition (v). Furthermore, all $+1$ and -1 entries of $\mathbf{A}$ come in pairs at each row and column, so all rows and columns of $\mathbf{A}$ sum zero. This implies that the rows and columns of $\mathbf{Q}_+$ and $\mathbf{Q}_-$ have the same sums as those of $\mathbf{Q}$ (that is, 0 or 1), so they comply with conditions (ii) and (iii), and are, therefore, $\mathrm{ds_s}$-matrices. Note, however, that $\mathbf{Q} = \frac{1}{2}\mathbf{Q}_+ + \frac{1}{2}\mathbf{Q}_-$, so $\mathbf{Q}$ cannot be an extreme point of $\mathcal{DS}_\mathrm{s}(p_1, p_2)$.

b) There is at least one column[6] in $\mathbf{Q}$ whose sum belongs to $]0, 1[$. In this case, choose the first entry $0 < \mathbf{Q}_{i_1,j_1} < 1$ in one of these columns, and proceed the process described in a) by choosing a new entry $0 < \mathbf{Q}_{i_1,j_2} < 1$, and so on. The difference now is that it may not be possible to close a cycle, because we can find, for instance, a column with only one nonzero entry $(0 < \mathbf{Q}_{i_k,j_k} < 1)$. When this happens stop and use this incomplete

[6] the process is identical if a row is considered

$\mathbf{A}$ matrix, where unpaired $+1$ or -1 entries may occur. Note, however, that $\mathbf{Q}_+ = \mathbf{Q} + \epsilon\mathbf{A}$ and $\mathbf{Q}_- = \mathbf{Q} - \epsilon\mathbf{A}$ are again both $\mathrm{ds_s}$-matrices, provided that $\epsilon < \min(\epsilon_1, \epsilon_2)$.
The constraint $\epsilon_1 = \min(\mathbf{Q}_{i',j'}, 1 - \mathbf{Q}_{i'',j''})$ assures condition (v) while

$$\epsilon_2 \;=\; \min(1 - \mathbf{Q}_{i_1,j_1}, 1 - \sum_{i=1}^{p_1} \mathbf{Q}_{i,j_1}, 1 - \mathbf{Q}_{i_k,j_k}, 1 - \sum_{i=1}^{p_1} \mathbf{Q}_{i,j_k})$$

assures conditions (ii) and (iii). Use row sums instead where they sum less than 1. Again, $\mathbf{Q} = \frac{1}{2}\mathbf{Q}_+ + \frac{1}{2}\mathbf{Q}_-$, so $\mathbf{Q}$ cannot be an extreme point of $\mathcal{DS}_\mathrm{s}(p_1, p_2)$.

In conclusion, all extreme points of $\mathcal{DS}_\mathrm{s}(p_1, p_2)$ are $\mathrm{p_p}$-matrices.

□

D A proof for Theorem 3

Build $\mathcal{DS}_{\mathrm{s}}^{p_{\mathrm{t}}}(p_1, p_2)$ by cutting $\mathcal{DS}_{\mathrm{s}}(p_1, p_2)$ with a plane of constant L_1-norm. We just have to check whether or not new extreme points are created with this process, or if the final result is the convex hull of some extreme points of $\mathcal{DS}_{\mathrm{s}}(p_1, p_2)$. This can be shown by recalling Farkas' Lemma.

E $\mathcal{P}_{\rm p}^{\rm c}(p_1, p_2)$ are the vertices of $\mathcal{S}_{\rm s}^{\rm c}(p_1, p_2)$

We're about to prove that a matrix $\mathbf{Q}$ is a columnwise $\mathrm{ds_s}$-matrix *iff* there is a finite number N of matrices $\mathbf{P}_1, \dots, \mathbf{P}_N \in \mathcal{P}_{\rm p}^{\rm c}(p_1, p_2)$ and nonnegative scalars $\alpha_1, \dots, \alpha_N$ such that $\sum_{k=1}^{N} \alpha_k = 1$ and $\mathbf{Q} = \sum_{k=1}^{N} \alpha_k \mathbf{P}_k$.

This proof is substantially supported by the proof in C. It is also divided in three steps. Consider $\mathbf{P} \in \mathcal{P}_{\rm p}^{\rm c}(p_1, p_2)$ and $\mathbf{Q} \in \mathcal{S}_{\rm s}^{\rm c}(p_1, p_2)$. Steps 1 and 2 are similar to the corresponding steps in Appendix C, with the difference that condition (ii) becomes the equality (ii').

$$(ii')\ \sum_{i=1}^{p_1} \mathbf{Q}_{i,j} = 1\ ,\ \forall j = 1 \dots p_2$$

This way we can show that that any convex combination $\sum_{k=1}^{N} \alpha_k \mathbf{P}_k$ of columnwise $\mathrm{p_p}$-matrices with gives a columnwise $\mathrm{ds_s}$-matrix $\mathbf{Q}$ and that every $p_1 \times p_2$ columnwise $\mathrm{p_p}$-matrix is an extreme point of $\mathcal{S}_{\rm s}^{\rm c}(p_1, p_2)$. Step 3 is considerably different.

Step 3 We must prove that there is no extreme point of $\mathcal{S}_{\rm s}^{\rm c}(p_1, p_2)$ outside $\mathcal{P}_{\rm p}^{\rm c}(p_1, p_2)$ by showing that $\mathcal{P}_{\rm p}^{\rm c}(p_1, p_2) \subset \mathcal{P}_{\rm p}(p_1, p_2)$ (complies with condition (ii') because the extreme points are themselves elements of $\mathcal{S}_{\rm s}^{\rm c}(p_1, p_2)$).
Suppose that $\mathbf{Q}$ is an extreme point of $\mathcal{S}_{\rm s}^{\rm c}(p_1, p_2)$ but not of $\mathcal{DS}_{\rm s}(p_1, p_2)$. This means that it would be possible to find two matrices $\mathbf{A}, \mathbf{B} \in \mathcal{DS}_{\rm s}(p_1, p_2)$ such that $\mathbf{A} \neq \mathbf{B} \neq \mathbf{Q}$ and $\mathbf{Q} = \frac{1}{2}\mathbf{A} + \frac{1}{2}\mathbf{B}$. $\mathbf{Q} \in \mathcal{P}_{\rm p}^{\rm c}(p_1, p_2)$, so this matrix obeys to condition (ii). Assume that $j_1 \dots j_k$ are the columns of $\mathbf{Q}$ that sum zero, and $j_{k+1} \dots j_{p_2}$ those that sum 1. Then

$$\begin{aligned}
\forall j = j_1 \dots j_k\ ,\ \ 0 &= \sum_{i=1}^{p_1} \mathbf{Q}_{i,j} \\
&= \sum_{i=1}^{p_1} \left(\frac{1}{2}\mathbf{A}_{i,j} + \frac{1}{2}\mathbf{B}_{i,j} \right) \\
&= \frac{1}{2}\sum_{i=1}^{p_1} \mathbf{A}_{i,j} + \frac{1}{2}\sum_{i=1}^{p_1} \mathbf{B}_{i,j}
\end{aligned} \tag{35}$$

so the rows $j_1 \dots j_k$ of both $\mathbf{A}$ and $\mathbf{B}$ sum zero.
Now, in one hand

$$\begin{aligned}
\forall j = j_{k+1} \dots j_{p_2}\ ,\ \ 1 &= \sum_{i=1}^{p_1} \mathbf{Q}_{i,j} \\
&= \sum_{i=1}^{p_1} \left(\frac{1}{2}\mathbf{A}_{i,j} + \frac{1}{2}\mathbf{B}_{i,j} \right) \\
&= \frac{1}{2}\sum_{i=1}^{p_1} \mathbf{A}_{i,j} + \frac{1}{2}\sum_{i=1}^{p_1} \mathbf{B}_{i,j}
\end{aligned} \tag{36}$$

On the other hand, since $\mathbf{A}, \mathbf{B} \in \mathcal{DS}_{\mathrm{s}}(p_1, p_2)$, these obey to condition (ii), that is

$$\sum_{i=1}^{p_1} \mathbf{A}_{i,j} \leq 1 \ , \ \forall j = 1 \ldots p_2 \tag{37}$$

$$\sum_{i=1}^{p_1} \mathbf{B}_{i,j} \leq 1 \ , \ \forall j = 1 \ldots p_2 \tag{38}$$

Solving equations 36, 37 and 38 together for rows $j_{k+1} \ldots j_{p_2}$, results on a single solution per row, requiring that $\sum_{i=1}^{p_1} \mathbf{A}_{i,j} = \sum_{i=1}^{p_1} \mathbf{B}_{i,j} = 1$. So $\mathbf{A}, \mathbf{B} \in \mathcal{P}_{\mathrm{p}}^{\mathrm{c}}(p_1, p_2)$, but this contradicts the initial supposition.
In conclusion: $\mathcal{P}_{\mathrm{p}}^{\mathrm{c}}(p_1, p_2) \subseteq \mathcal{P}_{\mathrm{p}}(p_1, p_2)$

□

F Cardinality of $\mathcal{P}_{\mathrm{p}}(\boldsymbol{p_1}, \boldsymbol{p_2})$, $\mathcal{P}_{\mathrm{p}}^{p_{\mathrm{t}}}(\boldsymbol{p_1}, \boldsymbol{p_2})$ and $\mathcal{P}_{\mathrm{p}}^{\mathrm{c}}(\boldsymbol{p_1}, \boldsymbol{p_2})$

The $p_1 \times p_2$ $\mathrm{p_p}$-matrices — with $p_1 \leq p_2$ — are, for all possible values of $p \leq p_1$, the $p \times p$ identity matrices with their rows placed in all possible arrangements of p_1 positions and their columns placed in all possible arrangements of p_2 positions. So, the cardinality of $\mathcal{P}_{\mathrm{p}}(p_1, p_2)$ is

$$\sum_{p=0}^{p_1} \mathrm{A}_p^{p_1} \mathrm{A}_p^{p_2} \;=\; \sum_{p=0}^{p_1} \frac{p_1! p_2!}{(p_1 - p)!(p_2 - p)!} \tag{39}$$

The $p_1 \times p_2$ rank-p_{t} $\mathrm{p_p}$-matrices are all possible arrangements of the rows and columns of a $p_{\mathrm{t}} \times p_{\mathrm{t}}$ identity matrix respectively in p_1 and p_2 positions. So, the cardinality of $\mathcal{P}_{\mathrm{p}}^{p_{\mathrm{t}}}(p_1, p_2)$ is

$$\mathrm{A}_{p_{\mathrm{t}}}^{p_1} \mathrm{A}_{p_{\mathrm{t}}}^{p_2} \;=\; \frac{p_1! p_2!}{p_{\mathrm{t}}! p_{\mathrm{t}}!} \tag{40}$$

The $p_1 \times p_2$ columnwise $\mathrm{p_p}$-matrices are the rank-p_{t} $\mathrm{p_p}$-matrices in the special case when $p_{\mathrm{t}} = p_1$. So, the cardinality of $\mathcal{P}_{\mathrm{p}}^{\mathrm{c}}(p_1, p_2)$ is

$$\mathrm{A}_{p_1}^{p_1} \mathrm{A}_{p_1}^{p_2} \;=\; \frac{p_2!}{p_1!} \tag{41}$$

G A proof for Theorem 4

Theorem 4 is a generalization of the Rank Theorem of [18], and our proof follows their approach.

The theorem assumes that the image coordinates of the extracted feature-points are grouped together in a matrix of centered observations

$$\mathbf{O}_{[p_1 \times 2F]} = \left[\tilde{\mathbf{X}} \mid \tilde{\mathbf{Y}}\right] = [\mathbf{CX} \mid \mathbf{CPY}] \tag{42}$$

and it sates that when $\mathbf{P} = \mathbf{P}^*$, the correct $\mathcal{P}^{\mathrm{c}}_{\mathrm{p}}(p_1, p_2)$ matrix, and the measurements are taken without noise, matrix $\mathbf{O}$ is, at most, of rank three. Refer to

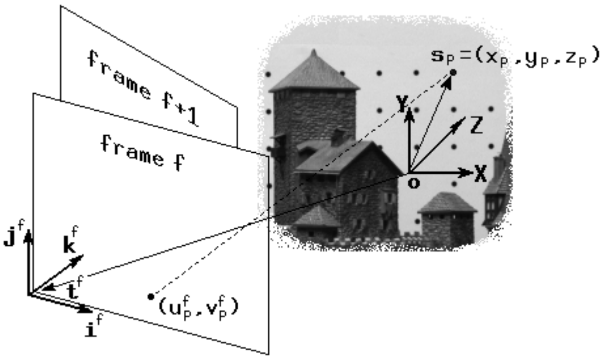

Fig. 7. The camera, the object and their coordinate systems.

Figure 7, where the origin $\mathbf{o}$ of the world coordinate system is at the centroid of the object points $\mathbf{s}_p = (x_p, y_p, z_p)$. We consider $\mathbf{s}_p$ and $\mathbf{t}^f$ as row vectors, in order to simplify the notation. The projection $\left(u_p^f, v_p^f\right)$ of each point $\mathbf{s}_p$ onto frame f is given by

$$u_p^f = \alpha^f \left(\mathbf{s}_p - \mathbf{t}^f\right) \mathbf{i}^f \tag{43}$$

$$v_p^f = \alpha^f \left(\mathbf{s}_p - \mathbf{t}^f\right) \mathbf{j}^f \tag{44}$$

The image coordinate systems $\left(\mathbf{i}^f, \mathbf{j}^f\right)$ are orthonormal so, under scaled-orthography, points are projected along $\mathbf{k}^f = \mathbf{i}^f \times \mathbf{j}^f$ and scaled with the scale factors α^f. When $\mathbf{P} = \mathbf{P}^*$ is correct then $\mathbf{PY}$ has the features placed in the same order as the corresponding features in $\mathbf{X}$, so the centered observation matrices yield

$$\begin{aligned}\tilde{u}_p^f &= u_p^f - \frac{1}{p_1}\sum_{p=1}^{p_1} u_p^f \\ &= \alpha^f \left(\mathbf{s}_p - \mathbf{t}^f\right)\mathbf{i}^f - \frac{1}{p_1}\sum_{p=1}^{p_1} \alpha^f \left(\mathbf{s}_p - \mathbf{t}^f\right)\mathbf{i}^f\end{aligned}$$

$$= \alpha^f \left(\mathbf{s}_p - \frac{1}{p_1} \sum_{p=1}^{p_1} \mathbf{s}_p \right) \mathbf{i}^f$$
$$= \alpha^f \mathbf{s}_p \mathbf{i}^f \tag{45}$$

and, similarly $\tilde{v}_p^f = \alpha^f \mathbf{s}_p \mathbf{j}^f$. Because of these two equations, matrix $\mathbf{O}$ can be written as

$$\mathbf{O} = \left[\tilde{\mathbf{X}} \mid \tilde{\mathbf{Y}} \right] = \mathbf{S}\mathbf{R}\boldsymbol{\Lambda} \tag{46}$$

where $\mathbf{S}$ is the shape matrix — its rows are the $\mathbf{s}_p$ vectors — $\mathbf{R} = \left[\mathbf{i}^1 \, \mathbf{j}^1 \cdots \mathbf{i}^f \, \mathbf{j}^f \right]$ represents the camera rotation, and $\boldsymbol{\Lambda}$ contains the scale factors organized in the following way

$$\boldsymbol{\Lambda} = \begin{bmatrix} \alpha^1 & & \\ & \ddots & \\ & & \alpha^{p_1} \end{bmatrix} \otimes \mathbf{I}_{[2]} \tag{47}$$

The result of the theorem follows from the observation that $\mathbf{S}$ is $p_1 \times 3$ and $\mathbf{R}$ is $3 \times 2F$ □

H A proof for Theorem 5

Theorem 5 states the equivalence of Problem 4 and the correspondence problem for points of a nondegenerate rigid object observed under a scaled orthographic camera. To prove this we introduce two intermediate Problems 5 and 6.

Problem 5

$$\mathbf{P}^* = \arg\min_{\mathbf{P}} \quad J_4(\mathbf{P}) \qquad \text{s.t.} \quad \mathbf{P} \in \mathcal{P}_{\mathrm{p}}(p_1, p_2)$$

with

$$J_4(\mathbf{P}) = \lambda_4(\mathbf{O}) \tag{48}$$

$\lambda_4(\mathbf{O})$ is the fourth singular value of $\mathbf{O}$.

Problem 6

$$\mathbf{P}^* = \arg\min_{\mathbf{P}} \quad \det\left(\mathbf{O}^\top \mathbf{O}\right) \qquad \text{s.t.} \quad \mathbf{P} \in \mathcal{P}_{\mathrm{p}}(p_1, p_2)$$

We also assume that the solution of Problem 5 is nondegenerate.

We start by validating the $J_4(\mathbf{P}) = \lambda_4(\mathbf{O})$ criterion of Problem 5 when a suitable bound ϵ_1 is imposed on the norm of the noise. Next, we prove the equivalence between Problems 5 and 6 when the norm of the noise is bounded by an existing positive ϵ_2. We then prove the equivalence of Problems 6 and 4, independently of the noise level. The full theorem is proven by transitivity of the equivalence relation, and using $\epsilon = \min(\epsilon_1, \epsilon_2)$. The chain of equivalences is shown in Figure 8.

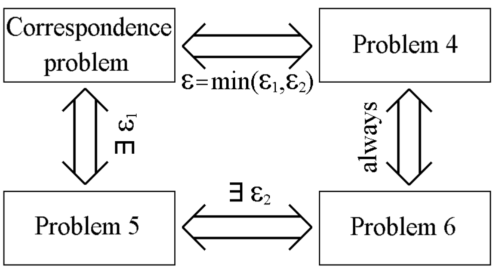

Fig. 8. The chain of equivalences that proves Theorem 5 for $\epsilon = \min(\epsilon_1, \epsilon_2)$.

H.1 Validity of the $\lambda_4(\mathbf{O})$ criterion

Theorem 4 states that in the absence of noise, and for the correct $\mathbf{P} = \mathbf{P}^*$ matrix, $\mathrm{rank}(\mathbf{O}) \leq 3$. For any $\mathbf{O}$ matrix we know that $\mathrm{rank}(\mathbf{O}) \leq 3 \Leftrightarrow \lambda_4(\mathbf{O}) = 0$. Since we are assuming that Problem 5 has a single nondegenerate solution, then there is no other $\mathbf{P}$ matrix that makes $J_4(\mathbf{P}) = \lambda_4(\mathbf{O}) = 0$ in the absence of noise, and there is a positive scalar δ such that the intervals

$$\{J \,:\, |J - J_4(\mathbf{P})| < \delta\}\ ,\ \forall \mathbf{P} \neq \mathbf{P}^* \in \mathcal{P}_{\mathrm{p}}(p_1, p_2) \qquad (49)$$

don't intersect

$$\{J \,:\, |J - J_4(\mathbf{P}^*)| < \delta\} \qquad (50)$$

The singular values of $\mathbf{O}$ are continuous functions of the entries of $\mathbf{O}$ — see [7] — so, in the presence of noise, J_4 is a continuous function of the entries of $\mathbf{X}' = \mathbf{X} + \mathbf{E}^X$ and $\mathbf{Y}' = \mathbf{X} + \mathbf{E}^Y$. By definition of continuity we know that there is a positive scalar ϵ_1 such that if $\left\|\mathbf{E}^X\right\| < \epsilon_1$ and $\left\|\mathbf{E}^Y\right\| < \epsilon_1$ then $J_4(\mathbf{X}', \mathbf{Y}', \mathbf{P})$ is always inside the intervals of Equation 49, $\forall \mathbf{P} \neq \mathbf{P}^* \in \mathcal{P}_{\mathrm{p}}(p_1, p_2)$, while $J_4(\mathbf{X}', \mathbf{Y}', \mathbf{P}^*)$ is inside the interval of Equation 50. This guarantees that, under these noise constraints, $\mathbf{P}^*$ is still the optimal solution of Problem 5, so this is an equivalent of the correspondence problem under the assumptions of Theorem 5. □

H.2 Equivalence of Problems 5 and 6

For any matrix $\mathbf{O}$, $\operatorname{rank}(\mathbf{O}) \leq 3 \;\Leftrightarrow\; \lambda_4(\mathbf{O}) = 0 \;\Leftrightarrow\; \det\left(\mathbf{O}^\top\mathbf{O}\right)$. Since we are assuming that Problem 5 has a single nondegenerate solution, then there is no $\mathbf{P}$ other than $\mathbf{P}^*$ that makes $\det\left(\mathbf{O}^\top\mathbf{O}\right) = 0$. So, in the absence of noise, the optimal solutions of Problems 6 and 5 are the same, that is to say, both problems are equivalent.

Using the same kind of reasoning of Section H.1, we could prove that there is a new bound ϵ_2 for the norm of the noise matrices for which Problems 6 and 5 remain equivalent. □

H.3 Equivalence of Problems 6 and 4

We start by using the result 20 with $\mathbf{M} = \tilde{\mathbf{X}}$ and $\mathbf{N} = \tilde{\mathbf{Y}}$ then $\mathbf{L} = \mathbf{O}$, so

$$\begin{aligned}\arg\min_{\mathbf{Q}} \det\left(\mathbf{O}^\top\mathbf{O}\right) &= \arg\min_{\mathbf{Q}} \left[\det\left(\mathbf{X}^\top\mathbf{C}^\top\mathbf{C}\mathbf{X}\right)\det\left(\mathbf{Y}^\top\mathbf{Q}^\top\boldsymbol{\Xi}\mathbf{Q}\mathbf{Y}\right)\right] \\ &= \arg\min_{\mathbf{Q}} \det\left(\mathbf{Y}^\top\mathbf{Q}^\top\boldsymbol{\Xi}\mathbf{Q}\mathbf{Y}\right)\end{aligned} \qquad (51)$$

where $\boldsymbol{\Xi}_{[p_1\times p_1]} = \mathbf{C}^\top\left[\mathbf{I}_{[p_1]} - \mathbf{C}\mathbf{X}(\mathbf{C}\mathbf{X})^+\right]\mathbf{C}$ is independent of $\mathbf{Q}$. The last step uses the fact that $\det\left(\mathbf{X}^\top\mathbf{C}^\top\mathbf{C}\mathbf{X}\right)$ is positive and independent of $\mathbf{Q}$. Now observe that $\mathbf{Y}^\top\mathbf{Q}^\top\boldsymbol{\Xi}\mathbf{Q}\mathbf{Y}$ is a 2×2 matrix, so we can use the result 21, obtaining

$$\det\left(\mathbf{Y}^\top\mathbf{Q}^\top\boldsymbol{\Xi}\mathbf{Q}\mathbf{Y}\right) = \operatorname{vec}\left(\mathbf{Y}^\top\mathbf{Q}^\top\boldsymbol{\Xi}\mathbf{Q}\mathbf{Y}\right)^\top \boldsymbol{\Delta}\operatorname{vec}\left(\mathbf{Y}^\top\mathbf{Q}^\top\boldsymbol{\Xi}\mathbf{Q}\mathbf{Y}\right) \qquad (52)$$

where

$$\boldsymbol{\Delta} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \otimes \begin{bmatrix} 0 & \frac{1}{2} \\ -\frac{1}{2} & 0 \end{bmatrix} \qquad (53)$$

Using the result 14 with $\mathbf{A} = \mathbf{B} = \mathbf{Q}\mathbf{Y}$

$$\mathrm{vec}\left(\mathbf{Y}^\top \mathbf{Q}^\top \boldsymbol{\Xi} \mathbf{Q}\mathbf{Y}\right) = \left(\mathbf{Y}^\top \mathbf{Q}^\top \otimes \mathbf{Y}^\top \mathbf{Q}^\top\right) \mathrm{vec}(\boldsymbol{\Xi}) \tag{54}$$

so, substituting this in equation 52 and keeping in mind that the determinant is a scalar function, so $\det(\mathbf{A}) = \mathrm{tr}\left[\det(\mathbf{A})\right]$, we get

$$\begin{aligned} \det\left(\mathbf{Y}^\top \mathbf{Q}^\top \boldsymbol{\Xi} \mathbf{Q}\mathbf{Y}\right) &= \mathrm{tr}\left[\det\left(\mathbf{Y}^\top \mathbf{Q}^\top \boldsymbol{\Xi} \mathbf{Q}\mathbf{Y}\right)\right] \\ &= \mathrm{tr}\left[\mathrm{vec}(\boldsymbol{\Xi})^\top (\mathbf{Q}\mathbf{Y} \otimes \mathbf{Q}\mathbf{Y}) \boldsymbol{\Delta} (\mathbf{Q}\mathbf{Y} \otimes \mathbf{Q}\mathbf{Y})^\top \mathrm{vec}(\boldsymbol{\Xi})\right] \end{aligned} \tag{55}$$

Using the result 19 with $\mathbf{A} = \mathrm{vec}(\boldsymbol{\Xi})^\top (\mathbf{Q}\mathbf{Y} \otimes \mathbf{Q}\mathbf{Y})$, $\mathbf{B} = \boldsymbol{\Delta}$ and $\mathbf{C} = \mathbf{A}^\top$, then

$$\begin{aligned} \det\left(\mathbf{Y}^\top \mathbf{Q}^\top \boldsymbol{\Xi} \mathbf{Q}\mathbf{Y}\right) = \mathrm{vec}\left(\boldsymbol{\Delta}^\top\right)^\top \Big(\left[\mathrm{vec}(\boldsymbol{\Xi})^\top (\mathbf{Q}\mathbf{Y} \otimes \mathbf{Q}\mathbf{Y})\right] \otimes \\ \otimes \left[\mathrm{vec}(\boldsymbol{\Xi})^\top (\mathbf{Q}\mathbf{Y} \otimes \mathbf{Q}\mathbf{Y})\right] \Big) \mathrm{vec}\left(\mathbf{I}_{[p_1^2]}\right) \end{aligned} \tag{56}$$

Using property 23 with $\mathbf{A} = \mathbf{C} = \mathrm{vec}(\boldsymbol{\Xi})^\top$ and $\mathbf{B} = \mathbf{D} = \mathbf{Q}\mathbf{Y} \otimes \mathbf{Q}\mathbf{Y}$

$$\det\left(\mathbf{Y}^\top \mathbf{Q}^\top \boldsymbol{\Xi} \mathbf{Q}\mathbf{Y}\right) = \mathrm{vec}\left(\boldsymbol{\Delta}^\top\right)^\top \mathrm{vec}(\boldsymbol{\Xi})^{[2]\top} (\mathbf{Q}\mathbf{Y})^{[4]} \mathrm{vec}\left(\mathbf{I}_{[p_1^2]}\right) \tag{57}$$

where $\mathbf{A}^{[k]} = \mathbf{A} \otimes \cdots \otimes \mathbf{A}$ stands for Kronecker power. Again using $\det(\mathbf{A}) = \mathrm{tr}\left[\det(\mathbf{A})\right]$ and property 18 with $\mathbf{A} = \mathrm{vec}\left(\boldsymbol{\Delta}^\top\right)^\top$, $\mathbf{B} = \mathrm{vec}(\boldsymbol{\Xi})^{[2]\top}$, $\mathbf{C} = (\mathbf{Q}\mathbf{Y})^{[4]}$ and $\mathbf{D} = \mathrm{vec}\left(\mathbf{I}_{[p_1^2]}\right)$ we get

$$\det\left(\mathbf{Y}^\top \mathbf{Q}^\top \boldsymbol{\Xi} \mathbf{Q}\mathbf{Y}\right) = \mathrm{vec}\left(\boldsymbol{\Delta}^\top\right)^\top \left[\mathrm{vec}\left(\mathbf{I}_{[p_1^2]}\right)^\top \otimes \mathrm{vec}(\boldsymbol{\Xi})^{[2]\top}\right] \mathrm{vec}\left[(\mathbf{Q}\mathbf{Y})^{[4]}\right] \tag{58}$$

Using property 25 we know that

$$\mathrm{vec}\left[(\mathbf{Q}\mathbf{Y})^{[4]}\right] = \mathrm{vec}\left[\mathbf{Q}^{[4]} \mathbf{Y}^{[4]}\right] \tag{59}$$

so applying property 14 with $\mathbf{A} = \mathbf{Q}^{[4]}$ and $\mathbf{B} = \mathbf{Y}^{[4]}$ we obtain

$$\mathrm{vec}\left[(\mathbf{Q}\mathbf{Y})^{[4]}\right] = \left(\mathbf{Y}^{[4]\top} \otimes \mathbf{I}_{p_1}\right) \mathrm{vec}\left(\mathbf{Q}^{[4]}\right) \tag{60}$$

Using property 24 with $\mathbf{A} = \mathbf{Q}$ and $k = 4$, we obtain

$$\mathrm{vec}\left[(\mathbf{Q}\mathbf{Y})^{[4]}\right] = \left(\mathbf{Y}^{[4]\top} \otimes \mathbf{I}_{p_1}\right) \boldsymbol{\Pi}_{p_1}^4 \mathrm{vec}(\mathbf{Q})^{[4]} \tag{61}$$

Finally we can use 61 in 58 to obtain

$$\det\left(\mathbf{Y}^\top \mathbf{Q}^\top \boldsymbol{\Xi} \mathbf{Q}\mathbf{Y}\right) = \boldsymbol{\omega}^\top \mathbf{q}^{[4]} \tag{62}$$

which is a fourth-order polynomial only with the fourth-order terms where

$$\boldsymbol{\omega}^\top = \mathrm{vec}\left(\boldsymbol{\Delta}^\top\right)^\top \left[\mathrm{vec}\left(\mathbf{I}_{[p_1]}\right)^\top \otimes \mathrm{vec}(\boldsymbol{\Xi})^{[2]\top}\right] \left(\mathbf{Y}^{[4]\top} \otimes \mathbf{I}_{[p_1]}\right) \boldsymbol{\Pi}_{p_1}^4 \tag{63}$$

□

This article was processed using the LaTeX macro package with LLNCS style