Maximizing Rigidity: Optimal Matching under Scaled-Orthography

João Maciel[1] and João Costeira[1]

Instituto de Sistemas e Robotica, Instituto Superior Tecnico, Portugal
{maciel,jpc}@isr.ist.utl.pt

Abstract. Establishing point correspondences between images is a key step for 3D-shape computation. Nevertheless, shape extraction and point correspondence are treated, usually, as two different computational processes. We propose a new method for solving the correspondence problem between points of a fully uncalibrated scaled-orthographic image sequence. Among all possible point selections and permutations, our method chooses the one that minimizes the fourth singular value of the observation matrix in the factorization method. This way, correspondences are set such that shape and motion computation are optimal. Furthermore, we show this is an optimal criterion under bounded noise conditions.
Also, our formulation takes feature selection and outlier rejection into account, in a compact and integrated way. The resulting combinatorial problem is cast as a concave minimization problem that can be efficiently solved. Experiments show the practical validity of the assumptions and the overall performance of the method.

1 Introduction

Extracting 3D-shape information from images is one of the most important capabilities of computer vision systems. In general, computing 3D coordinates from 2D images requires that projections of the same physical world point in two or more images are put to correspondence.

Shape extraction and point correspondence are treated, usually, as two different computational processes. Quite often, the assumptions and models used to match image points are unrelated to those used to estimate their 3D coordinates. On one hand, shape estimation algorithms usually require known correspondences [21], solving for the unknown shape and motion. On the other hand, image feature matching algorithms often disregard the 3D estimation process, requiring knowledge of camera parameters [15] or use other specific assumptions [18]. Furthermore, while matching algorithms tend to rely on local information — *e.g.* brightness [10,18] — shape computation algorithms [23,20,21, 19] rely on rigidity as a global scene attribute. These methods recover the rigid camera motion and object shape that best fit the data.

A. Heyden et al. (Eds.): ECCV 2002, LNCS 2351, pp. 232–246, 2002.

We present a new method that links shape computation to image feature matching by choosing the point correspondences that maximize a single global criterion — rigidity. In other words, correspondences are set such that they optimize one criterion for which we know how to compute the optimal solution for shape and motion. Also, our formulation takes feature selection and outlier rejection into account, in a compact and integrated way. This is made possible by formulating the matching process as an integer programming problem where a polynomial function — representing rigidity deviation — is minimized over the whole set of possible point correspondences. Combinatorial explosion is avoided by relaxing to continuous domain.

1.1 Previous Work

Rigidity has been used before in the correspondence framework [1,20,15,18], though used in conjunction with other assumptions about the scene or camera. The work of [22] is an example of a successful use of global geometrical reasoning to devise a pruning mechanism that is able to perform outlier rejection in sets of previously matched features. Optimality is guaranteed in a statistical sense.

Other approaches use a minimal set of correspondences which help computing the correspondences for the complete set of features [19]. This is related to using prior knowledge about camera geometry in order to impose epipolar constraints [24] or multi-view motion constraints.

Finally the approach of [4] is an example where matching and 3D reconstruction are deeply related. Correspondences, shape and motion are simultaneously optimized by an Expectation Maximization algorithm. Spurious features are not explicitly taken into account.

2 Maximizing Rigidity: Problem Statement

Consider the images of a static scene shown in Figure 1[1]. Segment p_1 feature-

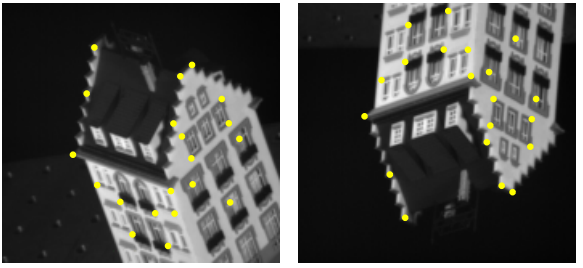

Fig. 1. Two images from the Hotel sequence, with extracted corners.

[1] Data was provided by the Modeling by Video group in the Robotics Institute, CMU

points on the first image and $p_2 > p_1$ on the second — the white dots — arrange their image coordinates u_p and v_p in two matrices $\mathbf{X}$ and $\mathbf{Y}$:

$$\mathbf{X} = \begin{bmatrix} u_1^1 & v_1^1 \\ \vdots & \vdots \\ u_{p_1}^1 & v_{p_1}^1 \end{bmatrix} \quad , \quad \mathbf{Y} = \begin{bmatrix} u_1^2 & v_1^2 \\ \vdots & \vdots \\ u_{p_2}^2 & v_{p_2}^2 \end{bmatrix} \tag{1}$$

Some of these features are projections of the same 3D points. We wish to recover their 3D coordinates assuming no prior knowledge except that the object is rigid and the camera is scaled-orthographic. To do so a selection mechanism should arrange some of the observed features in a matrix of centered measurements $\mathbf{W}$, as in [21][2]. Matched features must lie in the same row of $\mathbf{W}$. Note that no local image assumptions are made, no calibration information is known and no constraints in the disparity field are used.

Without noise, matrix $\mathbf{W}$ is, at most, rank 3, even with scale changes. We propose to solve this problem by searching for the correspondences that best generate a rank-three $\mathbf{W}$ matrix. This is only a necessary condition for rigidity, so multiple solutions minimize this criterion[3]. Within the optimization framework used — Section 4 — retrieving solutions of similar cost is trivial. The number of solutions decreases dramatically with increasing number of frames, or decreasing number of rejections. In any case, for orthographic cameras we can choose the only solution leading to an orthonormal motion matrix — the rigid solution.

With noisy measurements, $\mathbf{W}$ is allways full-rank, so we must be able to answer the following questions:

1. Is it possible to generalize the rank criterion in the presence of noise?
2. Is there any procedure to search for the best solution of this problem with reasonable time complexity?

This paper tries to give a positive answer to these questions, by formulating the correspondence problem as an optimization problem with polynomial cost function.

3 Optimal Matching

For the sake of simplicity, we start with the two-image case. Our goal is to determine a special permutation matrix $\mathbf{P}^* \in \mathcal{P}_{\mathrm{p}}^{\mathrm{c}}(p_1, p_2)$, such that $\mathbf{X}$ and $\mathbf{P}^*\mathbf{Y}$ have corresponding features in the same rows. $\mathbf{P}$ is constrained to $\mathcal{P}_{\mathrm{p}}^{\mathrm{c}}(p_1, p_2)$, the set of $p_1 \times p_2$ *columnwise partial permutation matrices* ($\mathrm{p_p}$-matrices). A $\mathrm{p_p}$-matrix is a permutation matrix with added columns of zeros. The optimal $\mathbf{P}^*$ is a zero-one variable that selects and sorts some rows of $\mathbf{Y}$, putting them to correspondence with the rows of $\mathbf{X}$. Each entry $\mathbf{P}_{i,j}$ when 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. Figure 2 shows an example. Such a matrix guarantees robustness in the presence

[2] Our $\mathbf{W}$ corresponds to their $\mathbf{W}^\top$

[3] Without noise, any object deforming according to a linear transformation in 3D space generates rank-3 projections.

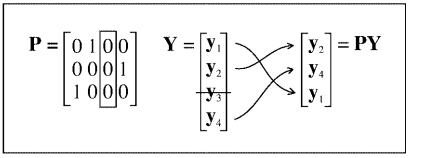

Fig. 2. A partial permutation matrix representing a particular selection and permutation of rows of $\mathbf{Y}$.

of outliers by allowing some features to be "unmatched". It encodes one way of grouping the measurements in a matrix of centered observations

$$\mathbf{W_P} \;=\; \mathbf{W}(\mathbf{P}) \;=\; [\mathbf{CX} \mid \mathbf{CPY}]_{[p_1 \times 4]} \tag{2}$$

Matrix $\mathbf{C}_{[p_1 \times p_1]} = \mathbf{I} - \frac{1}{p_1}\mathbf{1}_{[p_1 \times p_1]}$ normalizes the columns of the observation matrices to zero mean. The correct $\mathrm{p_p}$-matrix $\mathbf{P}^*$ generates $\mathbf{W}^*$ which is the measurement matrix of Tomasi-Kanade [21]. With noise-free measurements, non-degenerate full 3D objects produce rank-3 observation matrices $\mathbf{W}^\star = \mathbf{W}(\mathbf{P}^*)$ whenever $\mathbf{P}^*$ is the correct partial permutation. A single mismatched point will generate two columns of $\mathbf{W_P}$ that are outside the original 3-dimensional column-space — also called the shape-space. This makes $\mathbf{W_P}$ full-rank even in the absence of noise. In conclusion, the noise-free correspondence problem can be stated as Problem 1

Problem 1

$$\begin{aligned} \mathbf{P}^* = \arg\min_{\mathbf{P}} \quad & \mathit{rank}(\mathbf{W_P}) \\ \text{s.t.} \quad & \mathbf{P} \in \mathcal{P}_p^c(p_1, p_2) \end{aligned}$$

3.1 Approximate Rank

Consider now the case of noisy measurements. The observation matrix includes two additive noise terms $\mathbf{E}^X$ and $\mathbf{E}^Y$

$$\mathbf{W'_P} = \left[\mathbf{C}\left(\mathbf{X} + \mathbf{E}^X\right) \mid \mathbf{CP}\left(\mathbf{Y} + \mathbf{E}^Y\right)\right] \tag{3}$$

The factorization method [21] provides an efficient way of finding the best rigid interpretation of the observations. It deals with noise by disregarding all but the largest 3 singular values of $\mathbf{W'_P}$. The approximation error is measured by $\lambda_4(\mathbf{W'_P})$, the fourth singular value of $\mathbf{W'_P}$. If we use $\lambda_4(\mathbf{W'_P})$ as the generalization of the criterion of Problem 1, then we should search for the correspondence that minimizes the approximation error made by the factorization method. When noise is present, we formulate the correspondence problem as follows

Problem 2

$$\begin{aligned} \mathbf{P}^* = \arg\min_{\mathbf{P}} \quad & \lambda_4(\mathbf{W'_P}) \\ \text{s.t.} \quad & \mathbf{P} \in \mathcal{P}_p^c(p_1, p_2) \end{aligned}$$

In fact, for bounded noise, the solution to Problem 2 is again $\mathbf{P}^*$, which is the solution to Problem 1 when no noise is present. This is precisely stated in Proposition 1

Proposition 1 *For every nondegenerate rigid object observed by a scaled orthographic camera it is possible to find a scalar* $\epsilon > 0$ *such that if* $\left|\mathbf{E}^X_{(i,j)}\right| < \epsilon$ *and* $\left|\mathbf{E}^Y_{(i,j)}\right| < \epsilon \; \forall \; i,j$ *then the solution of Problem 2 is exactly* $\mathbf{P}^*$*, which is the solution of Problem 1 in the absence of noise.*

Proof: In the absence of noise, the correct $\mathbf{P}^*$ matrix generates $\mathbf{W'}^*$, and $\text{rank}\left(\mathbf{W'}^*\right) = 3 \Leftrightarrow \lambda_4\left(\mathbf{W}^*\right) = 0$. Assuming that Problem 2 has a single nondegenerate solution in the absence of noise, then there is a nonzero difference between the best and second best values of its cost. That is to say

$$\exists\, \delta > 0 \;\; : \;\; \lambda_4(\mathbf{W'}^*) + \delta < \lambda_4(\mathbf{W'_P}) \;,\; \forall \;\; \mathbf{P} \neq \mathbf{P}^* \tag{4}$$

Since $\lambda_4(\mathbf{W'_P})$ is a continuous function of the entries of $\mathbf{W'_P}$ then this is also a continuous function of the entries of $\mathbf{E}^X$ and $\mathbf{E}^Y$. By definition of continuity, $\exists\; \epsilon > 0$ such that if $|\mathbf{E}^X_{(i,j)}| < \epsilon$ and $|\mathbf{E}^Y_{(i,j)}| < \epsilon \; \forall \; i,j$ then Equation 4 still holds. This guarantees that, under these noise constraints, $\mathbf{P}^*$ is still the optimal solution to Problem 2.

Our proof for Proposition 1 does not present a constructive way to compute ϵ, so we did an empirical evaluation about the practical validity of this noise bound.

We segmented a set of points on two images. For each number of wrong matches, a set of randomly generated $\mathbf{P}$ matrices were used to compute $\lambda_4(\mathbf{W'_P})$. Figure 3 shows its statistics. The global minimum is reached for the correct correspondence $\mathbf{P}^*$, even with noisy feature locations. This shows that the bound ϵ is realistic. It also validates $\lambda_4(\mathbf{W'_P})$ as a practical criterion. Finally note that the average values of $\lambda_4(\mathbf{W'_P})$ increase monotonously with number of mismatches. This means that suboptimal solutions with objective values close to optimal will, on average, have a small number of mismatches. This is most useful to devise a stoping criterion for our search algorithm (section 4).

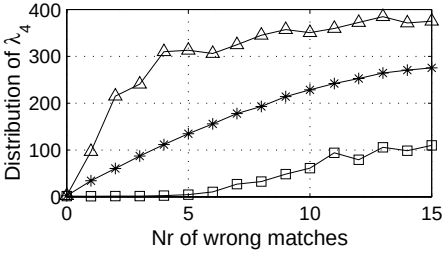

Fig. 3. Minimum, maximum and average of $\lambda_4(\mathbf{W'_P})$ as functions of the number of mismatches.

3.2 Explicit Polynomial Cost Function

We show here that the 4^{th} singular value of $\mathbf{W}$ (cost in Problem 2) has an equivalent explicit fourth-order polynomial cost function:

$$J(\mathbf{P}) = \boldsymbol{\omega}^{\top}\mathbf{q}^{[4]} \tag{5}$$

Here, $\boldsymbol{\omega}$ is a vector independent of $\mathbf{q} = \text{vec}(\mathbf{P})$, vec() is the vectorization operator[4] and $\mathbf{q}^{[4]} = \mathbf{q} \otimes \mathbf{q} \otimes \mathbf{q} \otimes \mathbf{q}$ (symbol $\otimes$ stands for Kronecker product). Considering the bounded noise assumption once more, we can use the same sort of reasoning used to prove Proposition 1 and show that the original cost function of Problem 2 can be changed to

$$J(\mathbf{P}) = \det\left(\mathbf{W'}_{\mathbf{P}}^{\top}\mathbf{W'}_{\mathbf{P}}\right) \tag{6}$$

In the absence of noise the result is immediate (they are both zero). The full proof of this result can be found in Appendix 1 (also in [12]). Figure 4 shows statistics of $\det(\mathbf{W'}_{\mathbf{P}}^{\top}\mathbf{W'}_{\mathbf{P}})$ computed for different matches of points that were segmented from real images. Once again, the global minimum is reached for the

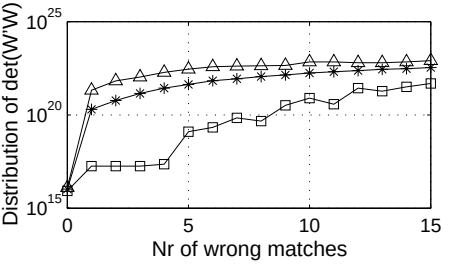

Fig. 4. Minimum, maximum and average of $\det(\mathbf{W'}_{\mathbf{P}}^{\top}\mathbf{W'}_{\mathbf{P}})$ as functions of the number of mismatches.

correct correspondence $\mathbf{P}^{*}$, so we conclude that our new noise bound is also valid in practice. Furthermore, the global minimum is even steeper — note that the plot is shown in logarithmic scale.

Finally, keeping in mind that $\mathbf{C}$ is symmetrical and idempotent and that $\mathbf{CX}$ is full rank, the determinant cost function of Equation 6 can be transformed into the polynomial form of Equation 5 (see appendix 2):

$$\begin{aligned} J(\mathbf{P}) &= \boldsymbol{\omega}^{\top}\mathbf{q}^{[4]} \\ \boldsymbol{\omega} &= \left[\left(\text{vec}(\Delta)^{\top}\mathbf{Y}^{[4]\top}\right) \otimes \left(\text{vec}(\boldsymbol{\Xi})^{[2]\top}\mathbf{I}_{p_1{}^4}\right)\right]\boldsymbol{\Pi} \\ \boldsymbol{\Xi} &= \mathbf{C} - \mathbf{CX}(\mathbf{X}^{\top}\mathbf{CX})^{-1}\mathbf{X}^{\top}\mathbf{C} \end{aligned}$$

[4] Stacks the columns of a matrix as a single column vector

$$\Delta = \frac{1}{2}\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}^{[2]}$$

$\boldsymbol{\Pi}$ is a fixed $(p_1 p_2)^4 \times (p_1 p_2)^4$ permutation such that $\text{vec}\left(\mathbf{P}^{[4]}\right) = \boldsymbol{\Pi}\text{vec}(\mathbf{P})^{[4]}$. The important fact is that $J(\mathbf{P})$ has a simple biquadratic structure, depending only on data matrices $\mathbf{X}$ and $\mathbf{Y}$.

3.3 Outline of the Complete Algorithm

The cost function of Equation 5 is ready to be formated in a global optimization framework, where the correct permutation is the solution to the following problem:

Problem 3

$$\mathbf{P}^* = \arg\min_{\mathbf{P}} \quad J(\mathbf{X}, \mathbf{PY}) \tag{7}$$

s.t.

$$\mathbf{P} \in \mathcal{P}_p^c(p_1, p_2) \iff \begin{cases} \mathbf{P}_{i,j} \in \{0,1\}\ , \quad \forall i = 1...p_1\ ,\ \forall j = 1...p_2 & .1 \\ \sum_{i=1}^{p_1} \mathbf{P}_{i,j} \leq 1\ , \forall j = 1 \dots p_2 & .2 \\ \sum_{j=1}^{p_2} \mathbf{P}_{i,j} = 1\ , \forall i = 1 \dots p_1 & .3 \end{cases} \tag{8}$$

3.4 Reformulation with a Compact Convex Domain

Problem 3 is a brute force NP-hard integer minimization problem. In general, there is no efficient way of (optimally) solving such type of problems. Nonetheless there is a related class of optimization problems for which there are efficient, optimal algorithms. Such a class can be defined as Problem 4.

Problem 4

$$\begin{aligned} \mathbf{P}^* \ = \ \arg\min_{\mathbf{P}} \quad & J_\epsilon(\mathbf{X}, \mathbf{PY}) \\ \text{s.t.} \quad & \mathbf{P} \in \mathcal{DS}_s(p_1, p_2) \end{aligned}$$

where J_ϵ is a concave version of J — to be def. later, equation 10— and $\mathcal{DS}_{\text{s}}(p_1, p_2)$ is the set of real $p_1 \times p_2$ *columnwise doubly sub-stochastic matrices*. This set is the convex hull of $\mathcal{P}_{\text{p}}^{\text{c}}(p_1, p_2)$, constructed by dropping the zero-one condition (8.1), and replacing it with

$$\mathbf{P}_{i,j} \geq 0\ ,\ \forall i = 1 \dots p_1\ ,\ \forall j = 1 \dots p_2 \tag{9}$$

Problems 3 and 4 can be made equivalent — same global optimal — by finding an adequate concave objective function J_ϵ. Also we must be sure that the vertices of $\mathcal{DS}_{\text{s}}(p_1, p_2)$ are the elements of $\mathcal{P}_{\text{p}}(p_1, p_2)$. Figure 5 summarizes the whole process.

The main idea is to transform the integer optimization problem into a global continuous problem, having exactly the same solution as the original, which minimum can be found efficiently. The full process is outlined as follows:

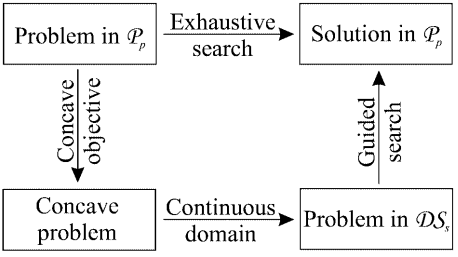

Fig. 5. Efficient solution of the combinatorial problem.

1. Extract interest points and build $\mathbf{X}$, $\mathbf{Y}$ — Equation 1.
2. Use $\mathbf{X}$, $\mathbf{Y}$ to build the cost $J(\mathbf{P})$ — Equation 5.
3. Build a concave equivalent $J_\epsilon(\mathbf{P})$ — Equation 10,11.
4. Write $\mathcal{S}_{\mathrm{s}}^{\mathrm{c}}(p_1, p_2)$ in canonical form — see section 3.6.
5. Build Problem 5 and solve it using a concave minimization algorithm — Section 4.

3.5 Equivalence of Problems 3 and 4

Theorem 1 states the fundamental reason for the equivalence. [7] contains its proof.

Theorem 1 *A strictly concave function $J : \mathcal{C} \to \mathbb{R}$ 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 $J_\epsilon : \mathcal{DS}_{\mathrm{s}}(p_1, p_2) \to \mathbb{R}$ 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 $\mathcal{P}_{\mathrm{p}}^{\mathrm{c}}(p_1, p_2)$ is $\mathcal{DS}_{\mathrm{s}}(p_1, p_2)$, and that all p_{p}-matrices are vertices of $\mathcal{DS}_{\mathrm{s}}(p_1, p_2)$, even in the presence of the rank-fixing constraint.

Consider Problem 3, where $J(\mathbf{q})$ is a class C^2 scalar function. Each entry of its Hessian is a continuous function $\mathbf{H}_{ij}(\mathbf{q})$. J can be changed to its concave version J_ϵ by

$$J_\epsilon(\mathbf{q}) = J(\mathbf{q}) + \textstyle\sum_{i=1}^{n} \epsilon_i q_i^2 - \sum_{i=1}^{n} \epsilon_i q_i \tag{10}$$

Note that the constraints of Problem 3 include $q_i \in \{0, 1\}$, $\forall i$. This means that $J_\epsilon(\mathbf{q}) = J(\mathbf{q})$, $\forall \mathbf{q}$. On the other hand $\mathcal{P}_{\mathrm{p}}(p_1, p_2)$ is bounded by a hypercube $\mathcal{B} = \{\mathbf{q} \in \mathbb{R}^n : 0 \leq q_i \leq 1,\ \forall i\}$. 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

$$\epsilon_r \leq -\tfrac{1}{2}\left(\max_{\mathbf{q}} \sum_{s=1,s\neq r}^{n} \left|\tfrac{\partial^2 J(q)}{\partial q_r \partial q_s}\right| - \min_{\mathbf{q}} \tfrac{\partial^2 J}{\partial q_r^2}\right) \tag{11}$$

which impose a negative stricly dominant diagonal to the Hessian of J_ϵ, that is to say, $|\mathbf{H}_{ii}| > \sum_{j=1,j\neq i}^{n} |\mathbf{H}_{ij}| \ , \ \forall i$. A strictly diagonally dominant matrix having only negative elements on its diagonal is strictly negative definite [6], so these values of ϵ_r will guaranty that $J_\epsilon(\mathbf{q})$ is concave for $\mathbf{q} \in \mathcal{B}$ and, therefore, also for $\mathbf{q} \in \mathcal{DS}_{\mathrm{s}}(p_1, p_2)$.

Finally, note that problem 4 is constrained to the set of doubly sub-stochastic matrices, defined by conditions 8.2, 8.3 and 9 This set has the structure of a compact convex set in $\mathbb{R}^{p_1 \times p_2}$. Its extreme points are the elements of $\mathcal{P}_{\mathrm{p}}^{\mathrm{c}}(p_1, p_2)$ — see [12] for greater details. This fact together with Theorem 1 proves that the continuous Problem 4 is equivalent to the original discrete Problem 3, since we're assuming that J_ϵ was conveniently made concave.

3.6 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 $\mathcal{DS}_{\mathrm{s}}(p_1, p_2)$ in canonical form, that is, how to state Problem 4 as

Problem 5

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

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 $\mathbf{P}$, so we use $\mathbf{q} = \mathrm{vec}(\mathbf{P})$, where vec() stacks the columns of its operand into a column vector. Condition 8.2 is equivalent to $\mathbf{P}.\mathbf{1}_{[p_2\times 1]} \leq \mathbf{1}_{[p_1\times 1]}$. Applying the vec operator [11] 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

$$\mathbf{A}_1 = \mathbf{1}_{[1\times p_2]}^{\top} \otimes \mathbf{I}_{[p_1]} \quad ; \quad \mathbf{b}_1 = \mathbf{1}_{[p_1\times 1]} \tag{12}$$

By the same token we express condition 8.3 as

$$\mathbf{1}^{\mathbf{T}}{}_{[p_1\times 1]}.\mathbf{P} \leq \mathbf{1}_{[p_1\times 1]} \wedge \mathbf{1}^{\mathbf{T}}{}_{[p_1\times 1]}.\mathbf{P} \geq \mathbf{1}_{[p_1\times 1]} \tag{13}$$

$$\mathbf{A}_2 = \mathbf{I}_{[p_2]} \otimes \mathbf{1}_{[1\times p_1]}^{\top} \quad ; \quad \mathbf{b}_2 = \mathbf{1}_{[p_2\times 1]} \tag{14}$$

The intersection of conditions 8.2 and 8.3 results on the constraints of Problem 5 with

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_1 \\ \mathbf{A}_2 \\ -\mathbf{A}_2 \end{bmatrix} \quad ; \quad \mathbf{b} = \begin{bmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \\ -\mathbf{b}_2 \end{bmatrix} \tag{15}$$

3.7 Image Sequences

With F frames, the observation matrix is

$$\mathbf{W} = [\mathbf{C}\mathbf{X}_1 \mid \cdots \mid \mathbf{C}\mathbf{P}_{F-1}\mathbf{X}_F]_{[p_1 \times 2F]} \tag{16}$$

The original Problem 1 must be extended to the $F-1$ variables $\mathbf{P}_1$ to $\mathbf{P}_{F-1}$. The obvious consequence is an increase of the dimensionality and number of constraints. Furthermore putting the cost function in explicit polynomial form is not trivial. However, $\det\left(\mathbf{W}^\top\mathbf{W}\right)$ is still a good criterion, as Figure 6 demonstrates.

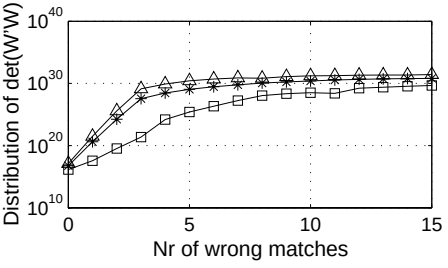

Fig. 6. Minimum, maximum and average of $\det(\mathbf{W'}_\mathbf{P}^\top\mathbf{W'}_\mathbf{P})$ as functions of the number of forced mismatches, in a 3 image example.

A slight modification in the formalism [12] makes possible to impose rejection mechanism in all images. In other words, it is possible to choose the best $p_t < p_1$ matches among all possible.

4 Minimizing Linearly Constrained Concave Functions

To minimize nonlinear concave cost functions constrained to convex sets we cannot rely on local methods, because many local minima may occur. Instead we apply *global optimization algorithms* that exploit both the concavity of the cost function and the convexity of the constraining set.

Concave programming is the best studied class of problems in global optimization [7,17], so our formulation has the advantage that several efficient and practical algorithms are available for its resolution. Among existing optimal methods, cutting-plane and cone-covering [14] provide the most efficient algorithms, but these are usually hard to implement. Enumerative techniques [16] are the most popular, mainly because their implementation is straightforward. We implemented the method of [3]. As iterations run, the *current best* solution follows an ever improving sequence of extreme points of the constraining polytope. On each iteration, global optimality is tested and a pair of upper and lower bounds are updated. Also, we use a threshold on the difference between bounds as stoping criterion. Since cost grows fast with the number of mismatches — Section 3.2 — this suboptimal strategy returns close to optimal solutions —

optimal most of the times — and dramatically reduces the number of iterations. Worst case complexity is non-polynomial but, like the *simplex* algorithm, it typically visits only a fraction of the extreme points. Our implementation takes advantage of the sparse structure of the constraints, and deals with redundancy and degeneracy using the techniques of [9].

Recently, special attention has been paid to sub-optimal concave minimization algorithms. [8] describes implementations of Frank and Wolfe [5] and Keller [8] algorithms and claims good performances in large-scale sparse problems. Simulated Annealing [2] is also having growing popularity.

5 Results

5.1 Experiment 1

Figure 7 shows two images of the Hotel sequence, with large disparity. No prior knowledge was used neither local image assumptions. Points were manually se-

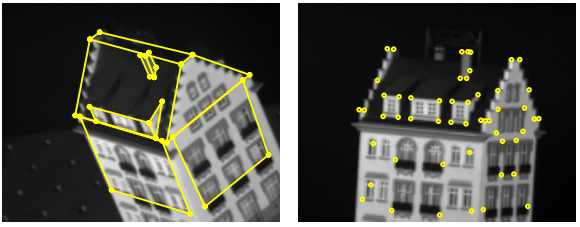

Fig. 7. Two images from the Hotel sequence, with manually segmented points.

lected in both images. In the second image, the number of points is double. The wireframe is shown just for a better perception of the object's shape. It is shown in the plot but was never used in the matching process.

The method was applied exactly as described before, using rigidity as the only criterion. Figure 8 shows the reconstruction of the matched points. The solution was found using an implementation of the optimal method of [3]. As expected, all matches are correct, corresponding to the global minimum of $\det\left(\mathbf{W}'^{\top}\mathbf{W}'\right)$.

5.2 Experiment 2

In this experiment we show how this optimal method can be used to match large points sets in a fully automatic way. At a first stage a corner detector selected a small number of points in each of 8 images from the Hotel sequence. Motion between each pair of images was then computed using the same procedure of experiment 1. An edge detector was then used to extract larger sets of points in each image. With known motion, the epipolar constraint could be used to

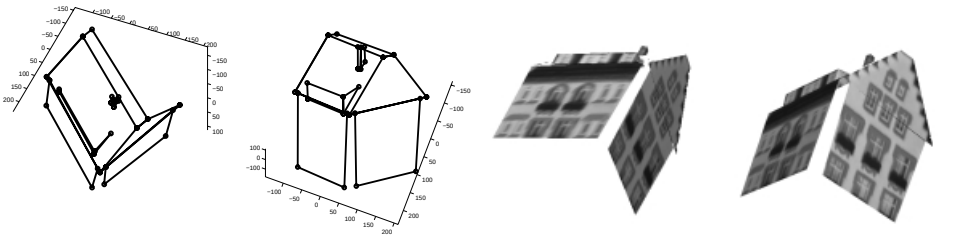

Fig. 8. Views of a reconstruction of the Hotel.

eliminate most of the matching candidates, so another global criterion — that of [13] — could be optimized with feasible computation. At the end a total of 1000 points were put into correspondence in all 8 images. Less than 10 were wrong matches, and all these were rejected by thresholding the distance between the observations and the rank-3 column space of $\mathbf{W}$. The set of remaining points was reconstructed using the factorization method. Figure 9 shows some views of the generated reconstruction.

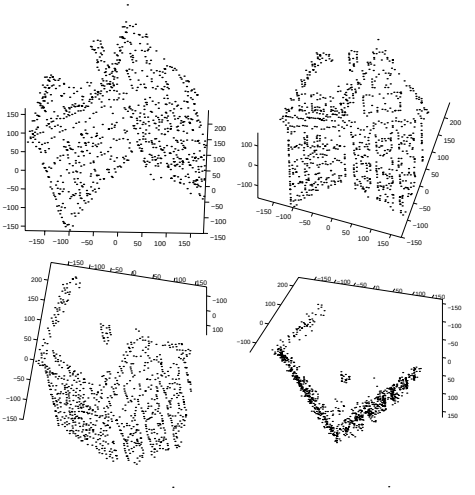

Fig. 9. Some views of an automatically generated 3D could of 900 points.

6 Conclusion

The described method solves the correspondence problem between points of a fully uncalibrated scaled-orthographic image sequence. Correspondences are set so that the shape and motion computation is optimal, by minimizing the fourth singular value of the observation matrix. We have shown that this is an optimal

criterion under bounded noise assumption. The method is also applicable to other problems where the rigidity assumption can be used, like 3D-to-3D matching, image-to-model matching and multibody factorization.

The most important limitation is the dimensionality of the resulting optimization problems. One practical way of handling this issue is the inclusion of additional *a priori* constraints — see [13] — with minor changes to the underlying problem formulation. Ongoing work is being conducted on experimenting different optimal and suboptimal algorithms, and testing their efficiency. Also, we are currently formulating and testing different ways of building explicit polynomial cost functions for multi-image matching problems.

Acknowledgments. João Maciel was supported by FCT through the Science and Technology Program of the 2nd EC Board of Support.
João Costeira was partially supported by FCT Project 3DVideo funded by program POSI-34121-SRI-2000

Appendix 1: Optimality of Rigidity Cost Functions

If Problem 1 is not degenerate, then both $\lambda_4\left(\mathbf{W'_P}\right)$ and $\det\left(\mathbf{W'_P}^{\top}\mathbf{W'_P}\right)$ are, in some sense, optimal criteria. This is stated in the following proposition:

Proposition 2 *If there is one single non-degenerate solution to Problem 1, it is possible to find a scalar* $\epsilon > 0$ *such that if* $\left|\mathbf{E}^X_{(i,j)}\right| < \epsilon$ *and* $\left|\mathbf{E}^Y_{(i,j)}\right| < \epsilon \; \forall \, i, j$ *then the solution to Problem 2 is exactly the same* $\mathbf{P}^*$ *if* $J(\mathbf{W'_P}) = \det\left(\mathbf{W'_P}^{\top}\mathbf{W'_P}\right)$. *Furthermore, this is the solution to Problem 1 without noise.*

Proof: Without noise, a unique $\mathbf{P}^*$ is also solution to Problem 1 with $J() = \lambda_4()$ or $J() = det()$ because

$$\operatorname{rank}\left(\mathbf{W}^*\right) = 3 \;\; \Leftrightarrow \;\; \lambda_4\left(\mathbf{W}^*\right) = 0 \;\; \Leftrightarrow \;\; \det\left(\mathbf{W}^{*\top}\mathbf{W}^*\right) = 0 \tag{17}$$

Non-degeneracy means that there is a nonzero difference between the best and second best cost values if $J() = \lambda_4()$ or $J() = det()$. This is to say that

$$\exists\, \delta_1 > 0 \;\; : \qquad \lambda_4(\mathbf{W}^*) + \delta_1 < \lambda_4(\mathbf{W_P}) \qquad , \; \forall \;\; \mathbf{P} \neq \mathbf{P}^* \tag{18}$$

$$\exists\, \delta_2 > 0 \;\; : \det\left(\mathbf{W}^{*\top}\mathbf{W}^*\right) + \delta_2 < \det\left(\mathbf{W_P}^{\top}\mathbf{W_P}\right) \, , \; \forall \;\; \mathbf{P} \neq \mathbf{P}^* \tag{19}$$

$\lambda_4(\mathbf{W'_P})$ and $\det\left(\mathbf{W'_P}^{\top}\mathbf{W'_P}\right)$ are continuous functions of the entries of $\mathbf{W'_P}$ so they are also continuous functions of the entries of $\mathbf{E}^X$ and $\mathbf{E}^Y$. By definition of continuity, $\exists\; \epsilon > 0$ such that if $|\mathbf{E}^X_{(i,j)}| < \epsilon$ and $|\mathbf{E}^Y_{(i,j)}| < \epsilon \; \forall \; i, j$ then equations (18) and (19) still hold for $\mathbf{W'}$. This guarantees that, under these noise constraints, $\mathbf{P}^*$ is still the optimal solution to Problem 1 with any of the two cost function.

Appendix 2: Writing the Determinant as a Polynomial Function

We will show here how to express the determinant cost function (6) as a bi-quadratic polynomial cost function

$$J_{rig}(\mathbf{P}) = \left(\mathbf{q}^\top \mathbf{B}_1 \mathbf{q}\right)\left(\mathbf{q}^\top \mathbf{B}_2 \mathbf{q}\right) - \left(\mathbf{q}^\top \mathbf{B}_3 \mathbf{q}\right)^2 \tag{20}$$

where $\mathbf{B}_i$ are matrices independent of $\mathbf{q} = \mathrm{vec}(\mathbf{P})$.

Start by using the fact that for any two matrices $\mathbf{M}_{[l\times m]}$ and $\mathbf{N}_{[l\times n]}$, if $\mathbf{L}_{[l\times(m+n)]} = [\mathbf{M} \mid \mathbf{N}]$ and $\mathbf{N}$ is full-rank — see [6,11] — then

$$\det\left(\mathbf{L}^\top \mathbf{L}\right) = \det\left(\mathbf{M}^\top \mathbf{M}\right) \det\left\{\mathbf{M}^\top \left[\mathbf{I}_{[m]} - \mathbf{N}\left(\mathbf{N}^\top \mathbf{N}\right)^{-1} \mathbf{N}^\top\right] \mathbf{M}\right\} \tag{21}$$

With $\mathbf{L} = \mathbf{W'_P} = [\mathbf{CX'} \mid \mathbf{CPY'}]$ and since $\mathbf{C}$ is symmetrical and idempotent, then

$$\det\left(\mathbf{W'_P}^\top \mathbf{W'_P}\right) = \det\left(\mathbf{X'}^\top \mathbf{C} \mathbf{X'}\right) \det\left(\mathbf{Y'}^\top \mathbf{P}^\top \boldsymbol{\Pi}^\perp \mathbf{P} \mathbf{Y'}\right) \tag{22}$$

with $\boldsymbol{\Pi}^\perp = \mathbf{C} - \mathbf{CX'}\left(\mathbf{X'}^\top \mathbf{CX'}\right)^{-1} \mathbf{X'}^\top \mathbf{C}$. Since the first determinant in equation (22) is positive and independent of $\mathbf{P}$, we can simplify the cost function, stating

$$\arg\min_{\mathbf{P}} \det\left(\mathbf{W'_P}^\top \mathbf{W'_P}\right) = \arg\min_{\mathbf{P}} \det\left(\mathbf{Y'}^\top \mathbf{P}^\top \boldsymbol{\Pi}^\perp \mathbf{P} \mathbf{Y'}\right) \tag{23}$$

Now define $\mathbf{Y'} = [\mathbf{u'}_2 \mid \mathbf{v'}_2]$, where $\mathbf{u'}_2$ and $\mathbf{v'}_2$ are respectively the row and column coordinates of points on the second image. This leads to

$$\det\left(\mathbf{W'_P}^\top \mathbf{W'_P}\right) = \det\left[\begin{array}{c|c} \mathbf{u'}_2^\top \mathbf{P}^\top \boldsymbol{\Pi}^\perp \mathbf{P} \mathbf{u'}_2 & \mathbf{u'}_2^\top \mathbf{P}^\top \boldsymbol{\Pi}^\perp \mathbf{P} \mathbf{v'}_2 \\ \hline \mathbf{v'}_2^\top \mathbf{P}^\top \boldsymbol{\Pi}^\perp \mathbf{P} \mathbf{u'}_2 & \mathbf{v'}_2^\top \mathbf{P}^\top \boldsymbol{\Pi}^\perp \mathbf{P} \mathbf{v'}_2 \end{array}\right] \tag{24}$$

For any two matrices $\mathbf{L}_{[l\times m]}$ and $\mathbf{M}_{[m\times n]}$

$$\mathrm{vec}(\mathbf{LM}) = \left(\mathbf{M}^\top \otimes \mathbf{I}_{[m]}\right) \mathrm{vec}(\mathbf{L}) \tag{25}$$

The observations $\mathbf{u'}_2$ and $\mathbf{v'}_2$ are vectors so, using equation (25), we obtain

$$\begin{aligned} \mathbf{u'}_2^\top \mathbf{P}^\top \boldsymbol{\Pi}^\perp \mathbf{P} \mathbf{v'}_2 &= \mathrm{vec}(\mathbf{P}\mathbf{u'}_2)^\top \boldsymbol{\Pi}^\perp \mathrm{vec}(\mathbf{P}\mathbf{v'}_2) \\ &= \mathbf{q}^\top \left(\mathbf{u'}_2 \otimes \mathbf{I}_{[p_2]}\right) \boldsymbol{\Pi}^\perp \left(\mathbf{v'}_2^\top \otimes \mathbf{I}_{[p_2]}\right) \mathbf{q} \end{aligned} \tag{26}$$

There are similar expressions for the other combinations of $\mathbf{u'}$ and $\mathbf{v'}$, so

$$\arg\min_{\mathbf{P}} \det\left(\mathbf{W'_P}^\top \mathbf{W'_P}\right) = \arg\min_{\mathbf{P}} \left[\left(\mathbf{q}^\top \mathbf{B}_1 \mathbf{q}\right)\left(\mathbf{q}^\top \mathbf{B}_2 \mathbf{q}\right) - \left(\mathbf{q}^\top \mathbf{B}_3 \mathbf{q}\right)^2\right] \tag{27}$$

with $\mathbf{q} = \mathrm{vec}(\mathbf{P})$ and

$$\mathbf{B}_1 = \left(\mathbf{u'}_2 \otimes \mathbf{I}_{[p_2]}\right) \boldsymbol{\Pi}^\perp \left(\mathbf{u'}_2^\top \otimes \mathbf{I}_{[p_2]}\right) \tag{28}$$

$$\mathbf{B}_2 = \left(\mathbf{v'}_2 \otimes \mathbf{I}_{[p_2]}\right) \boldsymbol{\Pi}^\perp \left(\mathbf{v'}_2^\top \otimes \mathbf{I}_{[p_2]}\right) \tag{29}$$

$$\mathbf{B}_3 = \left(\mathbf{u'}_2 \otimes \mathbf{I}_{[p_2]}\right) \boldsymbol{\Pi}^\perp \left(\mathbf{v'}_2^\top \otimes \mathbf{I}_{[p_2]}\right) \tag{30}$$

References

1. R. Berthilsson, K. Astrom, and A. Heyden. Projective reconstruction of 3d-curves from its 2d-images using error models and bundle adjustments. In *Scandinavian Conference on Image Analysis*, 1997.
2. R. Burkard, E. Çela, and Klinz B. On the biquadratic assignment problem. pages 117–146, May 1993.
3. A. Cabot and R. Francis. Solving certain nonconvex quadratic minimization problems by ranking the extreme points. *Operations Research*, 18(1):82–86, Feb 1970.
4. F. Dellaert, S. Seitz, C. Thorpe, and S. Thrun. Structure from motion without correspondence. In *Proc. CVPR*. IEEE Press, June 2000.
5. R. Fletcher. *Practical methods of optimization*. Wiley, second edition, 1987.
6. R. Horn and C. Johnson. *Matrix Analysis*. Cambridge U. Press, 1985.
7. R. Horst and P. Pardalos, editors. *Handbook of Global Optimization*. Kluwer, 1995.
8. J. Júdice and A. Faustino. Solution of the concave linear complementary problem. In *Recent Advances in Global Optimization*, pages 76–101, 1992.
9. M. Karwan, V. Lotfi, Telgen J., and S. Zionts. *Redundancy in Mathematical Programming*. Springer-Verlag, 1983.
10. B. Lucas and T. Kanade. An iterative image registration technique with app. to stereo vision. In *Proc. of the 7th International Joint Conference on AI*, 1981.
11. H. Lutkepohl. *Handbook of Matrices*. Wiley, 1996.
12. J. Maciel. *Global Matching: optimal solution to correspondence problems*. PhD thesis, Instituto Superior Técnico, 2002.
13. J. Maciel and J. Costeira. Robust point correspondence by concave minimization. In *Proc. of the 11th BMVC*, pages 626–635, Bristol, UK, 2000.
14. C. Meyer. A simple finite cone covering algorithm for concave minimization. Technical Report 193, Karl-Franzens-Universitat Graz, Apr. 2000.
15. Y. Ohta and T. Kanade. Stereo by intra- and inter-scanline search using dynamic programming. *IEEE Trans. PAMI*, 7(2):139–154, March 1985.
16. P. Pardalos. Enumerative techniques for solving some nonconvex global optimization problems. Technical Report CS–86–33, Dept. Computer Science, The Pennsylvania State University, Nov. 1986.
17. P. Pardalos and J. Rosen. Methods for global concave minimization: A bibliographic survey. *SIAM Review*, 28(3):367–379, Sep 1986.
18. S. Roy and I. Cox. A maximum-flow formulation of the n-camera stereo correspondence problem. In *Proc. ICCV*, 1997.
19. A. Shashua. Correspondence and affine shape from two orthographic views. Technical Report AIM 1327, MIT, AI Lab, 1991.
20. G. Sudhir, S. Banerjee, and A. Zisserman. Finding point correspondences in motion sequences preserving affine structure. *CVIU*, 68(2):237–246, November 1997.
21. C. Tomasi and T. Kanade. Shape from motion from image streams under orthography: a factorization method. *IJCV*, 9(2):137–154, November 1992.
22. P. Torr. *Motion Segmentation and Outlier Detection*. PhD thesis, Dept. Engineering Science, U. Oxford, 1995.
23. S. Ullman. Maximizing rigidity: the incremental recovery of 3-d structure from rigid and rubbery motion. Technical Report AIM 721, MIT, AI Lab, 1983.
24. Z. Zhang. Determining the epipolar geometry and its uncertainty - a review. *IJCV*, 27(2):161–195, 1998.