

On Computing the Translations Norm in the Epipolar Graph

Federica Arrigoni, Andrea Fusiello

DIEGM - University of Udine

Via Delle Scienze, 208 - Udine (Italy)

arrigoni.federica@spes.uniud.it, andrea.fusiello@uniud.it

Beatrice Rossi

AST Lab - STMicroelectronics

Via Olivetti, 2 - Agrate Brianza (Italy)

beatrice.rossi@st.com

Abstract

This paper deals with the problem of recovering the unknown norm of relative translations between cameras based on the knowledge of relative rotations and translation directions. We provide theoretical conditions for the solvability of such a problem, and we propose a two-stage method to solve it. First, a cycle basis for the epipolar graph is computed, then all the scaling factors are recovered simultaneously by solving a homogeneous linear system. We demonstrate the accuracy of our solution by means of synthetic and real experiments.

1. Introduction

In this paper we deal with the Epipolar Scales Computation (ESC) problem, namely the problem of recovering (up to a global scale factor) the unknown norms – also referred to as *epipolar scales* – of the relative translation directions extracted from the essential matrices. In fact, only the translation directions can be computed from the epipolar geometries, but not their norms, owing to the well-known depth-speed ambiguity. The problem can be usefully modeled by considering the *epipolar graph*, where nodes are the images and edges correspond to epipolar relationships between them.

The only solution in the literature [25] considers graphs with a special structure. This paper presents a more general technique, together with a formal analysis of the conditions under which the ESC problem is solvable.

The ESC problem finds application in Structure-from-motion (SfM), namely the problem of recovering 3D structure (of the scene) and motion (of the cameras) from point correspondences. A paradigm which is gaining increasing attention in the community consists in first computing the *relative* motion of all the cameras and then deriving their *absolute* position and angular attitude by considering the whole epipolar graph at the same time.

Almost all these global techniques [7, 15, 1, 16, 18] first solve for rotations and then for translations. The problem of

the unknown norms is bypassed either by exploiting implicit or explicit point triangulation (e.g. [1, 23, 5, 11, 15, 19]), or by solving a bearing-only network localization (e.g. [7, 4, 10, 16, 18]), where the relative translation directions expressed in an absolute frame are regarded as bearing measures that globally constraint the position of the cameras. Conditions under which positions are recoverable are studied in [18] and they refer to the concept of *parallel rigidity* [22]. The ESC problem is very related to this one, although there are some differences: in the ESC problem the input are *relative* rotations and relative translation directions, and the output are relative distances among cameras; in the bearing-only network localization problem the input are *absolute* rotations and relative translation directions, and the output are *absolute* positions of the cameras.

A different approach to global SfM consists in recovering rotations and translations *simultaneously*, by working on the manifold of rigid motions $SE(3)$. This approach, although being more principled, is less explored, probably due to the lack of a general solution to the ESC problem. Indeed, essential matrices do not fully specify elements of $SE(3)$, due to the scale ambiguity in the relative translations. The only approach of this type present in the literature is the iterative solution in [8], where the ESC problem is overlooked, though, since at each iteration the current estimates of the absolute motions are used to fix the scales of the corresponding relative translations.

In this paper we provide theoretical conditions that guarantee solvability of the ESC problem, and we propose a two-stage method to solve it. First, a cycle basis for the epipolar graph is computed, then all the scaling factors are recovered simultaneously by solving a homogeneous linear system. The key observation is that the compatibility constraints associated to cycles can be seen as equations in the unknown scales. Thus the ESC problem is cast to the resolution of a single linear system, and solvability depends on the algebraic properties of the coefficient matrix.

We consider two variants of our method, which differ for the algorithm used to obtain a cycle basis, namely computing a Fundamental Cycle Basis (FCB) or a Minimum Cycle

Basis (MCB). Experiments on synthetic and real data show that they both recover the epipolar scales accurately, and the lowest errors are obtained when using a MCB. Moreover, a MCB can be made resilient to outliers, which pays back for its higher computational cost.

The paper is organized as follows. First, we define the ESC Problem (Section 2) and we introduce the background necessary to address it (Section 3). Theoretical results about unique solvability of the problem are discussed in Section 4, and the derived method is detailed in Section 5. Finally, we evaluate the performances of our contributions to the ESC Problem via experiments on synthetic and real data (Section 6).

2. Problem Definition

Consider n pinhole cameras that capture the same (stationary) 3D scene. Let M_{ij} denote the relative transformation between cameras i and j , which can be represented as an element of the Special Euclidean Group $SE(3)$, namely the semi-direct product of the Special Orthogonal Group $SO(3)$ with \mathbb{R}^3 . Accordingly, each relative transformation can be expressed as

$$M_{ij} = \begin{pmatrix} R_{ij} & \mathbf{t}_{ij} \\ 0 & 1 \end{pmatrix} \quad (1)$$

where $R_{ij} \in SO(3)$ and $\mathbf{t}_{ij} \in \mathbb{R}^3$ respectively denote the relative rotation and translation between coordinate frames indexed by i and j .

Suppose that only some M_{ij} are known, represented by index pairs (i, j) in a set $\mathcal{E} \subseteq \{1, 2, \dots, n\} \times \{1, 2, \dots, n\}$. Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ denote the *epipolar graph* (also known as the *viewing graph* [14]), which has a vertex for each camera and edges in correspondence of the available pairwise transformations. \mathcal{G} is a directed finite simple graph with a labeling of its edge set by elements of $SE(3)$

$$\Lambda : \mathcal{E} \rightarrow SE(3), \quad \Lambda(i, j) = M_{ij} \quad (2)$$

such that if $(i, j) \in \mathcal{E}$ then $(j, i) \in \mathcal{E}$, and $\Lambda(j, i) = \Lambda(i, j)^{-1}$. Hence, \mathcal{G} may also be considered as an undirected graph. Let m denote the cardinality of \mathcal{E} , i.e. the number of edges of the underlying undirected graph.

In practice, the relative transformations M_{ij} are obtained by factorizing the essential matrices, which are computed from a collection of point matches across the input images. Each essential matrix is known up to scale due to the depth-speed ambiguity. Therefore, there is a scale ambiguity in the relative translations, i.e. what can be extracted are the relative translation *directions* $\hat{\mathbf{t}}_{ij} = \mathbf{t}_{ij} / \|\mathbf{t}_{ij}\|$. In other words, the scale factors $\alpha_{ij} = \|\mathbf{t}_{ij}\|$ of the relative translations are unknown. Note that the number of such unknowns is m since $\|\mathbf{t}_{ij}\| = \|\mathbf{t}_{ji}\|$.

The goal here is to reduce all the unknown scaling factors of the relative translations into a *single* global scaling factor, which cannot be eliminated. In other words, the present work addresses the following problem.

Epipolar Scales Computation (ESC) Problem. *Given the relative rotations $R_{ij} \in SO(3)$ and relative translation directions $\hat{\mathbf{t}}_{ij} \in \mathbb{R}^3$ for $(i, j) \in \mathcal{E}$, compute the scaling factors $\alpha_{ij} = \|\mathbf{t}_{ij}\|$ of the relative translations up to a single global scaling factor¹.*

In particular, the questions are: under which assumptions the ESC Problem admits solution, and which algorithm can solve it. A pair (\mathcal{G}, Λ) for which it is possible to solve the ESC problem is called a *solvable epipolar graph*.

3. Background

In this section we review some useful concepts from graph theory [12], and we describe the Zeller-Faugeras method [25], of which our method is a generalization.

3.1. Cycle Bases

Consider a finite simple graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where \mathcal{V} is the set of *vertices* (or *nodes*) of cardinality n and \mathcal{E} is the set of *edges* of cardinality m . If the edges are ordered pairs of vertices then \mathcal{G} is a *directed graph*, otherwise \mathcal{G} is an *undirected graph*. A *weighted graph* is a graph together with a weight function $w : \mathcal{E} \rightarrow \mathbb{R}^+$.

\mathcal{G} is called *connected* if there exists a path from each vertex to any other. \mathcal{G} is called *biconnected* if it has no articulation points, where a vertex $v \in \mathcal{V}$ is an *articulation point* if $\mathcal{G} \setminus \{v\}$ is disconnected. A graph is a *tree* if it is connected and it has $n-1$ edges. A subgraph of a connected graph \mathcal{G} is called a *spanning tree* if it has the same vertices of \mathcal{G} and it is a tree. A single spanning tree of a graph can be found in linear time $O(m+n)$ by either depth-first search or breadth-first search.

A *cycle* in an undirected graph is a subgraph in which every vertex has even degree, where the *degree* of a vertex is the number of times that the vertex occurs as the endpoint of an edge. A cycle is a *circuit* if it is connected and every one of its vertices has degree two. In this paper we use the notation $(i_1, i_2, \dots, i_{N-1}, i_N)$ to denote the N -length circuit formed by the edges $\{(i_1, i_2), (i_2, i_3), \dots, (i_{N-1}, i_N), (i_N, i_1)\}$.

If C_1, \dots, C_k are cycles of \mathcal{G} , then the *sum of cycles* $C_1 \oplus \dots \oplus C_k$ is defined as the cycle consisting of all the edges that are contained in an odd number in the cycles C_i , as illustrated in Figure 1. A *cycle basis* is a minimal set of circuits such that any cycle can be written as linear combination of the circuits in the basis. Viewing cycles as

¹Please note that when referring to a “unique” solution to the ESC problem we will include the global scale indeterminacy.

vectors indexed by edges, addition of cycles corresponds to modulo-2 sum of vectors, and the cycles of a graph form a vector space in $GF(2)^m$. The dimension of such a space is $m - n + cc$, where cc denotes the number of connected components in \mathcal{G} .

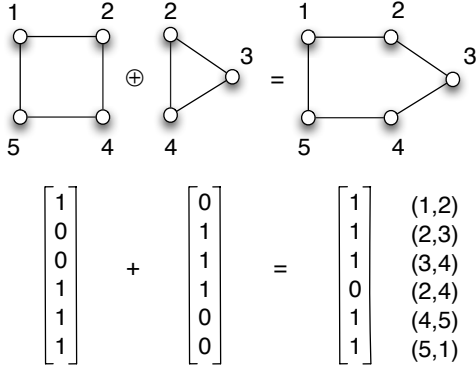


Figure 1: The sum of two cycles is a cycle where the common edges vanish.

If \mathcal{G} is connected and \mathcal{T} is any arbitrary spanning tree of \mathcal{G} , then adding any edge from $\mathcal{G} \setminus \mathcal{T}$ to \mathcal{T} will generate a circuit. The set of such circuits forms a cycle basis, which is referred to as *fundamental cycle basis* (FCB). This simple technique for extracting a cycle basis is summarized in Algorithm 1 and it runs in $O(m + n)$ time.

Algorithm 1 Spanning tree

Input: Connected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$

Output: Fundamental Cycle Basis \mathcal{B}

1. Initialize $\mathcal{B} = \emptyset$.
 2. Compute a spanning tree \mathcal{T} .
 3. **for** $(x, y) \in \mathcal{E} \setminus \mathcal{T}$ **do**
 Create the cycle $C(x, y) = P(x, y) \cup (x, y)$, where $P(x, y)$ is the shortest path in \mathcal{T} between x and y . Add $C(x, y)$ to \mathcal{B} .
 4. **end for**
-

The *length* of a cycle is either the number of edges in the cycle (in unweighted graphs) or the sum of the weights of the edges in the cycle (in weighted graphs). A *minimum cycle basis* (MCB) is a basis of total minimum length. In general a MCB is not unique. Horton's algorithm [9] finds a MCB in polynomial time, requiring at most $O(m^3n)$ steps, assuming that the underlying graph is biconnected. This method is described in Algorithm 2. The last step in Algorithm 2 can be implemented by applying Gaussian elimination to a 0, 1-matrix whose rows are the vectors in $GF(2)^m$

corresponding to the cycles generated in Step 2. Figure 2 outlines the difference between MCB and FCB.

Algorithm 2 Horton

Input: Biconnected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$

Output: Minimum Cycle Basis \mathcal{B}

1. Find the shortest path $P(x, y)$ between each pair of vertices $x, y \in \mathcal{V}$.
 2. **for** $v \in \mathcal{V}$ **do**
for $(x, y) \in \mathcal{E}$ **do**
 Create the cycle $C(v, x, y) = P(v, x) \cup P(v, y) \cup (x, y)$ and calculate its length. Degenerate cases in which $P(v, x)$ and $P(v, y)$ have vertices other than v in common can be omitted.
end for
 3. **end for**
 4. Order the cycles by increasing lengths.
 5. Initialize $\mathcal{B} = \emptyset$. Add to \mathcal{B} the next shortest cycle if it is independent from the already selected ones.
-

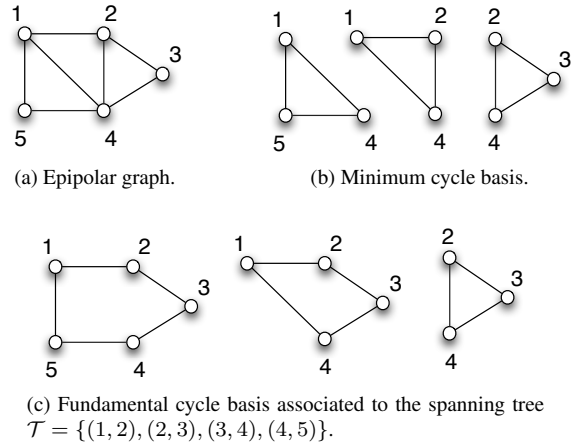


Figure 2: Example of a MCB and a FCB for a given epipolar graph. In general, the latter is composed of longer cycles.

3.2. Zeller-Faugeras method

Our method for solving the ESC Problem is inspired by [25], where the authors derive the scale factors from the composition of rigid motions.

If we consider a *sequence* of n images, whose epipolar graph is represented in Figure 3, then the following compositional rule holds

$$\mathbf{t}_{1i} = R_{12}\mathbf{t}_{2i} + \mathbf{t}_{12} \quad (3)$$

which is equivalent to

$$\alpha_{1i}\hat{\mathbf{t}}_{1i} = \alpha_{2i}R_{12}\hat{\mathbf{t}}_{2i} + \alpha_{12}\hat{\mathbf{t}}_{12}. \quad (4)$$

This leads to the following solution for the ratios of the scale factors

$$\frac{\alpha_{12}}{\alpha_{1i}} = \frac{(R_{12}\hat{\mathbf{t}}_{2i} \times \hat{\mathbf{t}}_{1i})^\top (R_{12}\hat{\mathbf{t}}_{2i} \times \hat{\mathbf{t}}_{12})}{\|R_{12}\hat{\mathbf{t}}_{2i} \times \hat{\mathbf{t}}_{12}\|^2}. \quad (5)$$

More precisely, if we arbitrarily fix the value of (e.g.) α_{12} , then we can compute the remaining scaling factors α_{1i} by using the equations above. The arbitrary choice of α_{12} corresponds to the global scaling factor, which can not be computed without external measurements.

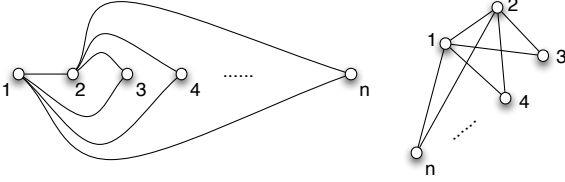


Figure 3: The epipolar graph corresponding to the Zeller-Faugeras method [25]. It is made of $n - 2$ circuits of length 3 all sharing a common edge.

Our method can be seen as an extension of this approach to general epipolar graphs, with a formal analysis of the conditions that guarantee solvability.

4. Theoretical Results

In order to address the ESC Problem, we consider the composition of pairwise motions along circuits, which must return the identity transformation.

We observe that it is impossible to solve the ESC Problem in the presence of edges not belonging to any cycle. Indeed, the norm of such edges can be chosen arbitrarily without any impact on the other scales, since they are not constrained by other edges. For this reason we assume that each edge in \mathcal{E} belongs to (at least) one cycle, namely the set of edges associated to a cycle basis coincides with \mathcal{E} itself. Such a graph is also called *bridgeless*.

4.1. A single circuit

For simplicity of exposition, we first consider the case where the epipolar graph consists of a single circuit C of length $N \geq 3$, e.g. $C = (1, 2, \dots, N - 1, N)$. The composition of the pairwise motions along C yields the 4×4 identity matrix, namely

$$M_{12}M_{23} \dots M_{N-1,N}M_{N1} = I. \quad (6)$$

Note that this equation is written by traversing the cycle in a given order (clockwise or anti-clockwise) while considering the *directed* epipolar graph. Equation (6) is called

the *compatibility constraint*, and it can also be expressed as $M_{12}M_{23} \dots M_{N-1,N} = M_{1N}$. Considering separately the rotation and translation terms, it results in

$$R_{12}R_{23} \dots R_{N-1,N} = R_{1N} \quad (7)$$

$$\alpha_{12}\hat{\mathbf{t}}_{12} + \sum_{k=2}^{N-1} \left(\prod_{i=1}^{k-1} R_{i,i+1} \right) \alpha_{k,k+1} \hat{\mathbf{t}}_{k,k+1} = \alpha_{1N} \hat{\mathbf{t}}_{1N} \quad (8)$$

where the relation between translations can be viewed as a homogeneous linear equation in the unknown scales. Note that the Zeller-Faugeras method considered the compatibility constraint for $N = 3$.

Equation (8) can also be expressed in terms of differences between the camera centers (i.e. the baselines), if the absolute rotations of the cameras are known. Let $R_1, \dots, R_n \in SO(3)$ denote the absolute rotations, let $\mathbf{b}_{ij} \in \mathbb{R}^3$ denote the baseline joining the optical centers of cameras i and j , and let $\hat{\mathbf{b}}_{ij} \in \mathbb{R}^3$ denote the versor of the baseline \mathbf{b}_{ij} . Using this additional information, the product of relative rotations in (8) reduces to $R_1 R_k^\top$. Indeed, the link between relative and absolute rotations is encoded by the formula $R_{ij} = R_i R_j^\top$, thus all the factors in (8) simplify except of the first and the last one. By multiplying both sides by $-R_1^\top$, we obtain

$$-\alpha_{12}R_1^\top \hat{\mathbf{t}}_{12} - \sum_{k=2}^{N-1} \alpha_{k,k+1} R_k^\top \hat{\mathbf{t}}_{k,k+1} = -\alpha_{1N} R_1^\top \hat{\mathbf{t}}_{1N} \quad (9)$$

which coincides with

$$\sum_{k=1}^{N-1} \alpha_{k,k+1} \hat{\mathbf{b}}_{k,k+1} = \alpha_{1N} \hat{\mathbf{b}}_{1N} \quad (10)$$

since the baselines are related to the relative translations through the formula $\mathbf{b}_{ij} = -R_i^\top \mathbf{t}_{ij}$. Note that the baseline versor can be viewed as the direction of the relative translation expressed in the absolute reference frame. For this reason, we can also regard $\hat{\mathbf{b}}_{ij}$ as the *bearing* of camera j as seen from camera i .

We now discuss under which conditions Equation (8) gives means to compute the unknown scaling factors α_{ij} (up to a global scale). Let $A \in \mathbb{R}^{3 \times N}$ be the coefficient matrix associated to Equation (8), whose entries depend on the relative rotations and translation directions, and let $\alpha \in \mathbb{R}^N$ be the stack of the scales α_{ij} . Using this notation, the compatibility constraint reduces to a homogeneous linear system of the form $A\alpha = 0$. Thus the ESC Problem admits a unique non-trivial solution – that corresponds to the one-dimensional null space of A – if and only if $\text{rank}(A) = N - 1$. Moreover, we have $\text{rank}(A) \leq 3$, since A is a $3 \times N$ matrix. Thus, in a circuit of length N the ESC problem can be solved uniquely only if $N \leq 4$.

Observe that specific motions cause $\text{rank}(A)$ to drop. In particular, $\text{rank}(A) = 1$ if and only if the camera centers

are collinear and $\text{rank}(A) = 2$ if and only if the camera centers lie on a common plane.

This implies that for $N = 4$ the ESC problem has a unique solution provided that the cameras are in a general position, otherwise multiple solutions are possible. On the contrary, for $N = 3$, the camera centers *must* be coplanar (as it is the case if we assume correct measurements), otherwise the problem admits only the trivial solution $\alpha = 0$. If the centers are collinear then multiple solutions arise.

4.2. A generic epipolar graph

We now consider a generic epipolar graph, providing conditions for the ESC Problem to admit a unique solution. We have just shown that – if \mathcal{G} is formed by a single circuit – it is possible to recover the epipolar scales if and only if its length is 3 or 4 (provided that the cameras are in a general configuration). Thus a circuit of length $N \geq 5$ is not solvable *alone*, because the associated linear system yields multiple solutions. However, when several cycles are considered in a generic epipolar graph, it might be possible to recover the scaling factors also in the presence of circuits of length $N \geq 5$.

To see this, consider the case of Figure 4a. The key observation is that the 5-length circuit has two edges in common with a solvable subgraph of \mathcal{G} . Specifically, the epipolar scales can be recovered as follows by considering the circuits $(1, 6, 2)$, $(2, 6, 7)$, $(2, 7, 3)$ and $(1, 2, 3, 4, 5)$. First, we arbitrarily choose the scaling factor of an edge of the circuit $(1, 6, 2)$, and compute the remaining scales by solving the associated linear system, which has a unique solution since it has length 3. This cycle shares the edge $(2, 6)$ with the 3-length circuit $(2, 6, 7)$. We use such an edge to fix the global scaling factor of $(2, 6, 7)$, and solve for the remaining scales. The same happens when considering the 3-length circuit $(2, 7, 3)$. In this way the scales of the edges $(1, 2)$ and $(2, 3)$ are already determined when considering the 5-length circuit $(1, 2, 3, 4, 5)$, and only 3 unknowns remain, which can be recovered as in a circuit of length 4.

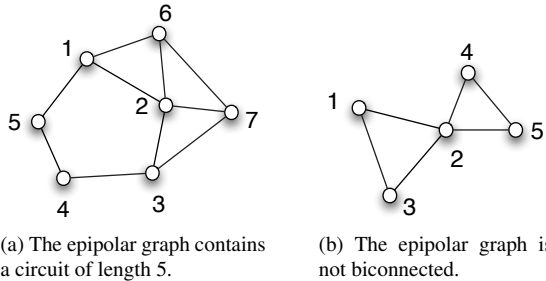


Figure 4: Examples of a solvable epipolar graph (left) and of an unsolvable epipolar graph (right).

An example of an unsolvable epipolar graph is reported in Figure 4b, where the graph is not biconnected. The cir-

cuits $(1,2,3)$ and $(2,4,5)$ do not have any edge in common, thus we can solve separately the ESC Problem for each circuit, but two unknowns remain which can not be reconciled to a single global scaling factor. It is straightforward to see that this generalizes to all the situations where articulation points are present, as mentioned also *en-passant* in [16]. In other words, the following proposition holds.

Proposition 1. *The ESC Problem admits a unique solution only if the epipolar graph is biconnected.*

Note that the requirement of being biconnected avoids both the situation of Figure 4b and the presence of edges not belonging to any cycle (biconnected \Rightarrow bridgeless).

It is straightforward to see that the necessary condition of Proposition 1 is not sufficient. (For instance, a single 5-length circuit is biconnected but the associated linear system admits multiple solutions). However, it gives a simple condition to detect non solvable graphs. Accordingly, if the epipolar graph is not biconnected, then our analysis applies to the largest biconnected component of \mathcal{G} .

We now provide a necessary and sufficient condition for the ESC Problem to admit solution. Let r be the total number of circuits present in the graph \mathcal{G} . Each circuit gives rise to a homogeneous linear equation of the form (8). All these equations can be stacked together to form a matrix A of dimensions $3r \times m$, whose entries depend on the relative rotations and translation directions. Each triplet of rows in A corresponds to a circuit, while each column corresponds to a relative translation. In this way all the edges are considered (since each edge belongs to at least one cycle by assumption) and all the existing constraints on the scales are taken into account (since we are considering all the circuits).

Thus the ESC Problem is equivalent to the resolution of a single homogeneous linear system

$$A\alpha = 0 \quad (11)$$

where $\alpha \in \mathbb{R}^m$ is the stack of the scaling factors α_{ij} . In other terms, unique solvability depends on the algebraic properties of the coefficient matrix A . More precisely, the ESC problem admits a unique solution if and only if $\text{nullity}(A) = 1$, i.e. if and only if $\text{rank}(A) = m - 1$. Such a solution is the 1-dimensional null-space of A , and it can be found by computing the eigenvector with zero eigenvalue of the matrix $A^T A$. This discussion is summarized in the following proposition.

Proposition 2. *Let $A \in \mathbb{R}^{3r \times m}$ be the coefficient matrix constructed by stacking the compatibility constraints associated to all the circuits in \mathcal{G} , where r is the number of such circuits. Let $\alpha \in \mathbb{R}^m$ be the stack of the scales α_{ij} . The ESC problem admits a unique (non-trivial) solution if and only if $\text{rank}(A) = m - 1$.*

Note that if an articulation point is present – as in the case of Figure 4b – then the matrix A can be partitioned into two independent blocks having both rank maximum minus 1. Thus the rank of the whole matrix is $m - 2$, i.e. the ESC Problem admits multiple solutions, according to Proposition 1.

Local vs global frames. Equation (11) can also be written in terms of the baselines (or bearings), generalizing Equation (10). In fact, the equation provided by a circuit C_k can be expressed as

$$B \text{diag}(\mathbf{c}_k^\top) \alpha = \mathbf{0} \quad (12)$$

where \mathbf{c}_k is the $m \times 1$ indicator vector of the circuit C_k , and B is a $3 \times m$ matrix whose columns are the baseline versors. Please note that this equation is written by traversing C_k in an arbitrary cyclic order (clockwise or anti-clockwise), hence the entries of \mathbf{c}_k have a sign that indicates whether the corresponding edge is traversed along the direction specified by its versor (the k -th column of B), or not.

Equivalently, we can use the Khatri-Rao [13] product \odot and write

$$(\mathbf{c}_k^\top \odot B) \alpha = \mathbf{0}. \quad (13)$$

In this way we can stack the equations coming from $r > 1$ circuits, obtaining

$$(C \odot B) \alpha = \mathbf{0} \quad (14)$$

where C is the $r \times m$ stack of the rows \mathbf{c}_k^\top .

Please observe that the matrix $(C \odot B)$ is *not* equal to A , but it has the same size and the same null-space (in the noise-free case). Each row in A is of the form

$$(\mathbf{c}_k^\top \odot -R_k B) \alpha = -R_k (\mathbf{c}_k^\top \odot B) \alpha = \mathbf{0} \quad (15)$$

where R_k is a rotation that takes into account the fact that in each circuit an arbitrary local reference system has been considered. Hence, there exists a choice of rotations R_1, \dots, R_m such that

$$\begin{bmatrix} -R_1(\mathbf{c}_1^\top \odot B) \\ \vdots \\ -R_m(\mathbf{c}_m^\top \odot B) \end{bmatrix} = \begin{bmatrix} -R_1 & & \\ & \ddots & \\ & & -R_m \end{bmatrix} (C \odot B) = A. \quad (16)$$

In summary, the equations involving the bearings and those involving the relative motions are equivalent in terms of constraints on the solution, however they configure two different approaches. The bearings in (12) require to compute the absolute rotations *before* the scale factors, and the problem gets very close to the bearing-only network localization. On the other hand, the equations in (11) – which give our solution to the ESC problem – are written with

respect to independent local frames, thereby avoiding the need to solve for the absolute rotations beforehand.

Nevertheless, it might be sometime useful to express the constraints in the “bearing form”, for it simplifies the discussion, as in the following paragraph.

How many circuits? As a matter of fact, considering *all* the circuits is redundant. The following result states that what is actually required is a set of *independent* circuits.

Proposition 3. *Let C_1, C_2, C_3 be three circuits in \mathcal{G} that satisfy $C_1 \oplus C_2 = C_3$. Then the equation obtained from the circuit C_3 is a linear combination of the equations obtained from the circuits C_1 and C_2 .*

Let us consider two circuits C_1 and C_2 that share one or more edges, and let $\mathbf{c}_1, \mathbf{c}_2$ be their *signed* indicator vectors. The sum of the equations derived from C_1 and C_2 writes

$$B \text{diag}(\mathbf{c}_1^\top + \mathbf{c}_2^\top) \alpha = \mathbf{0}. \quad (17)$$

Without loss of generality let us assume that C_1 and C_2 are traversed with a cyclic order such that the common edges to C_1 and C_2 are traversed in opposite directions, as in the case of Figure 5. Thanks to this assumptions the entries corresponding to common edges vanishes in $(\mathbf{c}_1^\top + \mathbf{c}_2^\top)$, and this is exactly the *signed* indicator vector of $C_1 \oplus C_2$.

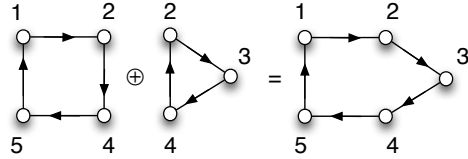


Figure 5: Sum of two circuits where the edge in common is traversed in opposite directions.

Thus, including a circuit which is the sum of other circuits does not add any independent constraint on the scaling factors.

5. Proposed Method

An immediate consequence of Proposition 3 is that we can consider a *cycle basis* rather than the set of all the circuits in Equation (11). Thus the epipolar scales can be recovered through the following steps.

1. Compute a cycle basis \mathcal{B} for the epipolar graph by using either Algorithm 1 or Algorithm 2.
2. Construct the $3(m - n + 1) \times m$ coefficient matrix A by stacking the compatibility constraints associated to the cycles in \mathcal{B} . If $\text{rank}(A) = m - 1$ then compute the unknown scales by solving system (11). Otherwise,

it is impossible to find a unique solution to the ESC Problem.

In this way, all the translation norms are recovered simultaneously (up to a global scale) by solving a single homogeneous linear system. Note that in order to guarantee solvability of the ESC problem, the number of rows in A must be greater than (or equal to) $m - 1$, i.e. the following necessary condition must be satisfied

$$m \geq \frac{3}{2}n - 2. \quad (18)$$

In the presence of noise unique solvability reduces to test if A has approximately rank $m - 1$. In this case, system (11) is solved in the least-squares sense, by computing the least eigenvector of the matrix $A^T A$, or – equivalently – the least right singular vector in the Singular Value Decomposition (SVD) of A .

Note that system (11) is *sparse*, since each row contains exactly N non-zero entries, if N is the length of the current circuit. Thus employing sparse eigen-solvers (such as MATLAB `eigs`) increases the efficiency of the method.

Which cycle basis? In the ideal (noise-free) case any cycle basis returns the desired solution. Thus the question is which basis is more suitable to our application when relative rotations and translation directions are corrupted by noise and outliers.

Intuitively, the performances of our method with respect to noise are better when using the shortest circuits, because this limits error accumulation. Therefore a MCB (Algorithm 2) should be preferred, because a MCB is characterized by the property that no circuit can be the sum of shorter circuits [21]. This does not hold for a fundamental cycle basis, which in general is composed of longer circuits.

Another advantage of using Algorithm 2 is that it can easily incorporate robustness to outliers among relative motions. Specifically, we take advantage of the redundancy of circuits generated in Step 2, without increasing the computational cost.

We say that a circuit in the epipolar graph is *null* if the composition of the relative rotations along it is equal to the identity. Non null circuits arise when one or more edges are outliers, they provide inconsistent constraints on the epipolar scales, and thus they cannot be part of the cycle basis. As a consequence, we modify Algorithm 2 by considering only *null circuits* in Step 2, while the remaining cycles are discarded.

In particular, a circuit $C = (1, 2, \dots, N - 1, N)$ generated in Step 2 of Algorithm 2 is kept if the following condition is satisfied

$$d(R_{12}R_{23} \dots R_{N-1,N}R_{N1}, I) \leq \epsilon\sqrt{N} \quad (19)$$

where $d(\cdot, \cdot) : SO(3) \times SO(3) \mapsto \mathbb{R}^+$ is a bi-invariant metric and ϵ is a given threshold. Note that this is an heuristic for finding a cycle basis for a *consistent* subgraph of \mathcal{G} , i.e. a subgraph containing only null cycles. Thus the number of circuits returned by this version of Algorithm 2 will be lower than $m - n + 1$, in general.

Compared to Algorithm 1, Horton’s algorithm has a higher computational cost, but this is balanced by increased accuracy and the possibility to discard outliers *while* computing the cycle basis. In contrast, Algorithm 1 generates only a minimum set of circuits, thus robustness can be achieved only by rejecting outliers *before* computing the basis. Available approaches for detecting outliers include [24, 17, 6, 16, 3, 23]. These techniques are computationally demanding and speed is always traded off with accuracy. Moreover, some of them [24, 6, 16, 3] are based anyway on the detection of non-null cycles.

6. Experiments

In this section we evaluate our approach on synthetic and real data, analyzing both accuracy and robustness to outliers. All the experiments are performed in Matlab on a dual-core 1.3 GHz machine. The code is available at www.diegm.uniud.it/fusiello/demo/gmf/.

Synthetic Data. We consider $n = 100$ cameras where absolute rotations $R_i \in SO(3)$ are sampled from random Euler angles, and the x, y, z -components of absolute translations $\mathbf{t}_i \in \mathbb{R}^3$ follow a standard Gaussian distribution. The edge set \mathcal{E} of the epipolar graph is sampled at random. The available pairwise motions are computed as $R_{ij} = R_i R_j^T$ and $\mathbf{t}_{ij} = -R_i R_j^T \mathbf{t}_j + \mathbf{t}_i$. All the instances simulated correspond to solvable epipolar graphs. The relative translation directions $\mathbf{t}_{ij}/\|\mathbf{t}_{ij}\|$ are corrupted by noise considering their representation in spherical coordinates, so as to remain on the unit sphere. Specifically, the spherical angles are corrupted by additive Gaussian noise with zero mean and standard deviation $\sigma \in [0.5^\circ, 5^\circ]$. The same perturbation is applied to the relative rotations, considering the angle-axis representation of $SO(3)$. All the results are averaged over 10 trials.

Theoretically, the estimated scales $\tilde{\alpha}$ should coincide with the ground truth ones α up to a multiplicative constant $s \in \mathbb{R}$, namely $\alpha = s\tilde{\alpha}$. We estimate such a constant in the least-squares sense, and we divide the mean of the residuals $r_{ij} = |\alpha_{ij} - s\tilde{\alpha}_{ij}|$ by the mean of α , to obtain a *relative mean error* on the scaling factors.

Figure 6 reports the relative mean errors on the epipolar scales as a function of σ . In this experiment we evaluate both Algorithm 1, in which a fundamental cycle basis (FCB) is extracted, and Algorithm 2, in which a minimum cycle basis (MCB) is computed. The former is highly de-

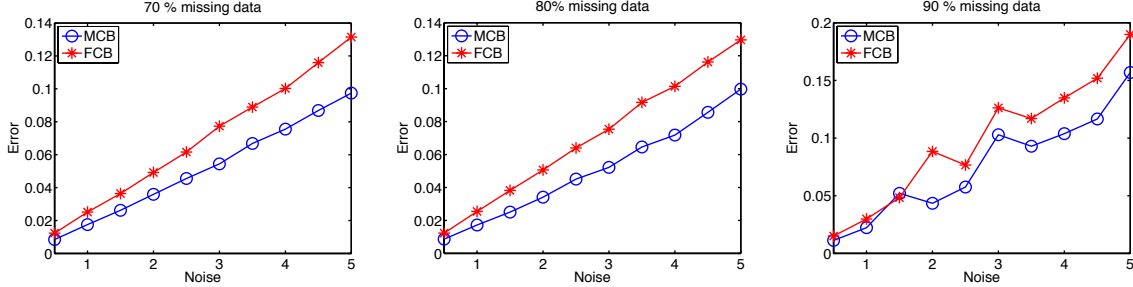


Figure 6: Relative mean error on the scale factors vs standard deviation of noise, for different percentages of missing pairs.

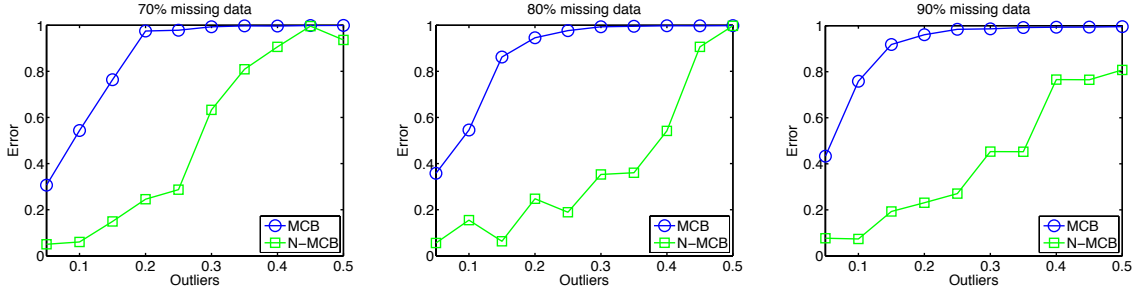


Figure 7: Relative mean error on the scale factors vs fraction of outliers, for different percentages of missing pairs.

pendent on the chosen spanning tree, thus for each trial we further average the results over 10 spanning trees simulated at random.

Both our methods give an accurate solution to the ESC Problem as noise increases, however the best resilience to noise is achieved by the MCB, as conjectured in the previous section. In the case of 90% of missing data (right subfigure) the graph is very sparse and the effect of randomness is amplified, thus producing irregular lines.

We now study the resilience to outliers of our variant of Algorithm 2 – henceforth dubbed “Null MCB” (N-MCB) – in which only null cycles are kept in Step 2. In this experiment we consider a fixed level of noise ($\sigma = 3^\circ$), while the fraction of wrong relative motions – randomly generated – varies from 5% to 50%. This percentage refers to the available pairwise motions (not to the complete epipolar graph), i.e. the number of outliers is a fraction of m .

Figure 7 reports the relative mean errors on the epipolar scales as a function of the fraction of outliers, obtained by MCB and N-MCB. While the former is non robust, the latter shows good resilience to rogue input, confirming the effectiveness of our heuristic for outlier handling. In particular, the lines corresponding to MCB converge to one since the scale factors converge to zero, thus indicating a complete failure. The lines of N-MCB are irregular due to the randomness of the data, which is amplified by the presence of both outliers and a high level of missing data.

In this experiment we also analyze the performance of N-MCB in terms of misclassification rate, which is the fraction

Table 1: Relative mean errors on the scale factors.

	% missing	FCB	MCB	N-MCB
Castle-P30	60	0.0990	0.0572	0.0326
Castle-P19	43	0.1872	0.0707	0.0359
Entry-P10	18	0.0402	0.0400	0.0124
Fountain-P11	2	0.0024	0.0017	0.0017
HerzJesu-P25	62	0.0808	0.0312	0.0044
HerzJesu-P8	18	0.0040	0.0108	0.0108
Average		0.0689	0.0353	0.0163

of effective outliers that are not removed. In all the trials we obtain a misclassification rate below 5%, thus our heuristic performs well as an outlier detector.

Real Data. We now consider the EPFL benchmark [20], a small-size real image dataset for which ground-truth motion is provided. From this the ground-truth scales can be easily computed, and they range from 0.7 to 43 meters. The relative rotations and translation directions are obtained following a standard approach based on the essential matrix factorization with a final bundle adjustment of camera pairs.

Table 1 shows the results obtained by all the variants of our method, namely FCB, MCB and N-MCB with threshold $\epsilon = 2^\circ$. As in the case of simulated data, they all recover the translations norm accurately, and the best precision is achieved, on the average, by N-MCB.

7. Conclusion

In this paper we have presented an in-depth study of the ESC problem, within the broader context of global structure from motion. After having provided theoretical conditions under which such a problem has a unique solution, we have presented an efficient algorithm to find it. The accuracy of our solution for computing the scaling factors has been demonstrated by means of synthetic and real experiments.

This method, in combination with a motion synchronization technique that works in $SE(3)$ [2], constitutes the core of a global structure-from-motion pipeline that will be characterized experimentally in future work.

On the theoretical side, we will explore the connection of our notion of ESC solvability and analogous concepts linked to the parallel (or bearing) rigidity [22, 26]. In this context we also aim at clarifying the separate role of C and B in Equation (14) in determining solvability.

References

- [1] M. Arie-Nachimson, S. Z. Kovalsky, I. Kemelmacher-Shlizerman, A. Singer, and R. Basri. Global motion estimation from point matches. In *International Conference on 3D Imaging, Modeling, Processing, Visualization and Transmission*, pages 81 – 88, 2012.
- [2] F. Arrigoni, A. Fusiello, and B. Rossi. Spectral motion synchronization in $SE(3)$. *ArXiv e-prints*, 1506.08765, 2015.
- [3] F. Arrigoni, B. Rossi, F. Malapelle, P. Fragneto, and A. Fusiello. Robust global motion estimation with matrix completion. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XL-5:63–70, 2014.
- [4] M. Brand, M. Antone, and S. Teller. Spectral solution of large-scale extrinsic camera calibration as a graph embedding problem. In *Proceedings of the European Conference on Computer Vision*, pages 262 – 273, 2004.
- [5] D. Crandall, A. Owens, N. Snavely, and D. P. Huttenlocher. Discrete-continuous optimization for large-scale structure from motion. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3001 – 3008, 2011.
- [6] O. Enqvist, F. Kahl, and C. Olsson. Non-sequential structure from motion. In *Eleventh Workshop on Omnidirectional Vision, Camera Networks and Non-classical Camera*, pages 264 – 271, 2011.
- [7] V. M. Govindu. Combining two-view constraints for motion estimation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 218 – 225, 2001.
- [8] V. M. Govindu. Lie-algebraic averaging for globally consistent motion estimation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 684 – 691, 2004.
- [9] J. D. Horton. A polynomial-time algorithm to find the shortest cycle basis of a graph. *SIAM Journal on Computing*, 16(2):358 – 366, 1987.
- [10] N. Jiang, Z. Cui, and P. Tan. A global linear method for camera pose registration. In *Proceedings of the International Conference on Computer Vision*, pages 481 – 488, 2013.
- [11] F. Kahl and R. Hartley. Multiple-view geometry under the l_∞ -norm. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(9):1603–1617, 2008.
- [12] T. Kavitha, C. Liebchen, K. Mehlhorn, D. Michail, R. Rizzi, T. Ueckerdt, and K. Zweig. Cycle bases in graphs: Characterization, algorithms, complexity, and applications. *Computer Science Review*, 3(4):199 – 243, 2009.
- [13] C. G. Khatri and C. R. Rao. Solutions to some functional equations and their applications to characterization of probability distributions. *Sankhy: The Indian Journal of Statistics, Series A (1961-2002)*, 30(2):pp. 167–180, 1968.
- [14] N. Levi and M. Werman. The viewing graph. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 518 – 522, 2003.
- [15] D. Martinec and T. Pajdla. Robust rotation and translation estimation in multiview reconstruction. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1 – 8, 2007.
- [16] P. Moulon, P. Monasse, and R. Marlet. Global Fusion of Relative Motions for Robust, Accurate and Scalable Structure from Motion. In *Proceedings of the International Conference on Computer Vision*, pages 3248 – 3255, 2013.
- [17] C. Olsson and O. Enqvist. Stable structure from motion for unordered image collections. In *Proc. of the Scandinavian conference on Image analysis*, pages 524–535, 2011.
- [18] O. Ozyesil, A. Singer, and R. Basri. Stable camera motion estimation using convex programming. *SIAM Journal on Imaging Sciences*, 8(2):1120 – 1262, 2015.
- [19] S. N. Sinha, D. Steedly, and R. Szeliski. A multi-stage linear approach to structure from motion. In *Proc. of the European Conference on Computer Vision*, pages 267 – 281, 2010.
- [20] C. Strecha, W. von Hansen, L. J. V. Gool, P. Fua, and U. Thoennessen. On benchmarking camera calibration and multi-view stereo for high resolution imagery. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1 – 8, 2008.
- [21] P. Vismara. Union of all the minimum cycle bases of a graph. *Electronic Journal of Combinatorics*, 4(1), 1997.
- [22] W. Whiteley. Matroids from discrete geometry. In J. Bonin, J. Oxley, and B. Servatius, editors, *Matroid Theory*, AMS Contemporary Mathematics, pages 171–313. 1997.
- [23] K. Wilson and N. Snavely. Robust global translations with 1DSfM. In *Proceedings of the European Conference on Computer Vision*, pages 61–75, 2014.
- [24] C. Zach, M. Klopschitz, and M. Pollefeys. Disambiguating visual relations using loop constraints. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1426 – 1433, 2010.
- [25] C. Zeller and O. Faugeras. Camera self-calibration from video sequences: the Kruppa equations revisited. Research Report 2793, INRIA, 1996.
- [26] S. Zhao and D. Zelazo. Bearing-only network localization: Localizability, sensitivity, and distributed protocols. *ArXiv e-prints*, 1502.00154, 2015.