

On the Solvability of Viewing Graphs

Matthew Trager^{1,2}, Brian Osserman³, and Jean Ponce¹

¹ Inria

² École Normale Supérieure, CNRS, PSL Research University

³ UC Davis

Abstract. A set of fundamental matrices relating pairs of cameras in some configuration can be represented as edges of a “viewing graph”. Whether or not these fundamental matrices are generically sufficient to recover the global camera configuration depends on the structure of this graph. We study characterizations of “solvable” viewing graphs, and present several new results that can be applied to determine which pairs of views may be used to recover all camera parameters. We also discuss strategies for verifying the solvability of a graph computationally.

Keywords: Viewing graph, fundamental matrix, 3D reconstruction.

1 Introduction

Multi-view geometry has been studied by photogrameters since the 1950s [1] and by computer vision researchers since the 1980s [2]. Still, most results to date are concerned with using multi-view tensors to characterize feature correspondences in 2, 3, or 4 views, and determine the corresponding projection matrices [3–6]. Although correspondences have also been characterized for arbitrary numbers of views [7–9], very little theoretical work has been devoted to understanding the geometric constraints imposed on configurations of $n > 4$ cameras by these tensors, including fundamental matrices [3], which are probably by far the most used in practice. Apart from a few works such as [10, 11], understanding how many and which fundamental matrices can be used to recover globally consistent camera parameters is a largely unexplored problem.

We address this topic in this paper. Following [10], we associate sets of fundamental matrices with edges of a “viewing graph”, and present a series of new results for determining whether a graph is “solvable”, *i.e.*, whether it represents fundamental matrices that determine a unique camera configuration. We also describe effective strategies for verifying solvability, and include some computational experiments using these methods. Our focus here is clearly of a theoretical nature, but understanding how subsets of fundamental matrices constrain the reconstruction process is clearly important in practice as well. Moreover, we will point out that one of our main results (Theorem 3) is constructive, and could potentially find applications in reconstruction algorithms (*e.g.*, it could be incorporated in large-scale systems such as [12], that incrementally build up networks of cameras to estimate their parameters).

Previous work. The first investigation of viewing graphs and their solvability can be found in [10]. In that work, Levi and Werman characterize all solvable viewing graphs with at most six vertices, and discuss a few larger solvable examples. Although they provide some useful necessary conditions (see our Proposition 2 and Example 2), they do not address the problem of solvability in general. In [11], Rudi *et al.* also consider viewing graphs, studying mainly whether a configuration can be recovered from a set of fundamental matrices using a linear system. They also present some “composition rules” for merging solvable graphs into larger ones. Trager *et. al* [13] provide a sufficient condition for solvability using $2n - 3$ fundamental matrices, and point to a possible connection with “Laman graphs” and graph rigidity theory. Indeed, Özyesil and Singer [14] show that if one uses *essential* matrices instead of fundamental ones then solvability can be characterized in terms of so-called “parallel-rigidity” for graphs. Their analysis however does not carry over to the more general setting of uncalibrated cameras. Finally, the viewing graph has also been considered in more practical work: for example, in [15, 16], it is used to enforce triple-wise consistency among fundamental matrices before estimating camera parameters.

Main contributions.

- We show that the minimum number of fundamental matrices that can be used to recover a configuration of n cameras is always $\lceil (11n - 15)/7 \rceil$ (Theorem 1).
- We present several criteria for deciding whether or not a viewing graph is solvable. After revisiting some results from [10, 11] (Section 3.1), we describe a new necessary condition for solvability that is based on the number of edges and vertices of subgraphs (Theorem 2), as well as a sufficient condition based on “moves” for adding new edges to a graph (Theorem 3).
- We describe an algebraic formulation for solvability that in principle can always be used to determine whether any viewing graph is solvable. Although this method is computationally challenging for larger graphs, we also introduce a much more practical *linear* test, that can be used to verify whether a viewing graph identifies a finite number of camera configurations (Section 4).
- Using an implementation of all the proposed methods, we analyze solvability for all minimal viewing graphs with at most 9 vertices. We also discuss some relevant examples (Section 5).

2 Background

To make our presentation mostly self-contained, we recall some basic theoretical facts that are used in the rest of the paper.

Notation. We write $\mathbb{P}^n = \mathbb{P}(\mathbb{R}^{n+1})$ for the n -dimensional real projective space. We use bold font for vectors and matrices, and normal font for projective objects. For example, a point in \mathbb{P}^3 will be written as $p = [\mathbf{p}]$ where \mathbf{p} is a vector in \mathbb{R}^4 and p is the equivalence class associated with \mathbf{p} . Similarly, a projective transformation represented by a matrix \mathbf{M} will be written as $M = [\mathbf{M}]$. We use $GL(n, \mathbb{R})$ for the group of $n \times n$ invertible real matrices.

2.1 Camera configurations and epipolar geometry

A projective camera $P = [\mathbf{P}]$ is represented by 3×4 matrix \mathbf{P} of full rank, defined up to scale. The matrix \mathbf{P} describes a linear projection $\mathbb{P}^3 \setminus \{c\} \rightarrow \mathbb{P}^2$ where $c = [c]$ is the pinhole of the camera, associated with the null-space of \mathbf{P} .

The matrix group $GL(4, \mathbb{R})$ acts on the set of cameras by multiplication on the right, and represents the group of projective transformations of \mathbb{P}^3 , or of changes of homogeneous coordinates. We will use the fact that the group of matrices in $GL(4, \mathbb{R})$ that fix a camera $P = [\mathbf{P}]$ with pinhole $c = [c]$ is given by

$$\text{Stab}(\mathbf{P}) = \{\alpha \mathbf{I}_4 + \mathbf{c}\mathbf{v}^T \mid \alpha \in \mathbb{R} \setminus \{0\}, \mathbf{v} \in \mathbb{R}^4\} \cap GL(4, \mathbb{R}). \quad (1)$$

Here $\text{Stab}(P)$ stands for “stabilizer”. Indeed, all the solutions for \mathbf{M} in $\mathbf{P}\mathbf{M} = \alpha \mathbf{P}$ are described by (1). Note that $\text{Stab}(P)$ only depends on the pinhole of P . The following important fact follows directly from the form of $\text{Stab}(P)$.

Lemma 1. *Given two cameras P_1, P_2 with distinct pinholes, we have that*

$$\text{Stab}(P_1) \cap \text{Stab}(P_2) = \{\alpha \mathbf{I}_4, \mid \alpha \in \mathbb{R} \setminus \{0\}\}. \quad (2)$$

In other words, the identity is the only projective transformation that fixes both P_1 and P_2 .

Two sets of cameras (P_1, \dots, P_n) and (P'_1, \dots, P'_n) with $P_i = [\mathbf{P}_i]$, $P'_i = [\mathbf{P}'_i]$ are *projectively equivalent* if there exists a single projective transformation T such that $P_i = P'_i T$ (so if $T = [\mathbf{T}]$ with \mathbf{T} in $GL(4, \mathbb{R})$, then $\mathbf{P}_i = \alpha_i \mathbf{P}'_i \mathbf{T}$ for non-zero constants α_i). The set of *configurations* of n cameras is the set of n -tuples of cameras up to projective equivalence. For any $n \geq 2$, the space of camera configurations can be viewed as a manifold of dimension $11n - 15$.

Given two cameras $P_1 = [\mathbf{P}_1]$, $P_2 = [\mathbf{P}_2]$, the associated *fundamental matrix* $F(P_1, P_2) = [\mathbf{F}]$ can be defined as the 3×3 matrix (up to scale) with entries

$$f_{il} = (-1)^{i+l} \det(\mathbf{P}_{1j}^T \mathbf{P}_{1k}^T \mathbf{P}_{2m}^T \mathbf{P}_{2n}^T), \quad (3)$$

where \mathbf{P}_{ar} denotes the r -th row of \mathbf{P}_a , and (i, j, k) and (l, m, n) are triples of distinct indices. The fundamental matrix can be used to characterize pairs of corresponding points in the two images, since $u_1 = [\mathbf{u}_1]$ and $u_2 = [\mathbf{u}_2]$ are projections of the same 3D point if and only if $\mathbf{u}_1^T \mathbf{F} \mathbf{u}_2 = 0$. For our purposes, the most important property of the fundamental matrix is that it is invariant under projective transformations, and that $F(P_1, P_2)$ uniquely identifies the configuration of P_1 and P_2 [17, Theorem 9.10].

Finally, viewed as a subset of \mathbb{P}^8 , the (closure of the) set of all fundamental matrices forms a hypersurface defined by $\det(\mathbf{F}) = 0$. If $F(P_1, P_2) = [\mathbf{F}]$, the left and right null-space of \mathbf{F} represent the two *epipoles* $e_{12} = P_1 c_2$ and $e_{21} = P_2 c_1$, which are the images of each pinhole viewed from the other camera. An epipole accounts for two of seven degrees of freedom of a fundamental matrix. In fact, the information encoded in the fundamental matrix can be seen as the pair of epipoles e_{12}, e_{21} , together with a projective transformation $\mathbb{P}^1 \rightarrow \mathbb{P}^1$ (known as

“epipolar line homography” [17]) between lines containing e_{12} in the first image and the lines containing e_{21} in the second image. In particular, the knowledge of two epipoles together with three point correspondences completely determines a fundamental matrix.

3 The viewing graph

The viewing graph is a graph in which vertices correspond to cameras, and edges represent fundamental matrices between them. More precisely, if $G = (V_G, E_G)$ is an undirected graph with n vertices, and P_1, \dots, P_n are projective cameras, we write

$$\mathcal{F}_G(P_1, \dots, P_n) = \{F_{ij} = F(P_i, P_j) \mid (i, j) \in E_G\}, \quad (4)$$

for the set of fundamental matrices defined by the edges of G . We say that the set $\mathcal{F}_G(P_1, \dots, P_n)$ is *solvable* if $\mathcal{F}_G(P_1, \dots, P_n) = \mathcal{F}_G(P'_1, \dots, P'_n)$ implies that (P_1, \dots, P_n) and (P'_1, \dots, P'_n) are in the same projective configuration. In other words, a set of fundamental matrices is solvable if and only if it uniquely determines a projective configuration of cameras.

Proposition 1. *The solvability of $\mathcal{F}_G(P_1, \dots, P_n)$ only depends on the graph G and on the pinholes c_1, \dots, c_n of P_1, \dots, P_n .*

Proof. The statement expresses the fact that changes of image coordinates are only a relabeling of a camera configuration and the associated fundamental matrices. More precisely, if S_1, \dots, S_n are arbitrary projective transformations of \mathbb{P}^2 , then (P_1, \dots, P_n) and (P'_1, \dots, P'_n) are in the same configuration if and only if the same is true for $(S_1 P_1, \dots, S_n P_n)$ and $(S_1 P'_1, \dots, S_n P'_n)$. This implies that $\mathcal{F}_G(P_1, \dots, P_n)$ is solvable if and only if $\mathcal{F}_G(S_1 P_1, \dots, S_n P_n)$ is. \square

Example 1. If G is a *complete* graph with $n \geq 3$ vertices, then $\mathcal{F}_G(P_1, \dots, P_n)$ is solvable if and only if the pinholes of the cameras P_1, \dots, P_n are not all aligned. Indeed, if the pinholes are aligned, then the fundamental matrices between *all* pairs of cameras are not sufficient to completely determine the configuration: replacing any $P_i = [\mathbf{P}_i]$ with $P'_i = [\mathbf{P}_i(\mathbf{I}_4 + \mathbf{c}_j \mathbf{v}^T)]$, where $c_j = [\mathbf{c}_j]$ is the pinhole of another camera and \mathbf{v}^T is arbitrary, yields a new set of cameras which belongs to a different configuration but has the same set of fundamental matrices. Conversely, it is known (see for example [9, 13]) that the complete set of fundamental matrices determines a unique camera configuration whenever there are at least three non-aligned pinholes. \diamond

In the rest of the paper we will only consider generic configurations of cameras/pinhole (so a complete graph will always be solvable). This covers most cases of practical interest, although in the future degenerate configurations (including some collinear or coplanar pinholes) could be studied as well.

Definition 1. *A viewing graph G is said to be solvable if $\mathcal{F}_G(P_1, \dots, P_n)$ is solvable for generic cameras P_1, \dots, P_n .*

In other words, solvable viewing graphs describe sets of fundamental matrices that are generically sufficient to recover a camera configuration. Despite its clear significance, the problem of characterizing which viewing graphs are solvable has not been studied much, and only partial answers are available in the literature (mainly in [10,11]). It is quite easy to produce examples of graphs that are solvable, but it is much more challenging, given a graph, to determine whether it is solvable or not. The following observation provides another useful formulation of solvability (note that the “if” part requires the genericity assumption, as shown in Example 1).

Lemma 2. *A viewing graph G is solvable if and only if, for generic cameras P_1, \dots, P_n , the fundamental matrices $\mathcal{F}_G(P_1, \dots, P_n) = \{F(P_i, P_j) \mid (i, j) \in E_G\}$ uniquely determine the remaining fundamental matrices $\{F(P_i, P_j) \mid (i, j) \notin E_G\}$.*

This viewpoint also suggests the idea that, given any graph G , we can define a “solvable closure” \overline{G} , as the graph obtained from G by adding edges corresponding to fundamental matrices that can be deduced generically from $\mathcal{F}_G(P_1, \dots, P_n)$. Hence, a graph is solvable if and only if its closure is a complete graph. We will return to this point in Section 3.4.

3.1 Simple criteria

We begin by recalling two necessary conditions for solvability that were shown in [10]. These provide simple criteria to show that a viewing graph is *not* solvable.

Proposition 2. [10] *If a viewing graph with $n > 3$ vertices is solvable, then: 1) All vertices have degree at least 2. 2) No two adjacent vertices have degree 2.*

We extend this result with the following necessary condition (which implies the first point in the previous statement).

Proposition 3. *Any solvable graph is 2-connected, i.e., it has the property that after removing any vertex the graph remains connected.*

Proof. Assume that a vertex i disconnects the graph G into two components G_1, G_2 , and let P_1, \dots, P_n be a set of n generic cameras, whose pairwise fundamental matrices are represented by the edges of G . If $c_i = [\mathbf{c}_i]$ is the pinhole of the camera P_i associated with i , then we consider two distinct projective transformations of the form $T_1 = [\mathbf{I}_4 + \alpha_1 \mathbf{c}_i \mathbf{v}_1^T]$ and $T_2 = [\mathbf{I}_4 + \alpha_2 \mathbf{c}_i \mathbf{v}_2^T]$. These transformations fix the camera P_i . If we apply T_1 to all cameras in G_1 and T_2 to all cameras in G_2 , while leaving P_i fixed, we obtain a different camera configuration that gives rise to the same set of fundamental matrices as P_1, \dots, P_n for all edges in G . \square

We also recall a result from [11] which will be used in the next section.

Proposition 4. [11] *If G_1 and G_2 are solvable viewing graphs, then the graph G obtained by identifying two vertices from G_1 and with two from G_2 is solvable.*

Note that if both pairs of vertices in the previous statement are connected by edges in G_1 and G_2 , then these two edges will automatically be identified in G .

3.2 How many fundamental matrices?

We now ask ourselves what is the minimal number of edges that a graph must have to be solvable (or, equivalently, how many fundamental matrices are required to recover a camera configuration). Since a single epipolar relation provides at most 7 constraints in the $(11n-15)$ -dimensional space camera configuration, we deduce that any solvable graph must have *at least* $e(n) = \lceil (11n-15)/7 \rceil$ edges. This fact was previously observed in [11, Theorem 2]. However, compared to [11], we show here that this bound is tight, *i.e.*, that there always exists a solvable graph with $e = e(n)$ edges. Concretely, this means that, for n generic views, there is always a way of recovering the corresponding camera configuration using $e(n)$ fundamental matrices.

Theorem 1. *The minimum number of edges of a solvable viewing graph with $n \geq 2$ views is*

$$e(n) = \left\lceil \frac{11n - 15}{7} \right\rceil.$$

Proof. For $n \leq 9$, examples of solvable viewing graphs with $e(n)$ edges are illustrated Figure 1. The solvability of these graphs will be shown in Section 3.4 (all but one of these also appear in [10]). In particular, let G_0 be a solvable viewing graph with 9 vertices and 12 edges. Using Proposition 4, we deduce that, starting from a solvable viewing graph G with n vertices and e edges, we can always construct a solvable graph G' with $n + 7$ vertices and $e + 11$ edges. The graph G' is simply obtained by merging G and G_0 as in Proposition 4, using two pairs of vertices both connected by edges.

Now, for any $n > 9$, we consider the unique integers q, r such that $n = 7q + r$ and $2 \leq r \leq 8$. It is easy to see that

$$e(n) = \left\lceil \frac{11n - 15}{7} \right\rceil = 11q + \left\lceil \frac{11r - 15}{7} \right\rceil.$$

To obtain a solvable viewing graph with n vertices and $e(n)$ edges, we start from a solvable graph with r vertices and $e(r)$ edges, and repeat the gluing construction described above q times. The resulting graph is solvable and has the desired number of vertices and edges. \square

Remark 1. It is worth pointing out that, in order to recover projection matrices for n views, it is quite common to use $2n - 3$ fundamental matrices (see for example [18, Sec.4.4]). In fact, as shown in [13, Proposition 7], a large class of solvable viewing graphs can be defined, starting for example from a 3-cycle, by adding vertices of degree two, one at the time: this always gives a total of $2n - 3$ edges. For this type of viewing graphs it is possible to recover projection matrices incrementally, using a pair of fundamental matrices for each camera. In fact, it is probably quite often erroneously believed that $2n - 3$ is the minimal number of fundamental matrices that are required for multi-view reconstruction. Part of the confusion may arise from the fact that the “joint image” [8, 13, 19], which characterizes multi-view point correspondences in $(\mathbb{P}^2)^n$, has dimension three

(or codimension $2n - 3$). This means means that we expect $2n - 3$ conditions to be necessary to cut out generically the set of image correspondences among n views. On the other hand, according to Theorem 1, fewer constraints are actually sufficient to determine camera geometry.⁴

Some values of $e(n)$ are listed in Table 1 (here $d(n)$ represents the minimal number of constraints on the fundamental matrices, and will be discussed in the next section). Note that $e(n) < 2n - 3$ for all $n \geq 5$.

Table 1. The relation between n , $e(n)$, and $d(n)$

n	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
$e(n)$	1	3	5	6	8	9	11	12	14	16	17	19	20	22	23
$d(n)$	0	3	6	2	5	1	4	0	3	6	2	5	1	4	0

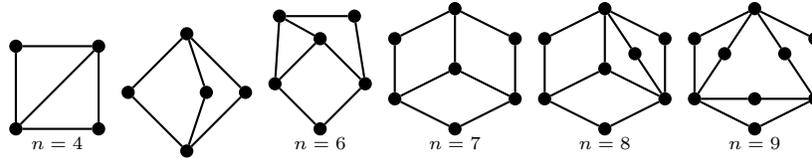


Fig. 1. Examples of minimal solvable viewing graphs for $n \leq 9$ views (see Section 3.4)

3.3 Constraints on fundamental matrices

Closely related to the solvability of viewing graphs is the problem of describing *compatibility* of fundamental matrices. Indeed, given a solvable graph G , it is not true in general that any set of fundamental matrices can be assigned to the edges of G , since fundamental matrices must satisfy some feasibility constraints in order to correspond to an actual camera configuration. For example, it is well known that the fundamental matrices F_{12}, F_{23}, F_{31} relating three pairs of cameras with non-aligned pinholes are compatible if and only if

$$\mathbf{e}_{13}^T \mathbf{F}_{12} \mathbf{e}_{23} = \mathbf{e}_{21}^T \mathbf{F}_{23} \mathbf{e}_{31} = \mathbf{e}_{32}^T \mathbf{F}_{31} \mathbf{e}_{12} = 0, \quad (5)$$

where $e_{ij} = [\mathbf{e}_{ij}]$ is the epipole in image i relative to the camera j [17, Theorem 15.6]. In most practical situations fundamental matrices are estimated separately, so these constraints need to be taken into account [15]. However, it is sometimes

⁴ This implies however that fewer than $2n - 3$ conditions can in fact determine a joint image in $(\mathbb{P}^2)^n$, at least “indirectly” through the camera configuration. Mathematically, this is an interesting phenomenon that could be investigated in the future.

incorrectly stated that compatibility for any set of fundamental matrices only arises from triples and equations of the form (5) [11, Theorem 1], [15, Definition 1]. While it is true that for a complete set of $\binom{n}{2}$ fundamental matrices triple-wise compatibility is sufficient to guarantee global compatibility, for smaller sets of fundamental matrices other types of constraints will be necessary. For example, there are many solvable viewing graphs with no three-cycles (*e.g.*, the graph in Figure 1 with $n = 5$), however the fundamental matrices cannot be unconstrained if $7e(n) > 11n - 15$, which always true unless $e = 2$ modulo 9 (cf. Table 1).

More formally, we can consider the set \mathcal{X} of compatible fundamental matrices between *all pairs* of n views, so that $\mathcal{X} \subset (\mathbb{P}^8)^N$ where $N = \binom{n}{2}$. Since each compatible N -tuple is associated with a camera configuration, we see that \mathcal{X} has dimension $11n - 15$. Given a viewing graph G with n views, we write $\mathcal{X}_G \subset (\mathbb{P}^8)^e$ for the projection of \mathcal{X} onto the factors in $(\mathbb{P}^8)^N$ corresponding to the edges of G . The set \mathcal{X}_G thus represents compatible fundamental matrices for pairs of views associated with the edges of G . The following result follows from dimensionality arguments (see the supplementary material for a complete proof).

Proposition 5. *If G is solvable with n vertices, \mathcal{X}_G has dimension $11n - 15$.*

If \mathcal{X}_G has dimension $11n - 15$, then the fundamental matrices assigned to the edges of G must satisfy $d(n, e) = 7e - 11n + 15$ constraints⁵. This also means that the minimum number of constraints on the fundamental matrices associated with a solvable graph is $d(n) = d(n, e(n))$ (see Table 1).

We now use Proposition 5 to deduce a new necessary condition for solvability.

Theorem 2. *Let G be a solvable graph with n vertices and e edges. Then for any subgraph G' of G with n' vertices and e' edges we must have*

$$d(n', e') \leq d(n, e), \quad (6)$$

where $d(n, e) = 7e - 11n + 15$. More generally, if G_1, \dots, G_k are subgraphs of G , each with n_i vertices and e_i edges, with the property that the edge sets $E_{G_i} \subset E_G$ are pairwise disjoint, then we must have

$$\sum_{i=1}^k d(n_i, e_i) \leq d(n, e). \quad (7)$$

Proof. Using the same notation as above, we note that that $\mathcal{X}_{G'}$ is a projection of \mathcal{X}_G onto e' factors of $(\mathbb{P}^8)^e$: this implies $\dim \mathcal{X}_{G'} + 7(e - e') \geq \dim \mathcal{X}_G$, or $7e' - \dim \mathcal{X}_{G'} \leq 7e - \dim \mathcal{X}_G$. Since $\dim \mathcal{X}_{G'} \leq 11n' - 15$ and $\dim \mathcal{X}_G = 11n - 15$ (because G is solvable), we obtain

$$7e' - 11n' + 15 \leq 7e' - \dim \mathcal{X}_{G'} \leq 7e - \dim \mathcal{X}_G = 7e - 11n + 15.$$

For the second statement, we consider the graph $G' = (\bigcup_i V_{G_i}, \bigcup_i E_{G_i})$. Since the edges of G_i are disjoint, we have

$$\dim \mathcal{X}_{G'} \leq \sum_{i=1}^k \dim \mathcal{X}_{G_i} \leq \sum_{i=1}^k (11n_i - 15),$$

⁵ This is the codimension of \mathcal{X}_G in \mathcal{H}^e where $\mathcal{H} \subset \mathbb{P}^8$ is the determinant hypersurface.

and $e' = \sum_i^k e_i$. The result follows again from $7e' - \dim \mathcal{X}_{G'} \leq 7e - \dim \mathcal{X}_G$. \square

Example 2. In [10], Levi and Werman observe that all viewing graphs of the form shown in Figure 2 are not solvable. This can be easily deduced from Theorem 2. Indeed, for a graph G of this form, the subgraphs G_1, G_2, G_3, G_4 have disjoint edges, however we have (using the same notation as in the proof of Theorem 2)

$$\sum_{i=1}^4 d(n_i, e_i) = d(n, e) - 4 \times 11 + 3 \times 15 > d(n, e).$$

According to Theorem 2 this means that G is not solvable. \diamond

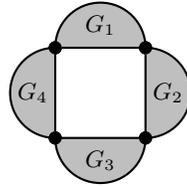


Fig. 2. A viewing graph of this form (where G_1, G_2, G_3, G_4 represent arbitrary subgraphs) is not solvable

3.4 Constructive approach for verifying solvability

Until now we have mainly discussed necessary conditions for solvability, which can be used to show that a given graph is not solvable. We next introduce a general strategy for proving that a graph is solvable. This method is not always guaranteed to work, but in practice it gives sufficient conditions for most of the graphs we tested (cf. Section 5).

Recall from the beginning of this section that we introduced the “viewing closure” \overline{G} of G as the graph obtained by adding to G all edges corresponding to fundamental matrices that can be deduced from $\mathcal{F}_G(P_1, \dots, P_n)$. Our approach consists of a series of “moves” which describe valid ways to add new edges to a viewing graph. For this it is convenient to introduce a new type of edge in the graph, which keeps track of the fact that partial information about a fundamental matrix is available. More precisely:

- A *solid (undirected) edge* between vertices i and j means that the fundamental matrix between the views i and j is fixed (as before).
- A *directed dashed edge* (for short, a *dashed arrow*) between vertices i and j means that the i -th epipole in the image j is fixed.

As these definitions suggest, a solid edge also counts as a dashed double-arrow, but the converse is not true. We next introduce three basic “moves” (cf. Figure 3).

- (I) If there are solid edges defining a four-cycle with one diagonal, draw the other diagonal.
- (II) If there are dashed arrows $1 \rightarrow 2$, $1 \rightarrow 3$, and solid edges $2 - 4$ and $3 - 4$, draw a dashed arrow $1 \rightarrow 4$.
- (III) If there are double dashed arrows $1 \leftrightarrow 2$, together with three pairs of dashed arrows $i \rightarrow 1, i \rightarrow 2$ for $i = 3, 4, 5$, make the arrow between 1 and 2 a solid (undirected) edge.

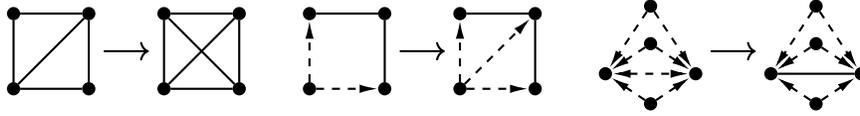


Fig. 3. Three moves (left: I, center: II, right: III) that can be used to prove solvability

Theorem 3. *Let G be a viewing graph. If applying the three moves described above iteratively to G we obtain a complete graph, then G is solvable.*

Proof. For each of the three moves we need to show that the new edges contain information about the unknown fundamental matrices that can actually be deduced from $\mathcal{F}_G(P_1, \dots, P_n)$.

Move I: The second diagonal of the square is deducible from the other edges because the square with one diagonal is a solvable graph (this is a simple consequence of Proposition 4).

Move II: Assume that $e_{21} = P_2c_1$ and $e_{32} = P_3c_2$ are fixed epipoles in images 2 and 3, and that the fundamental matrices $F_{24} = F(P_2, P_4)$, $F_{34} = F(P_3, P_4)$ are also fixed. If c_1, c_2, c_4 are not aligned, we can use F_{24} to “transfer” the point e_{21} , and obtain a line l_{41} in image 4 that contains epipole e_{41} . Similarly, if c_1, c_3, c_4 are not aligned, we obtain another line m_{41} using the same procedure with F_{34} and e_{31} . If the pinholes c_1, c_2, c_3, c_4 are not all coplanar, the lines l_{41} and m_{41} will be distinct, and their point of intersection will be e_{41} . This implies that we can draw a dashed arrow from 1 to 4.

Move III: Assume that the epipoles e_{21} and e_{12} are fixed, and that the images of three other pinholes c_3, c_4, c_5 are fixed in both images 1 and 2. If the planes c_1, c_2, c_i for $i = 3, 4, 5$ are distinct, then the images of c_3, c_4, c_5 give three correspondences that fix the epipolar line homography. This completely determines F_{12} , and we can draw a solid edge between 1 and 2. \square

In practice, the three moves can be applied cyclically until no new edges can be added (it is also easy to argue the order is irrelevant, because we are simply annotating information that is always deducible from the graph). Finally, we note that all three moves are *constructive* and *linear*, meaning they actually provide a practical strategy for computing all fundamental matrices: it is enough to transfer epipoles appropriately, and use them to impose linear conditions on the unknown fundamental matrices.

Example 3. Using Theorem 3, we can show that all graphs from Figure 1 are solvable. Figure 4 illustrates this explicitly for two cases ($n = 6$ and $n = 8$). \diamond

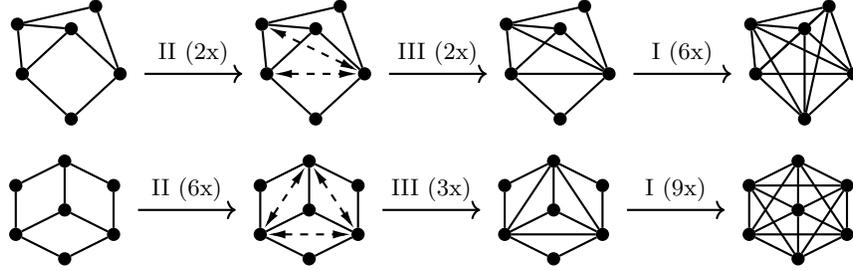


Fig. 4. Two applications of Theorem 3 to prove that viewing graphs are solvable

4 Algebraic tests for solvability and finite solvability

Given a viewing graph G , it is possible to write down a set of algebraic conditions that will in principle always determine whether G is solvable. One way to do this is by characterizing the set of projective transformations of \mathbb{P}^3 that can be applied to all cameras without affecting any of the fundamental matrices represented by the edges of the viewing graph. More precisely, since every pair of vertices connected by an edge represents a projectively rigid pair of cameras, we assign a matrix \mathbf{g}_λ in $GL(4, \mathbb{R})$ to each edge λ of the graph (so \mathbf{g}_λ describes a projective transformation applied to a pair of cameras). We then impose that matrices on adjacent edges act compatibly on the shared vertex/camera. If the edges λ and λ' share a vertex i , then from (1) we see that this compatibility can be written as

$$\mathbf{g}_\lambda \mathbf{g}_{\lambda'}^{-1} = \alpha \mathbf{I}_4 + \mathbf{c}_i \mathbf{v}^T, \quad (8)$$

where α is an arbitrary (nonzero) constant and \mathbf{v} is an arbitrary vector. Thus, if G is a viewing graph with e edges and c_1, \dots, c_n are a set of pinholes, we consider the set of all compatible assignments of matrices:

$$\mathcal{T}_G(c_1, \dots, c_n) = \{(\mathbf{g}_\lambda, \lambda \in E_G) \mid (8) \text{ holds for all adjacent edges in } G\} \subset GL(4, \mathbb{R})^e.$$

If G is solvable, then for general c_1, \dots, c_n the set $\mathcal{T}_G(c_1, \dots, c_n)$ will consist of e -tuples of matrices that are all scalar multiples of each other. This in fact means that the only way to act on all cameras without affecting the fixed fundamental matrices is to apply a single projective transformation.

By substituting random pinholes in (8), we can use these equations for $\mathcal{T}_G(c_1, \dots, c_n)$ as an algebraic test for verifying whether a viewing graph is solvable. This approach however is computationally very challenging, since it requires solving a non-linear algebraic system with a large number of variables.

On the other hand, if we are only interested in the *dimension* of $\mathcal{T}_G(c_1, \dots, c_n)$, then we can use a much simpler strategy: noting that $\mathcal{T}_G(c_1, \dots, c_n)$ may be viewed as an algebraic group (it is a subgroup of $GL(4, \mathbb{R}^e)$, it is sufficient to compute the dimension of its *tangent space* at any point, and in particular at the identity (*i.e.*, the product of identity matrices).⁶ An explicit representation of the tangent space of $\mathcal{T}_G(c_1, \dots, c_n)$ is provided by the following result (see the supplementary material for a proof).

Proposition 6. *The tangent space of $\mathcal{T}_G(c_1, \dots, c_n)$ at the identity can be represented as the space of e -tuple of matrices $(\mathbf{h}_\lambda, \lambda \in E_G)$ where each \mathbf{h}_λ is in $\mathbb{R}^{4 \times 4}$ (not necessary invertible), and with compatibility conditions of the form*

$$\mathbf{h}_\lambda - \mathbf{h}_{\lambda'} = \alpha \mathbf{I}_4 + \mathbf{c}_i \mathbf{v}^T, \quad (9)$$

where $\alpha \in \mathbb{R} \setminus \{0\}$ and $\mathbf{v} \in \mathbb{R}^4$ are arbitrary, and λ and λ' share the vertex i .

When the pinholes have been fixed, the compatibility constraints (9) can be expressed as *linear* equations in the entries of the matrices \mathbf{h}_λ . These equations are obtained by eliminating the variables α and \mathbf{v} from (9). The resulting conditions in terms of $\mathbf{h}_\lambda, \mathbf{h}_{\lambda'}, \mathbf{c}_i$ are rather simple, and listed explicitly in the supplementary material. Using this approach, the dimension of $\mathcal{T}_G(c_1, \dots, c_n)$ is easy to determine: it is enough to fix the pinholes randomly, and compute the dimension of the induced linear system.

When $\mathcal{T}_G(c_1, \dots, c_n)$ has dimension $d = 15 + e$ (which accounts for the group of projective transformations, and scale factors for each matrix \mathbf{g}_λ), we deduce that there are at most a *finite* number of projectively inequivalent ways in which we can act on all the cameras without affecting the fixed fundamental matrices. In other words, the fundamental matrices associated with the edges of G determine at most a finite set of camera configurations (rather than a single configuration, which is our definition for solvability). When this happens, we say that G is *finite solvable*. On the other hand, we were not able to find an example of a finite solvable graph that is provably not solvable, nor to find a proof that “finite solvability” implies “solvability”. To our knowledge, whether a set of fundamental matrices can characterize a finite number of configurations, but more than a single one, is a question that has never been addressed.

Open Question. Is it possible for a viewing graph to be finite solvable without being solvable?

Our experiments show that this behavior does not occur for a small number of vertices, but we see no reason why this should be true for larger graphs. This is certainly an important issue that we hope to investigate in the future.

⁶ Here we actually need that $\mathcal{T}_G(c_1, \dots, c_n)$ is smooth: this follows from a technical result, which states that an algebraic group (more properly a “group scheme”) over a field of characteristic zero is always smooth [20, Sec.11].

5 Experiments and examples

We have implemented and tested all of the discussed criteria and methods using the free mathematical software `SageMath` [21].⁷ We then analyzed solvability for all minimal viewing graphs with $n \leq 9$ vertices and $e(n) = \lceil (11n - 15)/7 \rceil$ edges. The results are summarized in Table 5. For every pair $(n, e(n))$, we list the number of all non-isomorphic connected graphs of that size (“connected”), the number of graphs that satisfy the necessary condition from Theorem 2 (“candidates”), the number of those that satisfy the sufficient condition from Theorem 3 (“solvable with moves”), and the number of graphs that are finite solvable (“finite solvable”), using the linear method from 4. We see that Theorems 2 and 3 allow us to recover all minimal solvable graphs for $n \leq 7$, since candidate graphs are always solvable with moves. On the other hand, for $n = 8$, and particularly for the unconstrained case $n = 9$, there are some graphs that we could not classify with those methods (although finite solvability was easy to verify in all cases). For the undecided graphs, we were sometimes able to prove solvability with the general algebraic method from Section 4, or using other arguments. The following examples present a few interesting cases.

Table 2. Solvability of minimal viewing graphs using our methods

$(n, e(n))$	(3,3)	(4,5)	(5,6)	(6,8)	(7,9)	(8,11)	(9,12)
connected	1	1	5	22	107	814	4495
candidates	1	1	1	4	3	36	28
solvable with moves	1	1	1	4	3	31	5
finite solvable	1	1	1	4	3	36	27

Example 4. The graph shown in Figure 5 (left) is one of the five cases with $n = 8$, $e = 11$ that are “candidates” but are not “solvable with moves”. However, we can show that this graph is actually solvable by arguing that the image of the pinhole 1 in the view 7 is fixed, even if this is not a consequence of the moves of Theorem 3 (this is represented by the gray dashed arrow in the figure). To prove this fact, one needs to keep track of more information, and record also when an epipole is constrained to a line (rather than only when an epipole is fixed, which is the purpose of dashed edges).⁸ After drawing the dashed arrow from 1 to 7, solvability can be shown using the moves from Theorem 3. \diamond

Example 5. The graph shown in Figure 5 (center) is the only viewing graph with $n = 9$ and $e = 12$ that is “candidate” but is not “finite solvable”. The fact that it is not finite solvable can also be deduced without computations. Indeed, any finite solvable viewing graph of this size cannot impose any constraints on

⁷ Our code is available at <https://github.com/mtrager/viewing-graphs>.

⁸ This information can be taken into account by defining a new type of edge together with additional moves. We did not do this in Theorem 3 because this type of edge is never necessary for smaller graphs.

the fundamental matrices associated with its edges (this is because $d(9, 12) = 7 \times 12 - 11 \times 9 + 15 = 0$). However, the image of the pinhole 7 in the view 2 is over-constrained, because we can draw a dashed arrow $7 \rightarrow 2$ using move II for two distinct four-cycles $((7, 1, 2, 5)$ and $(7, 3, 2, 5))$. This implies that the fundamental matrices associated with the edges of the graph cannot be arbitrary. \diamond

Example 6. The graph shown in Figure 5 (right) is not “solvable with moves”, however one can show that it is solvable: indeed, the general algebraic compatibility equations from Section 4 are in this case simple and can be solved explicitly (see the supplementary material for the computations). The fundamental matrices associated with the edges of the graph are unconstrained, so 12 arbitrary fundamental matrices determine a unique configuration of 9 cameras. \diamond

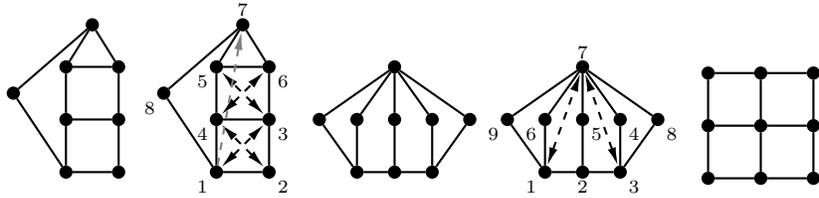


Fig. 5. The graphs described in Examples 4, 5, 6

6 Conclusions

We have studied the problem of solvability of viewing graphs, presenting a series of new theoretical results that can be applied to determine whether a graph is solvable. We have also pointed out some open questions (particularly, the relation between finite solvability and solvability, discussed in Section 4), and we hope that this paper can lead to further work on these issues.

Our main focus here was to understand whether the camera-estimation problem is well-posed, and we did not directly address the task of determining the configuration computationally. Properly recovering a global camera configuration that is consistent with local measurements is challenging, and is arguably the main obstacle for any structure-from-motion algorithm. For this reason, we believe that a complete understanding of the algebraic constraints that characterize the compatibility of fundamental matrices would be very useful. This is an issue that has not been considered much in classical multi-view geometry, and is very closely related to the topic of this paper. We plan to investigate it next.

Acknowledgments. This work was supported in part by the ERC grant VideoWorld, the Institut Universitaire de France, the Inria-CMU associated team GAYA, a collaboration agreement between Inria and NYU, and a grant from the Simons Foundation #279151.

References

1. Thompson, M., Eller, R., Radlinski, W., Speert, J., eds.: Manual of Photogrammetry. American Society of Photogrammetry (1966) Third Edition.
2. Longuet-Higgins, H.C.: A computer algorithm for reconstructing a scene from two projections. *Nature* **293**(5828) (1981) 133
3. Luong, Q.T., Faugeras, O.: The fundamental matrix: theory, algorithms, and stability analysis. **17**(1) (1996) 43–76
4. Shashua, A.: Algebraic functions for recognition. **17**(8) (1995) 779–789
5. Hartley, R.I.: Lines and points in three views and the trifocal tensor. *International Journal of Computer Vision* **22**(2) (1997) 125–140
6. Hartley, R.: Computation of the quadrifocal tensor. (1998) 20–35
7. Faugeras, O., Mourrain, B.: On the geometry and algebra of the point and line correspondences between n images. In: *Computer Vision, 1995. Proceedings., Fifth International Conference on, IEEE (1995)* 951–956
8. Triggs, B.: Matching constraints and the joint image. In: *Computer Vision, 1995. Proceedings., Fifth International Conference on, IEEE (1995)* 338–343
9. Heyden, A., Åström, K.: Algebraic properties of multilinear constraints. *Mathematical Methods in the Applied Sciences* **20**(13) (1997) 1135–1162
10. Levi, N., Werman, M.: The viewing graph. In: *Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on. Volume 1., IEEE (2003)* I–I
11. Rudi, A., Pizzoli, M., Pirri, F.: Linear solvability in the viewing graph. In: *Asian Conference on Computer Vision, Springer (2010)* 369–381
12. Snavely, N., Seitz, S., Szeliski, R.: Photo tourism: Exploring image collections in 3D. In: *SIGGRAPH. (2006)*
13. Trager, M., Hebert, M., Ponce, J.: The joint image handbook. In: *Proceedings of the IEEE International Conference on Computer Vision. (2015)* 909–917
14. Ozyesil, O., Singer, A.: Robust camera location estimation by convex programming. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. (2015)* 2674–2683
15. Sweeney, C., Sattler, T., Hollerer, T., Turk, M., Pollefeys, M.: Optimizing the viewing graph for structure-from-motion. In: *Proceedings of the IEEE International Conference on Computer Vision. (2015)* 801–809
16. Sinha, S.N., Pollefeys, M.: Camera network calibration and synchronization from silhouettes in archived video. *International journal of computer vision* **87**(3) (2010) 266–283
17. Hartley, R., Zisserman, A.: *Multiple view geometry in computer vision*. Cambridge university press (2003)
18. Heyden, A.: Tensorial properties of multiple view constraints. *Mathematical Methods in the Applied Sciences* **23**(2) (2000) 169–202
19. Aholt, C., Sturmfels, B., Thomas, R.: A hilbert scheme in computer vision. *Canad. J. Math* **65**(5) (2013) 961–988
20. Mumford, D.: *Abelian Varieties*. Studies in mathematics. Hindustan Book Agency (2008)
21. Developers, T.S.: SageMath, the Sage Mathematics Software System (Version 8.0.0). (2017) <http://www.sagemath.org>.