

Relative Multi-View Geometry

Matthew Trager

This document collects some informal notes on relative multi-view geometry from the perspective of point configurations. Some of the material presented is well-known, and some overlaps with [7].

Notation: We use bold font for vectors and matrices, and normal font for projective objects.

Point Configurations. We write \mathbb{P}_k^n for the space of configurations of k ordered points in \mathbb{P}^n . This means that we consider ordered sets of k points, and say that two k -tuples of points (p_1, \dots, p_k) and (q_1, \dots, q_k) in $(\mathbb{P}^n)^k$ are equivalent if there exists a projective transformation of \mathbb{P}^n that maps p_i to q_i for $i = 1, \dots, k$. We write $\langle x_1, \dots, x_k \rangle$ for the configuration in \mathbb{P}_k^n of the points x_1, \dots, x_k in \mathbb{P}^n .

It will be convenient to parameterize elements in \mathbb{P}_k^n using $k \times (n+1)$ matrices, with each row corresponding to the projective coordinates of a point:

$$\begin{pmatrix} - & \mathbf{x}_1^T & - \\ & \vdots & \\ - & \mathbf{x}_k^T & - \end{pmatrix} \mapsto \langle x_1, \dots, x_k \rangle, \quad \mathbf{x}_i \in \mathbb{R}^{n+1}. \quad (1)$$

We write $\langle \mathbf{M} \rangle$ for the configuration defined by the $k \times (n+1)$ matrix \mathbf{M} .

Lemma 1. *Two $k \times (n+1)$ matrices $\mathbf{M}_1, \mathbf{M}_2$ give rise to the same configuration of points in \mathbb{P}_k^n if and only if there exists \mathbf{T} in $GL_{n+1}(\mathbb{R})$ and a non-singular diagonal $k \times k$ matrix \mathbf{D} such that $\mathbf{D}\mathbf{M}_1\mathbf{T} = \mathbf{M}_2$.*

Here the diagonal matrix \mathbf{D} is necessary to eliminate the dependence on the choices of homogeneous coordinates. Note that if $k < n+2$, then there is only one generic configuration. Assuming $k \geq n+2$, we can easily associate a generic configuration $\langle \mathbf{M} \rangle$ in \mathbb{P}_k^n with $n(k-n-2)$ invariant coefficients: it is sufficient to remove the projective ambiguity by assuming that the first $n+2$ rows of \mathbf{M} identify a projective basis z_1, \dots, z_{n+2} , and rescale the remaining rows so that (say) the last column is always one. The remaining free $n(k-n-2)$ coefficients uniquely determine the point configuration. This can be seen as a generalization of the classical cross-ratio for four points in \mathbb{P}^1 .

Another language for expressing geometric properties of point configurations is based on the *bracket algebra*, developed in *algebraic invariant theory* [6]. The “brackets” of a configuration $\langle \mathbf{M} \rangle$ are the set of all $(n+1) \times (n+1)$ minors of \mathbf{M} . Any projectively invariant property of a set of points can be expressed as a (multihomogeneous) polynomial in brackets, so brackets can be viewed as a set of “coordinates” for the configuration. On the other hand, brackets are not algebraically independent, since they satisfy the quadratic *Plücker-Grassmann relations*. Furthermore, a configuration can be represented by many possible sets of brackets, corresponding to different choices for the matrix \mathbf{M} .

Gale duality. There is a natural duality that associates configurations in \mathbb{P}_k^n with configurations in \mathbb{P}_k^m with $m+n+2=k$. This association was known at least since the nineteenth century, and was extensively studied by mathematicians such as Castelnuovo and Coble [1]. It was independently rediscovered (in affine space) in the context of polytopes and linear programming. A study of Gale duality from the perspective of modern algebraic geometry can be found in [2].

Definition 2. Let m, n, k be positive integers such that $m+n+2=k$. Two point configurations $\langle \mathbf{M} \rangle$ in \mathbb{P}_k^n and $\langle \mathbf{N} \rangle$ in \mathbb{P}_k^m , associated with matrices \mathbf{M}, \mathbf{N} of sizes $k \times (n+1)$ and $k \times (m+1)$ respectively, are said to be Gale transforms of each other if there exists a non-singular diagonal $k \times k$ matrix \mathbf{D} such that $\mathbf{M}^T \mathbf{D} \mathbf{N} = 0$.

The diagonal matrix \mathbf{D} serves once again to eliminate the dependence on the choices of homogeneous coordinates (in both $\langle \mathbf{M} \rangle$ and $\langle \mathbf{N} \rangle$). The idea is that the transform of the configuration points defined by the rows \mathbf{M} is the configuration of points represented by the kernel of \mathbf{M}^T . This algebraic definition is very simple, but for the moment its geometric interpretation is not clear. It is also easy to see that if the k points in $\langle \mathbf{M} \rangle$ span all of \mathbb{P}^n (i.e., the matrix \mathbf{M} has full rank), then the Gale transform determines a single configuration $\langle \mathbf{N} \rangle$ in \mathbb{P}^m : this means that for general configurations we can talk about *Gale duality*.

The following example illustrates how to compute the dual of a configuration.

Example 3. If the first $n+1$ points of a configuration $\langle \mathbf{M} \rangle$ in \mathbb{P}_k^n are in general position, we may assume that \mathbf{M} has the form

$$\begin{pmatrix} \mathbf{I}_{n+1} \\ \mathbf{A} \end{pmatrix}.$$

The *Gale dual* of this configuration is now defined by the rows of

$$\begin{pmatrix} \mathbf{A}^T \\ \mathbf{I}_{m+1} \end{pmatrix}.$$

Indeed, these matrices satisfy Definition 2 with $D = \text{diag}(\mathbf{I}_{n+1}, -\mathbf{I}_{m+1})$.

Example 4. The Gale dual of a configuration of six points in \mathbb{P}^3 is a configuration of six points in \mathbb{P}^1 . If we assume that p_1, \dots, p_5 are the five projective reference points, and $p_6 = (c_1 : c_2 : c_3 : c_4)$, then the Gale dual is the configuration of $(1 : c_1), (1 : c_2), (1 : c_3), (1 : c_4), (1 : 0), (0 : 1)$ in \mathbb{P}^1 . Geometrically, the configuration in \mathbb{P}^1 is the configuration of p_1, \dots, p_6 along the unique twisted cubic in \mathbb{P}^3 passing through p_1, \dots, p_6 [2, Example (a)].

Maps on configurations. We consider two simple maps associating a configuration of k points $\langle x_1, \dots, x_k \rangle$ in \mathbb{P}_k^n with configurations of $k-1$ points:

$$\begin{aligned} \rho_k : \langle x_1, \dots, x_k \rangle &\mapsto \langle x_1, \dots, x_{k-1} \rangle \in \mathbb{P}_{k-1}^n, \\ \pi_k : \langle x_1, \dots, x_k \rangle &\mapsto \langle u_1, \dots, u_{k-1} \rangle \in \mathbb{P}_{k-1}^{n-1}. \end{aligned}$$

where $u_i = P_{x_k}(x_i) \in \mathbb{P}^{n-1}$ and $P_{x_k} : \mathbb{P}^n \dashrightarrow \mathbb{P}^{n-1}$ is any linear projection with center x_k . In other words, the map ρ_k simply discards the last point x_k from the configuration, while the map π_k uses x_k to project the remaining points into \mathbb{P}^{n-1} . These two maps are related by Gale duality as follows [1].

Proposition 5. *Let $\langle x_1, \dots, x_k \rangle$ and $\langle y_1, \dots, y_k \rangle$ be two point configurations in \mathbb{P}_k^n and \mathbb{P}_k^m that are Gale transforms of each other (so $n + m + 2 = k$). Then the configurations $\rho_k(x_1, \dots, x_k)$ and $\pi_k(y_1, \dots, y_k)$ in \mathbb{P}_{k-1}^n and \mathbb{P}_{k-1}^{m-1} are also Gale transforms of each other.*

Finally, we observe that configurations are not invariant under reorderings of points: if σ is a permutation of $\{1, \dots, k\}$, then in general $\langle x_1, \dots, x_k \rangle \neq \langle x_{\sigma(1)}, \dots, x_{\sigma(k)} \rangle$. The effect of permutations on point configurations was studied by Coble [1]. Here, we limit ourselves to the following special situation (see [7, Proposition A.2]).

Lemma 6. *Given $n+1$ points x_1, \dots, x_{n+1} in \mathbb{P}^{n+1} in general position, there exists a family of birational involutions $T : x \mapsto \hat{x}$, defined on a dense open set of \mathbb{P}^3 , such that for any points y_1 and y_2 in that set $\langle x_1, \dots, x_{n+1}, y_1, y_2 \rangle = \langle x_1, \dots, x_{n+1}, \hat{y}_2, \hat{y}_1 \rangle$ holds. Any two such involutions are related by a projective transformation of \mathbb{P}^3 that fixes x_1, \dots, x_{n+1} .*

If x_1, \dots, x_{n+1} are reference points $(1 : 0 : \dots : 0), \dots, (0 : \dots : 1)$, then T can be chosen to be the standard Cremona involution which inverts the coordinates of generic points in \mathbb{P}^n .

Cameras and Scenes. In multi-view geometry, we are concerned with configurations of points in \mathbb{P}^3 and \mathbb{P}^2 .

Definition 7. *A viewing configuration is a configuration*

$$S_{n,k} = \langle c_1, \dots, c_n, x_1, \dots, x_k \rangle \in \mathbb{P}_{n+k}^3$$

of $n+k$ points in \mathbb{P}^3 , where the first n points are viewed as “pinholes” and the remaining points are “scene points”. The image configurations of a viewing configuration are

$$I_k^i = \langle u_1^i, \dots, u_k^i \rangle \in \mathbb{P}_k^2, \quad i = 1, \dots, n,$$

where u_1^i, \dots, u_k^i are points obtained by projecting x_1, \dots, x_k from c_i . The image configurations I_k^i are all uniquely determined by $S_{n,k}$.

We will often use $\mathbb{P}_{n,k}^3$ instead of \mathbb{P}_{n+k}^3 for the space of viewing configurations with n pinholes and k scene points. In this setting, the problem of *relative multi-view reconstruction* (from n views and k scene points) consists in using image configurations I_k^1, \dots, I_k^n in \mathbb{P}_k^2 to recover the unknown viewing configuration $S_{n,k}$ in $\mathbb{P}_{n,k}^3$ which generated them. This problem is completely equivalent to traditional formulations of multi-view reconstruction, but it “factors out” projective equivalence in both \mathbb{P}^3 and \mathbb{P}^2 . This is convenient, since:

1. A projective reference frame in \mathbb{P}^3 is not physically defined, and indeed all traditional projective reconstruction methods only yield solutions up to “projective ambiguity”.
2. A projective reference frame in the images \mathbb{P}^2 is usually determined by image measurements, but the equivalence classes of the images are in fact sufficient to determine the equivalence class of the scene.

Of course, arbitrary image configurations I_k^1, \dots, I_k^n will not be associated with any scene $S_{n,k}$. Classical results in multi-view geometry can be used to characterize the set of “compatible” image configurations, as well as the set of viewing configurations that can generate them.

One camera. Let us first consider viewing configurations $S_{1,k} \in \mathbb{P}_{1,k}^3$ with one camera. Since generic configurations in \mathbb{P}^2 are equivalent for $k \leq 4$, we assume $k \geq 5$. The following result captures the general situation.

Proposition 8. *Given a general image configuration I_k in \mathbb{P}_k^2 with $k \geq 5$ points, there exists an $(k - 4)$ -dimensional family of viewing configurations $S_{1,k} = \langle c, x_1, \dots, x_k \rangle$ which generate I_k : if c, x_1, \dots, x_4 are independent and fixed, then x_5, \dots, x_k must each lie on a line through c ; if x_1, \dots, x_5 are independent and fixed, then c must lie on a twisted cubic through x_1, \dots, x_5 , while x_6, \dots, x_k must lie on lines through c and, independently from c , they must each lie on (ruled) quadrics containing x_1, \dots, x_5 .*

These facts are known (see for example [5]). To justify them here, we use the following general fact.

Lemma 9. *Let $X = (x_1, \dots, x_4)$ be a quadruple points in \mathbb{P}^3 no three of which are collinear. For any c not aligned with any pair in X , there exists a unique pinhole camera $P_{c,X} : \mathbb{P}^3 \dashrightarrow \mathbb{P}^2$ with pinhole c such that $P(x_i) = e_i$, where e_1, \dots, e_4 is the standard basis of \mathbb{P}^2 . The projection mapping is defined by*

$$P_{c,X}(x) = \left(\frac{[cx_2x_3x]}{[cx_2x_3x_4]} : \frac{[cx_1x_3x]}{[cx_1x_3x_4]} : \frac{[cx_1x_2x]}{[cx_1x_2x_4]} \right) \in \mathbb{P}^2, \quad (2)$$

where $[p_1p_2p_3p_4]$ denotes the determinant of the 4×4 matrix defined by a choice of vector coordinates for the projective points p_1, p_2, p_3, p_4 (a bracket for the point configuration) We refer to $P_{c,X}$ as the reduced camera with pinhole c , relative to X .

Note that if the points X are the reference points of \mathbb{P}^3 , $P_{c,X}$ is described by the well-known reduced projection matrix

$$\begin{pmatrix} \frac{1}{c^1} & 0 & 0 & -\frac{1}{c^4} \\ 0 & \frac{1}{c^2} & 0 & -\frac{1}{c^4} \\ 0 & 0 & \frac{1}{c^3} & -\frac{1}{c^4} \end{pmatrix}, \quad c = (c^1 : c^2 : c^3 : c^4). \quad (3)$$

Furthermore, if

$$\hat{c} = \left(\frac{x_1}{[xx_2x_3x_4]} \right)$$

Using (2), we deduce from $P_{c,X}(x) = u$ that

$$\text{rk} \begin{pmatrix} [cx_2x_3x] & u^1[cx_2x_3x_4] \\ [cx_1x_3x] & u^2[cx_1x_3x_4] \\ [cx_1x_2x] & u^3[cx_1x_2x_4] \end{pmatrix} = 1, \quad (4)$$

where $u = (u^1 : u^2 : u^3)$. If we fix x and u , this expression constrains the pinhole c to lie on a twisted cubic in \mathbb{P}^3 (see [7]). Moreover, if further $P_{c,X}(y) = v$, then by using (4) and eliminating c we obtain a relation of the form

$$\mathbf{u}^T \mathbf{G}_X(x, y) \mathbf{v} = 0. \quad (5)$$

where $\mathbf{G}_X(x, y)$ is a 3×3 matrix whose entries are biquadratic in the coordinates of x and y . This can be viewed as an instance of a *dual reduced fundamental matrix* [3]. If the points in X are the reference points of \mathbb{P}^3 , then

$$G_X(x, y) = \begin{pmatrix} 0 & -x^2x^4y^1y^3 + x^2x^3y^1y^4 & x^3x^4y^1y^2 - x^2x^3y^1y^4 \\ x^1x^4y^2y^3 - x^1x^3y^2y^4 & 0 & -x^3x^4y^1y^2 + x^1x^3y^2y^4 \\ -x^1x^4y^2y^3 + x^1x^2y^3y^4 & x^2x^4y^1y^3 - x^1x^2y^3y^4 & 0 \end{pmatrix}$$

If we assume that x is fixed, (5) describes a quadric in y that passes through x_1, x_2, x_3, x_4, x .

Example 10. We present an illustration of Gale duality and Carlsson-Weinshall duality for camera projections. Consider a viewing configuration that we write $S = \langle x_1, x_2, x_3, x_4, x_5, c \rangle$ in \mathbb{P}_6^3 . If S is the Gale dual of $S' = \langle u_1, u_2, u_3, u_4, u_5, u_6 \rangle$ in \mathbb{P}_6^1 , then the image $I \in \mathbb{P}_5^2$ of S (with a projection from c) is the Gale dual of $\langle u_1, u_2, u_3, u_4, u_5 \rangle \in \mathbb{P}_5^1$. To see all this more concretely, let assume that S is described by the matrix

$$\begin{pmatrix} I_4 \\ x_5 \\ 1 \ 1 \ 1 \ 1 \end{pmatrix}.$$

If $x_5 = (x^1 : x^2 : x^3 : x^4)$, the Gale dual of this configuration is described by the rows of the matrix

$$\begin{pmatrix} x^1 & 1 \\ x^2 & 1 \\ x^3 & 1 \\ x^4 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}. \quad (6)$$

Note that the same configuration in \mathbb{P}_6^1 is defined by the rows $(1, 1/x^i)$, corresponding to using $\hat{c}, x_1, x_2, x_3, x_4, \hat{x}$ for defining S . The image configuration I is the Gale dual of the configuration in \mathbb{P}_5^1 obtained by removing the last row of (6). It is easy to verify that this corresponds to the rows of

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & -1 & -1 \\ x^1 - x^4 & x^2 - x^4 & x^3 - x^4 \end{pmatrix}.$$

This agrees with the expression (3) for projecting x_5 using a reduced camera with pinhole $c = (1 : 1 : 1 : 1)$.

Multiple cameras. The following fact follows immediately from Proposition 8.

Proposition 11. *Given n general image configurations I_5^1, \dots, I_5^n in \mathbb{P}_5^2 with $k = 5$ points, there exists an n -dimensional family of viewing configurations $S_{n,5} = \langle c_1, \dots, c_n, x_1, \dots, x_5 \rangle$ which generates all the I_5^i : if x_1, \dots, x_5 are independent and fixed, then each c_i must lie on a twisted cubic through x_1, \dots, x_5 . Given n images I_6^1, \dots, I_6^n in \mathbb{P}_6^2 , there exists a scene $\langle c_1, \dots, c_n, x_1, \dots, x_6 \rangle$ in $\mathbb{P}_{n,6}^3$ that generates all the I_6^i if and only if the n quadrics from Proposition 8 intersect at a point different from x_1, \dots, x_5 . Over \mathbb{C} , we expect $2^3 - 5 = 3$ solutions for $n = 3$, and zero solutions for $n > 3$.*

Reduced joint images. We recall that the *joint image* of a set of $n \geq 2$ cameras P_1, \dots, P_n is the subvariety of $(\mathbb{P}^2)^n$ defined as the closure of all points $(P_1(x), \dots, P_n(x))$ for all admissible x in \mathbb{P}^3 .

Definition 12. A reduced joint image is the joint image in $(\mathbb{P}^2)^n$ associated with $n \geq 2$ reduced cameras $P_{c_1, X}, \dots, P_{c_n, X}$ for a fixed quadruple of points $X = (x_1, x_2, x_3, x_4)$ in \mathbb{P}^3 .

It is straightforward to see that a joint image is reduced if and only if it contains the n -tuples $(e_1, \dots, e_1), (e_2, \dots, e_2), (e_3, \dots, e_3), (e_4, \dots, e_4)$ where e_1, \dots, e_4 is the standard basis in $(\mathbb{P}^2)^n$. Moreover, the well known fact that the joint image determines the corresponding cameras up to projective equivalence, implies that a reduced joint image determines (and is also determined by) the *point configuration* $\langle c_1, \dots, c_n, x_1, x_2, x_3, x_4 \rangle$ in \mathbb{P}_{n+4}^3 . For this reason, we write $V(\langle c_1, \dots, c_n, x_1, x_2, x_3, x_4 \rangle)$ for the reduced joint image associated with the cameras $P_{c_i, X}$. We mention however that the configuration of the pinholes alone is sufficient to determine the (reduced) joint image's isomorphism class. Indeed, if we denote by $M(c_1, \dots, c_n)$ the *multi-image variety* [4], consisting of n -tuples of lines in $\text{Gr}(1, \mathbb{P}^3)^n$ passing through c_1, \dots, c_n that meet at a point, then for any choice of x_1, x_2, x_3, x_4 we have that

$$V(\langle c_1, \dots, c_n, x_1, x_2, x_3, x_4 \rangle) \cong M(c_1, \dots, c_n) \quad (7)$$

where \cong denotes isomorphism as algebraic varieties.

Finally, we can define a “dual” reduced joint image associated with $k \geq 2$ scene points and four fixed points X :

Definition 13. The dual reduced joint image $\hat{V}(\langle x_1, x_2, x_3, x_4, x'_1, \dots, x'_n \rangle)$ in $(\mathbb{P}^2)^n$ associated with a configuration in \mathbb{P}_{n+4}^3 of n scene points x'_1, \dots, x'_n and a quadruple of “reference points” $X = (x_1, x_2, x_3, x_4)$ in \mathbb{P}^3 is the (the closure of the) set of tuples (u_1, \dots, u_n) given by $u_i = P_{c, X}(x'_i)$ for a varying pinhole c .

According to Carlsson-Weinshall duality we have that

$$V(\langle c_1, \dots, c_n, x_1, x_2, x_3, x_4 \rangle) = \hat{V}(\langle x_1, x_2, x_3, x_4, \hat{c}_1, \dots, \hat{c}_n \rangle), \quad (8)$$

as subvarieties of $(\mathbb{P}^2)^n$, where \hat{c} denotes a Cremona inversion of c relative to X .

References

- [1] A. Coble: *Point sets and allied Cremona groups*. Proceedings of the National Academy of Sciences of the United States of America 1.4, 1915.
- [2] D. Eisenbud, and S Popescu: *The projective geometry of the Gale transform*. Journal of algebra 230, (2000).
- [3] R. Hartley, and A. Zisserman: *Multiple view geometry in computer vision*. Cambridge university press, 2003.
- [4] J. Ponce, B. Sturmfels, and M. Trager: *Congruences and concurrent lines in multi-view geometry*. Advances in Applied Mathematics 88, 2017.
- [5] F. Schaffalitzky, A. Zisserman, R. Hartley, and P.H. Torr: *A six point solution for structure and motion*. In European Conference on Computer Vision 2000.
- [6] B. Sturmfels: *Algorithms in invariant theory*. Springer Science & Business Media, 2008.
- [7] M. Trager, M. Hebert, and J. Ponce: *Coordinate-Free Carlsson-Weinshall Duality and Relative Multi-View Geometry*. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2019.