



HAL
open science

Consistency of silhouettes and their duals

Matthew Trager, Martial Hebert, Jean Ponce

► **To cite this version:**

Matthew Trager, Martial Hebert, Jean Ponce. Consistency of silhouettes and their duals. IEEE Conference on Computer Vision and Pattern Recognition, 2016, Jun 2016, Las Vegas, United States. hal-01287180v2

HAL Id: hal-01287180

<https://hal.science/hal-01287180v2>

Submitted on 13 Apr 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Consistency of silhouettes and their duals

Matthew Trager*
Inria

Martial Hebert
Carnegie Mellon University

Jean Ponce*
École Normale Supérieure / PSL Research University

Abstract

Silhouettes provide rich information on three-dimensional shape, since the intersection of the associated visual cones generates the “visual hull”, which encloses and approximates the original shape. However, not all silhouettes can actually be projections of the same object in space: this simple observation has implications in object recognition and multi-view segmentation, and has been (often implicitly) used as a basis for camera calibration. In this paper, we investigate the conditions for multiple silhouettes, or more generally arbitrary closed image sets, to be geometrically “consistent”. We present this notion as a natural generalization of traditional multi-view geometry, which deals with consistency for points. After discussing some general results, we present a “dual” formulation for consistency, that gives conditions for a family of planar sets to be sections of the same object. Finally, we introduce a more general notion of silhouette “compatibility” under partial knowledge of the camera projections, and point out some possible directions for future research.

1. Introduction

When can a set of 2D silhouettes be projections of the same object in space? This seemingly simple question is related to a variety of practical problems in computer vision, such as multi-view segmentation [10], object recognition [37], and multi-view stereo [12]. Geometric consistency is sometimes taken for granted, when appearance-based features give reasonable evidence that the silhouettes are associated to the same object. However, it is clear that incorporating geometric constraints can be important, either in the process of inference, or for correcting the effects of noisy data.

In this paper, we analyze the notion of “consistency” for silhouettes and for more general closed image sets (Figure 1). We consider opaque objects projected in different images and assume, in our initial setting, the knowledge of all camera parameters. The theory can be seen as a natural extension of classical multi-view geometry, which provides

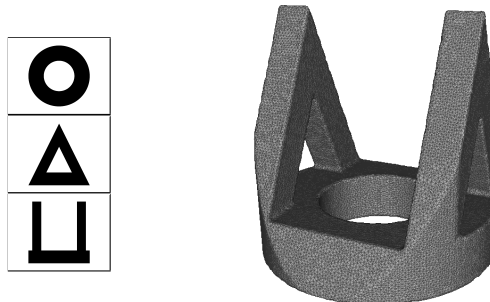


Figure 1: Geometrically consistent silhouettes are feasible projections of a single object.

conditions for *points* (and sometimes *lines*) to be consistent (*i.e.*, to correspond) in terms of given camera projections [11, 16]. This also expands an analogy initiated by the generalized epipolar constraint introduced in [1]. Throughout our discussion, we point out several results that hold for arbitrary closed image sets (or sometimes convex image sets), that have identical counterparts in the theory of point correspondences.

As consistency is clearly not a metric property of silhouettes, it is natural to adopt the framework of *projective geometry* [7], since this eliminates various degenerate situations and allows, for example, to unify the cases of orthographic and perspective projections. This is also typical in multi-view geometry.¹ Another advantage of the projective language is that it provides a homogeneous formulation of *duality*. In particular, we exploit the fact that perspective projections are related to *planar sections* in the dual space [27], to define a very natural “dual” notion of consistency, expressing conditions for a family of planar sets to be sections of the same object.

Finally, we also consider the case of having only *partial* knowledge of the camera parameters, and discuss a more general concept of “compatibility” for silhouettes. This extends a setting first considered in [2].

¹The most significant difference between the euclidean and projective frameworks in our setting is that in the latter case visual cones are *two-sided*. An alternative (but perhaps less natural) approach would have been the use of *oriented projective geometry* [31].

*Willow project team. DI/ENS, ENS/CNRS/Inria UMR 8548.

The overall goal of the paper is not to give the “final” answer to the complex problem of silhouette consistency, but rather to make a first formal foray in that field, with precise definitions that have been missing so far in a general setting, spelling out rigorously what is known in this area and adding a set of new results. For example, we do not deal with algorithmic issues here, and we assume throughout the presentation an ideal setting with no noise. We believe this to be a necessary first step, much in the same way as multi-view geometry initially characterizes *exact* point correspondences, and then makes use of the theory to infer camera parameters from real world data.

Previous work.

The most widespread application of silhouette consistency has been for designing alternatives to point-based methods for camera calibration, required for dealing with smooth and textureless surfaces. Indeed, ever since the seminal work of Rieger [28], the problem of estimating camera motion or calibration parameters using only silhouettes has received considerable attention: see, *e.g.*, [4, 13, 17, 20, 25, 35]. Albeit with some variations, all these methods exploit (more or less directly) the geometric constraints provided by the *epipolar tangencies* [1].

In addition to camera calibration, silhouette consistency has been enforced explicitly for other tasks; for example multi-view segmentation [5], or 3D-reconstruction [8] and recognition [24]. Another interesting “artistic” application is discussed in [26].

There exists limited theoretical work on silhouette consistency, and it is always restricted to special situations. In particular, the problem of determining whether a family of silhouettes can correspond to a real object is considered in [2, 19, 36], but results are only given for the case of orthographic projections and somewhat restricted camera motion. Some theoretical facts, *e.g.*, the fact that epipolar tangency conditions do not imply global consistency, can be found in [4, 6].

Finally, the duality between projections and planar sections is well known in both the euclidean (orthographic) setting [34, 30], as well as in the projective (perspective) case [27, 32]. However, we are not aware of work on consistency for planar sections.

Main contributions.

- We formally introduce a notion of *geometric consistency* for arbitrary closed sets in the case of general projective (perspective) cameras. We present some new results (Propositions 1, 3), and collect others which are scattered in more applied work (Propositions 2, 4).
- We investigate in detail the relationship between pairwise consistency, epipolar tangencies, and our more general notion of consistency (Sections 2.2 and 2.3). We also observe

the existence of special “tangential points”, which play an important role in constraining the consistency for silhouettes.

- We restate the notion of consistency in terms of duality, expressing the condition for planar sets to be sections of the same object. For convex silhouettes, we show that the dual of the visual hull coincides with the convex hull of the dual image of the silhouettes (Proposition 6).
- We define the notion of *compatibility* for silhouettes, which characterizes silhouettes that *may* be geometrically consistent for appropriate camera parameters. This setting generalizes a viewpoint first introduced in [2].

Notation. Our analysis will be coordinate free, so we consider projective cameras as linear maps $\mathcal{M} : \mathbb{P}^3 \setminus \{c\} \rightarrow \mathbb{P}^2$, where c is the camera pinhole or center. The action of a camera will be indicated with $\mathcal{M}(p) = u$. For any camera \mathcal{M} and set of points T in an image, we define the associated *visual cone* as $\mathcal{M}^{-1}(T)$, where \mathcal{M}^{-1} denotes the pre-image set (this definition excludes the pinhole from the visual cone, however this will be irrelevant for our discussion).

2. Consistency of image sets

In this section, we introduce a notion of geometric consistency for arbitrary closed sets in different images. Our definition is very natural, and similar concepts have previously been used to introduce “incoherence” measures for silhouettes [4, 17]. Compared to these works, we focus on analyzing some theoretical properties of consistency, rather than finding strategies for putting it into practice. After a general discussion, we consider the important case of two silhouettes (Section 2.2), pointing out how the consistency condition is basically equivalent to the popular “epipolar tangency” constraint, but only applied to *extremal* tangents. We then investigate in more detail the difference between pairwise consistency, and general consistency, and argue the importance of special “tangential points”, which constrain consistent silhouettes (Section 2.3).

2.1. Basic definitions

Let $\mathcal{M}_1, \dots, \mathcal{M}_n$ be n perspective cameras with distinct centers c_1, \dots, c_n , and let T_1, \dots, T_n be a family of closed sets, one in each image. For example, the sets T_i could be a finite collection of points, curves or closed regions. For each $i = 1, \dots, n$, we let $C_i = \mathcal{M}_i^{-1}(T_i)$ be the visual cone associated to T_i .

Definition 1. *The sets T_1, \dots, T_n are said to be consistent (relative to the cameras $\mathcal{M}_1, \dots, \mathcal{M}_n$) if there exists a non-empty set $R \subseteq \mathbb{P}^3 \setminus \{c_1, \dots, c_n\}$ such that $\mathcal{M}_i(R) = T_i$ for all $i = 1, \dots, n$.*

When T_1, \dots, T_n are consistent for $\mathcal{M}_1, \dots, \mathcal{M}_n$, the

visual hull associated with T_1, \dots, T_n is given by $H = \bigcap_i C_i$, and it is the largest set that projects onto T_1, \dots, T_n .

If all T_1, \dots, T_n are singletons, then consistency reduces to the classical notion of point correspondence [11, 16]; in this case, the visual hull is simply the triangulated 3D-point. Extending this analogy, consistent image sets can be seen as n -tuples of projections of all possible objects in space. Moreover, it is clear that there is a one-to-one correspondence between n -tuples of consistent sets, and the (exact) visual hulls associated with a fixed set of n cameras.

In principle, the concept of visual hull is not well defined if the original silhouettes (or image sets) are not geometrically consistent. This is rarely taken into consideration, and it is customary to define the visual hull simply as the intersection of the cones $\bigcap_i C_i$ for arbitrary (non necessarily consistent) silhouettes: this operation can be justified by noting that if $\bigcap_i C_i$ is not empty, then it is the visual hull associated with the subsets $\tilde{T}_i = \mathcal{M}_i(\bigcap_j C_j) \subseteq T_i$, which will always be consistent. In fact, consistency is clearly equivalent to the fact that $\tilde{T}_i = \mathcal{M}_i(\bigcap_j C_j) = T_i$ for all $i = 1, \dots, n$, or to $T_i \subseteq \mathcal{M}_i(\bigcap_j C_j)$, since the opposite inclusion is always true. We collect a few other simple but useful properties:

Proposition 1. *Let T_1, \dots, T_n be arbitrary closed image sets.*

1. T_1, \dots, T_n are consistent if and only if for each $i = 1, \dots, n$, and for all $\mathbf{u}_i \in T_i$, the visual ray $\mathcal{M}^{-1}(\mathbf{u}_i)$ intersects $\bigcap_{j \neq i} C_j$.
2. T_1, \dots, T_n are consistent if and only if

$$T_i \subseteq \mathcal{M}_i \left(\bigcap_{j \neq i} C_j \right), \quad \forall i \in \{1, \dots, n\}. \quad (1)$$

See Figure 2.

3. If T_1, \dots, T_n are consistent, then any subfamily T_{i_1}, \dots, T_{i_s} is consistent (for the associated cameras $\mathcal{M}_{i_1}, \dots, \mathcal{M}_{i_s}$).

Proof. The first property follows from the fact that $T_i \subseteq \mathcal{M}_i(\bigcap_j C_j)$ can be expressed as $\mathcal{M}_i^{-1}(\mathbf{u}_i) \cap \bigcap_j C_j \neq \emptyset$ for all $\mathbf{u}_i \in T_i$, which in turn is equivalent to $\mathcal{M}_i^{-1}(\mathbf{u}_i) \cap \bigcap_{j \neq i} C_j \neq \emptyset$, since $\mathcal{M}_i^{-1}(\mathbf{u}_i) \subseteq C_i$. The second and third properties are consequences of the first one. \square

This might be a good moment to point out that the notion of geometric consistency discussed in this paper is somewhat independent from a more intuitive (but less formal) concept of “similarity” of appearance. Indeed, consistent silhouettes may actually look completely different (as in the



Figure 2: The silhouettes from Figure 1 are geometrically consistent (relative to three orthogonal orthographic projections), since each silhouette is contained in the reprojection (in gray) of the intersected visual cones associated to the other views. See Proposition 1.

example in Figure 1), while, on the other hand, almost identical silhouettes may be geometrically inconsistent. Thus, the concept might be well suited for being used alongside more traditional feature-based methods for recognition.

Finally, for the rest of the paper we will make the following assumption for all n -tuples of closed image sets T_1, \dots, T_n and cameras $\mathcal{M}_1, \dots, \mathcal{M}_n$:

(A) For each camera center \mathbf{c}_i and visual cone C_j , with $i \neq j$, \mathbf{c}_i does not belong to C_j .

This condition is useful for excluding various degenerate situations; for example, it guarantees that the visual hulls associated with all subfamilies of T_1, \dots, T_n are closed sets.²

2.2. Pairwise consistency

Let us assume that we are given only two image sets T_1, T_2 and two cameras $\mathcal{M}_1, \mathcal{M}_2$. According to Proposition 1, we know that T_1, T_2 are consistent for $\mathcal{M}_1, \mathcal{M}_2$ if and only if

$$T_1 \subseteq \mathcal{M}_1(C_2), \quad \text{and} \quad T_2 \subseteq \mathcal{M}_2(C_1) \quad (2)$$

In other words, we require for each set to be contained in the projection of the visual cone associated to the other one. Pairwise consistency is closely related to the popular *epipolar tangency constraint* [1, 35]. Indeed, we can restate the condition (2) in terms of epipolar geometry as follows.

Proposition 2. *Two image sets T_1, T_2 are consistent if and only if the set of epipolar lines in the first image which intersect T_1 is in epipolar correspondence with the set of epipolar lines in the second image which intersect T_2 .*

Proof. The statement can be seen as a consequence of the first property in Proposition 1. In fact, the epipolar correspondence condition guarantees that for all $i = 1, 2$, and for every point $\mathbf{u}_i \in T_i$ there exists at least one corresponding point $\mathbf{u}_j \in T_j$ ($j \neq i$), so that triangulating all pairs of associated points (i.e., intersecting the cones C_1, C_2) we obtain a set $R \subseteq \mathbb{P}^3 \setminus \{\mathbf{c}_1, \mathbf{c}_2\}$ that projects exactly onto T_1 and T_2 . Note that assumption (A) is equivalent to the fact that, in each image, the epipole lies outside of the given set. \square

²This condition is not actually necessary for all of our results, and weaker assumptions may often be considered. However, for the sake of simplicity, we give a single condition that is valid throughout the paper.

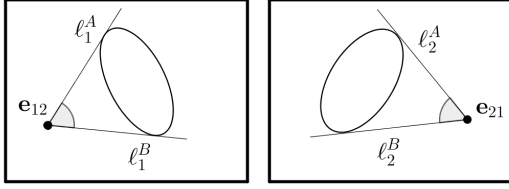


Figure 3: Pairwise consistency: the sets of epipolar lines intersecting each silhouette must be in epipolar correspondence. In particular, the *extremal lines* lines ℓ_1^A, ℓ_2^A and ℓ_1^B, ℓ_2^B must correspond.

This result can also be stated in terms of visual cones: if we consider the pencil of planes passing through the two camera centers, then pairwise consistency is equivalent to the fact that each visual cone meets the same set of planes in the pencil.

If we assume that T_1, T_2 are connected regions bounded by smooth curves, then pairwise consistency basically reduces to the fact that *extremal epipolar tangents* (i.e., “outermost” epipolar lines that are tangent to the contours) are in epipolar correspondence [4] (Figure 3). On the other hand, pairwise consistency *does not* require non-extremal epipolar tangents to be matched. In fact, perhaps somewhat unintuitively, the visual hull generated by two silhouettes with non-extremal epipolar tangencies (matched or unmatched) will always present a complicated topology because it must take into account for all possible self-occlusions (Figure 4). The difference between the two examples shown is just that in the case (*top*) of unmatched internal epipolar tangencies, *any* object which is consistent with the silhouettes must project with occlusions (i.e., the inconsistent “branches” are necessarily projections of different parts of the observed object); on the other hand, in the case (*bottom*) where internal epipolar tangencies are matched, there *exists* an object (different from the visual hull) that has a simple topology and that projects onto the silhouettes without occlusions.

Returning to the case of an arbitrary number of sets, we will say that T_1, \dots, T_n are *pairwise consistent* if each pair $T_i, T_j, i \neq j$ is consistent. From Proposition 1, we know that this holds whenever T_1, \dots, T_n are consistent. The converse is *not* true, as pointed out in [4, 6] (see also Figure 6). However, it would be useful to clarify the practical distinction between these two notions: much of our discussion in the following will be aimed at a better understanding of this issue. For example, in the case of three sets, pairwise consistency often implies that there is at least an “approximate” consistency: this property directly generalizes the fact that three *non-coplanar* visual rays that converge pairwise will always converge [16].

Proposition 3. *Let T_1, T_2, T_3 be connected closed sets that are pairwise consistent. If the centers of the cameras c_1, c_2, c_3 are not collinear, and if each visual cone C_i does*

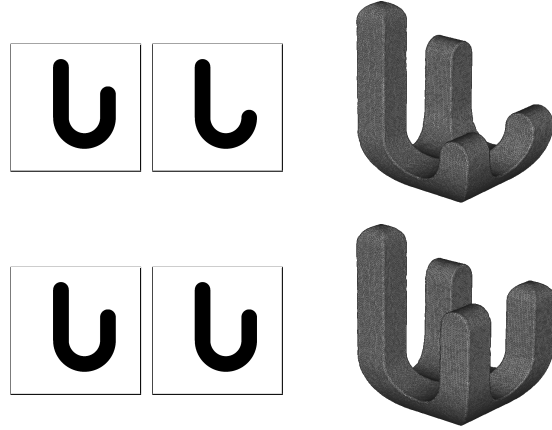


Figure 4: The visual hull generated by two “hook-shaped” silhouettes always presents self-occlusions, whether the internal epipolar tangents are unmatched (*top*) or matched (*bottom*). In these examples we considered orthogonal orthographic projections, but the behavior is completely general.

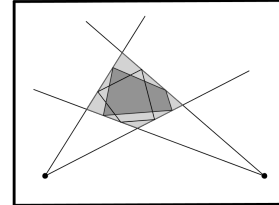


Figure 5: Proof of Proposition 3. The set T_1 (dark gray) and $\mathcal{M}_1(C_2 \cap C_3)$ are both “inscribed” in the quadrilateral $\mathcal{M}_1(C_2) \cap \mathcal{M}_1(C_3)$ (light gray), and must thus intersect.

not intersect the plane spanned by c_1, c_2, c_3 , then $\bigcap_i C_i$ is not empty.

Proof. It is sufficient to prove that, say, $T_1 \cap \mathcal{M}_1(C_2 \cap C_3)$ is not empty. Let $R = \mathcal{M}_1(C_2) \cap \mathcal{M}_1(C_3)$. The assumptions on the centers guarantee that R is a quadrilateral (it is the intersection of two connected projected cones; see Figure 5). From pairwise consistency (2), we know that T_1 is tightly “inscribed” in R , meaning that $T_1 \subseteq R$ and T_1 intersects of the four edges of R . The same holds for $\mathcal{M}_1(C_2 \cap C_3)$: indeed $\mathcal{M}_1(C_2 \cap C_3) \subseteq \mathcal{M}_1(C_2) \cap \mathcal{M}_1(C_3)$ holds, and $\mathcal{M}_1(C_2 \cap C_3)$ intersects all edges of R since C_2, C_3 are themselves pairwise consistent (an extremal epipolar line in $\mathcal{M}_1(C_2)$ is the projection of a line in C_2 which must intersect $C_2 \cap C_3$). The claim follows from continuity arguments: for example, consider paths in T_1 and $\mathcal{M}_1(C_2 \cap C_3)$ connecting different pairs of opposite edges. \square

2.3. From pairwise to general consistency

We have already noted that pairwise consistency does not imply general consistency [4, 6]: in fact, as shown in the example in Figure 6, the two notions are not equivalent even in

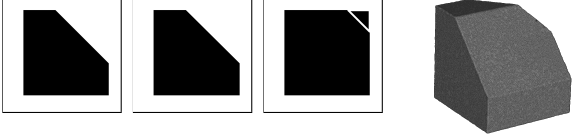


Figure 6: Using orthogonal orthographic projections (that may be thought as “side”, “top”, and “front” views), we observe that two appropriately “clipped” squares and one full square are pairwise consistent but not globally consistent: indeed, the reprojection of the visual hull associated to the first two silhouettes into the third view does not contain the corresponding silhouette (cf. Proposition 1), since the top right corner is missing. In fact, the volume intersection of the three visual cones (shown on the right) does not project onto the third silhouette.

the restricted case of convex image sets (contrary to a claim in [4]). Interestingly, however, we can show that for convex sets (*i.e.*, sets that are convex in an affine chart: we refer to the next section for a discussion on convexity in the projective setting) general consistency is actually guaranteed by “triplet-wise” consistency. A similar fact has been pointed out in [36].

Proposition 4. *Let T_1, \dots, T_n be closed convex sets. If T_i, T_j, T_k are consistent for every $\{i, j, k\} \subseteq \{1, \dots, n\}$ then T_1, \dots, T_n are consistent.*

Proof. From Proposition 1, it is sufficient to prove that for every $i = 1, \dots, n$, and for all $\mathbf{u}_i \in T_i$, we have $\mathcal{M}^{-1}(\mathbf{u}_i) \cap \left(\bigcap_{j \neq i} C_j\right) \neq \emptyset$. Because of convexity, the sets $U_j^i = \mathcal{M}_i^{-1}(\mathbf{u}_i) \cap C_j$ are intervals; moreover, they intersect pairwise because of the assumption of “triplet-wise” consistency: this implies that they all intersect.³ See Figure 7. \square

This statement closely resembles “Helly-type” theorems in computational geometry [33], and is a generalization of the fact that point correspondence is always implied by triplet-wise point correspondence. An inspection of the proof also shows that it is not actually necessary for the sets T_i to be convex, but only for their intersections with epipolar lines (rather than all lines) to be intervals.

Returning to the case of arbitrary image sets, we see from (2) that pairwise consistency is equivalent to

$$T_i \subseteq \bigcap_{j \neq i} \mathcal{M}_i(C_j), \quad i = 1, \dots, n. \quad (3)$$

Moreover, if the sets T_i are consistent there is a chain of inclusions $T_i \subseteq \mathcal{M}_i\left(\bigcap_{j \neq i} C_j\right) \subseteq \bigcap_{j \neq i} \mathcal{M}_i(C_j)$. In fact, as pointed out in [4], pairwise and general consistency can

³To be precise, it should be noted that, although in projective space, none of the cones C_j contain \mathbf{c}_i (because of our assumption (A)), so we may treat the sets U_j^i as intervals on a real line.

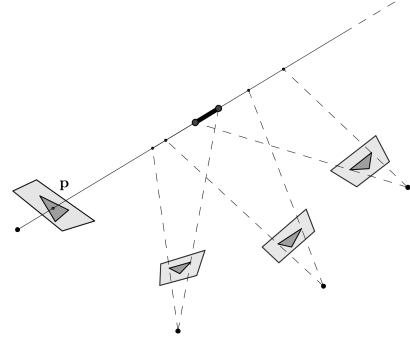


Figure 7: Proof of Proposition 4: the visual ray corresponding to \mathbf{p} meets the other visual cones in intervals that intersect pairwise, and thus must all intersect.

be seen as the same formal condition except for the inverted order of the projection and intersection of the visual cones.

It is also natural to ask ourselves how the consistency condition actually constrains the shape of the image sets. In this regard, we observe that if T_1, \dots, T_n are consistent for $\mathcal{M}_1, \dots, \mathcal{M}_n$, then a set, say T_1 , may be replaced by any strictly larger one \tilde{T}_1 without affecting consistency, provided that $\tilde{T}_1 \subseteq \mathcal{M}_1\left(\bigcap_{j \neq 1} C_j\right)$, because this guarantees that condition (1) remains satisfied in each image. For example, as shown in Figure 8, we are allowed to modify the shape of a set T_i with a small protrusion, in the neighborhood of any point that does not belong to the boundary of $\mathcal{M}_i\left(\bigcap_{j \neq i} C_j\right)$. On the other hands, points belonging to the intersection of T_i with the boundary of $\mathcal{M}_i\left(\bigcap_{j \neq i} C_j\right)$ are more constrained: we will refer to these points as “tangential points”, since the associated visual rays are “tangent” to the visual hull associated with the remaining silhouettes $\bigcap_{j \neq i} C_j$. Tangential points play a crucial role in characterizing consistency, and we intend to investigate their properties in more detail in the future. For the moment, we simply point out that the visual ray corresponding to a tangential point will typically graze the surface of one of the remaining visual cones, so that the ray will correspond to an epipolar tangent (not necessarily extremal) for the associated image (Figure 9, left). In general, it is also possible for the ray to intersect the set $\bigcap_{j \neq i} C_j$ only at an “intersection curve” [23], so that it is not tangent to any cone, and gives rise to a “tangential triple point” on the visual hull surface (Figure 9, right): however, it is not hard to realize that this case will not occur for generic projections of smooth solids, essentially because points on a smooth surface are never visible from three generic viewing directions (one also has to note tangential triple points are “hard points” [22], *i.e.*, points of the visual hull that are guaranteed to belong to the actual object being observed).

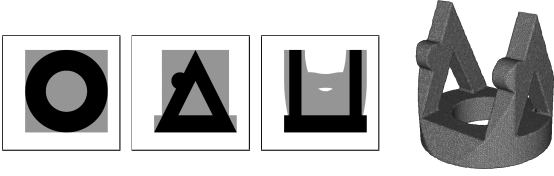


Figure 8: If we modify the second silhouette from Example 1 so that the new silhouette is still contained in the reprojection of the visual hull associated to the other views, consistency is preserved. Note that the reprojection onto the third view is not the same as in the previous case, however condition (1) remains satisfied since the new reprojected set is larger.

3. A dual view of consistency

In this section we revisit the notion of consistency from the viewpoint of *duality*. In particular, we discuss the relationship between consistency for projections, as presented in the previous section, and a different notion of consistency for planar sections of solids. We focus in our presentation on the case of convex sets, since the duality is much simpler in this setting. We defer a more general analysis to future work.

3.1. Duality

The basis of many similar notions of duality is the fact that points and hyperplanes in some n -dimensional space can play symmetric roles. For example, in \mathbb{R}^n any hyperplane through the origin can be represented with its orthogonal vector. In projective space \mathbb{P}^n , all hyperplanes correspond to points of the *dual projective space* $(\mathbb{P}^n)^*$: in fact, any k -dimensional linear subspace H in \mathbb{P}^n , can be associated with an $(n - k)$ -dimensional linear subspace H^* in $(\mathbb{P}^n)^*$, and this duality interchanges the role of “join” and “meet” [7].

This identification of hyperplanes in \mathbb{P}^n with points in $(\mathbb{P}^n)^*$ can be used to introduce notions of “duality” for more general objects in \mathbb{P}^n . For example, if S is a smooth or algebraic hypersurface in \mathbb{P}^n , (the closure of) the set of tangent hyperplanes at points of S forms a *dual hypersurface* in $(\mathbb{P}^n)^*$ [32]. However, dual hypersurfaces will typically have self-intersections (e.g., in the case of curves, crossings correspond to bitangents of the original curve). For this reason, we will consider instead duality for (*properly*) *convex sets* in projective space [15]. In the following, we will say that a set K in \mathbb{P}^n is *convex* when there exists an affine chart (i.e., an appropriate dehomogenization of projective coordinates) so that K is compact and convex in the usual affine sense. This is equivalent to a more analytical definition, for which K is convex if it is the projectivization of a cone \hat{K} in \mathbb{R}^{n+1} that is *convex, closed, and pointed* (i.e., it does not contain any line) [3]. To any convex set K in \mathbb{P}^n , we can now associate a *dual set* K° in $(\mathbb{P}^n)^*$, that can be

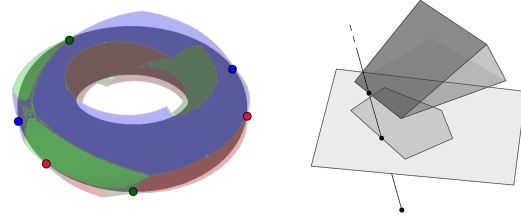


Figure 9: Tangential points are the intersections of a silhouette T_i with the boundary of $\mathcal{M}_i \left(\bigcap_{j \neq i} C_j \right)$. *Left*: the silhouette of a torus and the reprojection of the visual hull generated by three different views. The six tangential points are epipolar tangencies (colors indicate projections of cone surfaces and the associated epipolar tangency points). *Right*: a tangential point may also be the projection of a “tangential triple point”: in this case, the associated viewing ray is not tangent to any visual cone.

characterized geometrically as the closure of the set of all hyperplanes H such that $H \cap K$ is empty. Alternatively, the same set K° is the projectivization of the *polar cone* $(\hat{K})^\circ$ of a convex cone \hat{K} in \mathbb{R}^{n+1} that corresponds to K .

For an appropriate choice of coordinates, convex duality in projective space essentially coincides with the usual concept of polarity for affine sets [3]. However, while many properties of affine polarity only apply for convex sets containing the origin (e.g., biduality: $K^{\circ\circ} = K$), the projective framework is completely “homogeneous”, and does not require similar conditions. Moreover, this setting is useful for dealing with general perspective projections.

All the results in the next section can be understood geometrically (and proven using purely synthetic arguments), by exploiting the characterization of duality as the set of “complementary” hyperplanes. More formal analytical proofs can be given using cones in \mathbb{R}^{n+1} : see the supplementary material for details.

3.2. Duality and visual hulls

A perspective projection $\mathcal{M} : \mathbb{P}^3 \setminus \{c\} \rightarrow \mathbb{P}^2$ with center c also defines a *dual map* $\mathcal{M}^* : (\mathbb{P}^2)^* \rightarrow (\mathbb{P}^3)^*$ that associates lines in \mathbb{P}^2 with planes in \mathbb{P}^3 through c . In particular, the image of \mathcal{M}^* is a plane in the dual space $(\mathbb{P}^3)^*$, namely the dual of the center c^* (i.e., the set of planes containing c). The following important result holds:

Proposition 5. *Let \mathcal{M} be a perspective projection with center c , and let $K \subseteq \mathbb{P}^3 \setminus \{c\}$ be a convex set. Then $\mathcal{M}(K) = L$ is equivalent to*

$$\mathcal{M}^*(L^\circ) = K^\circ \cap c^*. \quad (4)$$

Indeed, the geometric intuition for condition (4) is that a line does not meet L if and only if its preimage does not meet K . This result basically states fact that *projecting* a convex object is equivalent to *taking a planar section* of its

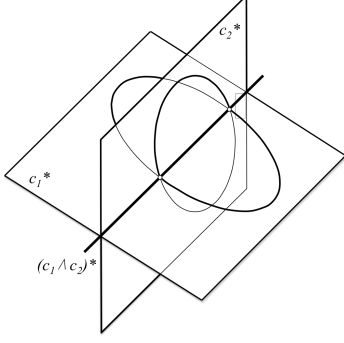


Figure 10: Dual-pairwise-consistency. The dual images of the silhouettes $\mathcal{M}_1^*(T_1^\circ)$ and $\mathcal{M}_2^*(T_2^\circ)$ are sectionally consistent if and only if they have the same intersection with the “dual baseline” $(c_1 \wedge c_2)^*$.

dual. We also refer to [27], where a similar idea is used in a “smooth” setting to investigate the qualitative relationship between image contours and projective shapes.

Motivated by Proposition 5, we introduce a notion of “sectional consistency”. We assume that we are given projective maps $\mathcal{P}_1, \dots, \mathcal{P}_n$ where each $\mathcal{P}_i : \mathbb{P}^2 \rightarrow \mathbb{P}^3$ identifies \mathbb{P}^2 with a plane π_i in \mathbb{P}^3 . The planes π_1, \dots, π_n are distinct.

Definition 2. A family L_1, \dots, L_n of sets in \mathbb{P}^2 is said to be sectionally consistent (relative to the embeddings $\mathcal{P}_1, \dots, \mathcal{P}_n$) if there exists $K \subseteq \mathbb{P}^3$ such that $K \cap \pi_i = \mathcal{P}_i(L_i)$ for all $i = 1, \dots, n$.

This notion of sectional consistency is “dual” to geometric consistency from Definition 1. Indeed:

Proposition 6. A family L_1, \dots, L_n of convex sets in \mathbb{P}^2 is consistent for a set of projections $\mathcal{M}_1, \dots, \mathcal{M}_n$ if and only if $L_1^\circ, \dots, L_n^\circ$ are sectionally consistent for the embeddings $\mathcal{M}_1^*, \dots, \mathcal{M}_n^*$. Moreover, if consistency holds, and H is the visual hull associated with L_1, \dots, L_n , then $H = K^\circ$, where K is the convex hull of $\mathcal{M}_1^*(L_1^\circ), \dots, \mathcal{M}_n^*(L_n^\circ)$.⁴

Proof. The first claim is a direct consequence of (4). For the second part, we can use the fact that for arbitrary closed convex cones K_1, \dots, K_n in \mathbb{R}^N :

$$[\text{Conv}(K_1 \cup \dots \cup K_n)]^\circ = K_1^\circ \cap \dots \cap K_n^\circ. \quad (5)$$

See for example [30]. In fact, it is sufficient to take $K_i = \mathcal{M}_i^*(L_i^\circ)$, and observe that $(\mathcal{M}_i^*(L_i^\circ))^\circ$ is in fact the visual cone C_i (see the supplementary material for details). \square

⁴The convex hull of a closed set contained in an affine chart in \mathbb{P}^n is independent of the choice of projective coordinates only when the set is connected: this is true for $\mathcal{M}_1^*(L_1^\circ) \cup \dots \cup \mathcal{M}_n^*(L_n^\circ)$ since assumption (A) guarantees that $\mathcal{M}_i^*(L_i^\circ) \cup \mathcal{M}_j^*(L_j^\circ)$ is connected for all i, j .

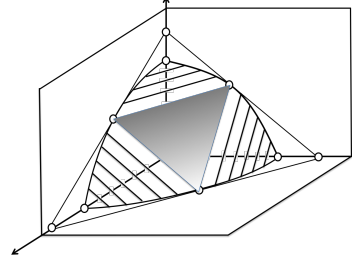


Figure 11: Structure of the “dual visual hull”. The boundary of the dual visual hull is composed of planar curves, ruled surface patches and planar triangular patches: each of these are associated to specific components of the visual hull. See text for details.

Consistency is arguably more intuitive in its dual formulation, since planar sets in space are conceptually easier to grasp than families of cones. For example, the pairwise consistency constraint for two convex silhouettes T_1, T_2 dualizes to the fact that $\mathcal{M}_1^*(T_1^\circ)$ and $\mathcal{M}_2^*(T_2^\circ)$ have the same intersection with the “dual baseline” $(c_1 \wedge c_2)^*$ (here \wedge denotes the “join” of two points): see Figure 10.

We also point out a duality between the “combinatorial” structure of the visual hull H and of its dual $K = H^\circ$. Indeed, by interpreting the supporting planes of H as points on the boundary of K , we see that a plane π supports a cone (patch) C_i if and only if, in the dual space, it represents a point belonging to the planar curve that is the boundary of $\mathcal{M}_i^*(L_i^\circ)$; on the other hand, if a supporting plane π meets the visual hull at the intersection of two cones C_i, C_j (i.e., at an intersection curve [23]), then it is dual to a point on a ruled patch, joining the two boundaries of $\mathcal{M}_i^*(L_i^\circ)$ and $\mathcal{M}_j^*(L_j^\circ)$; finally, if π supports H at the intersection of three cones C_i, C_j, C_k (i.e., a triple point), then it represents a point of the dual hull K that belongs to a planar patch, spanning across the three boundaries of $\mathcal{M}_i^*(L_i^\circ), \mathcal{M}_j^*(L_j^\circ), \mathcal{M}_k^*(L_k^\circ)$ (note that ruled and planar patches are typical for convex hulls of curves in space [21]). See Figure 11.

The notion of sectional consistency given in Definition 2 is reminiscent of questions in geometric tomography [14], or stereology [18]. In tomography, for example, the duality between projections and sections is well studied, but typically in an affine setting that considers only orthographic projections. Nevertheless, it is quite possible that tools from these related fields could provide interesting new insight for problems in computer vision.

4. Compatible silhouettes

Throughout the paper, we have always considered families of silhouettes (or image sets) T_1, \dots, T_n together with known camera projections $\mathcal{M}_1, \dots, \mathcal{M}_n$. However, we can introduce a more general notion of “geometric compatibility”, that can be applied in the case of incomplete knowl-

edge of the camera parameters. More precisely, we define an arbitrary family of silhouettes (or image sets) T_1, \dots, T_n to be *compatible with some partial knowledge* \mathcal{P} of the camera parameters when there *exist* projections $\mathcal{M}_1, \dots, \mathcal{M}_n$ that agree with \mathcal{P} and for which T_1, \dots, T_n are consistent according to Definition 1.

Simply put, this notion characterizes the most general condition for a family of silhouettes (or sets) to be feasible projections of a single object. The study of similar issues was initiated in [2], where the authors assume the external parameters of the cameras to be unknown, and analyze in detail the geometric constraints for compatibility in a particular case (orthographic viewing directions parallel to the same plane).

The first problem raised by the definition of compatibility is to understand whether it actually constrains the silhouettes and the camera parameters, or if instead it is just a vacuous condition. For example, in [2] it is pointed out that if one considers internal as well as external parameters to be unknown, then *any* family of silhouettes will be compatible (they note that considering a convex object, and applying local protrusions with appropriate shapes, one is able to produce arbitrary silhouettes by placing cameras near the surface). This interesting observation, however, would seem to imply that SfM methods can never exploit the geometry of the silhouettes in order to recover camera parameters, since a particular family of silhouettes would provide no information on the viewing conditions. On the other hand, we note that the construction proposed in [2] violates our assumption (A) for *all* pairs of cameras and silhouettes (viewing cones must be extremely “wide”, and thus contain all other centers which lie near the surface of the convex object). We argue instead that by considering only certain regions in the space of parameters, the compatibility of silhouettes *does* provide constraints on the camera projections, even without knowledge of internal parameters. For example, consider two pictures with two arbitrary silhouettes (closed connected regions with non empty interior) T_1, T_2 , one in each image: in this case, it is easy to realize that the fundamental matrices for which T_1, T_2 are consistent *and* such that each epipole lies outside the corresponding silhouette (as in condition (A)), form a space that always has 5 degrees of freedom, compared to 7 in the unconstrained case. Indeed, for any choice of epipoles $e_1 \notin T_1$ and $e_2 \notin T_2$, there is one-parameter family of feasible fundamental matrices, corresponding to the homographies between the pencil of lines through e_1 and e_2 that associate the projections of T_1 and T_2 (as segments of the pencil). On the other hand, assuming $e_1 \in \text{Int}(T_1)$ and $e_2 \in \text{Int}(T_2)$, the space of fundamental matrices for which T_1, T_2 are consistent has 7 degrees of freedom, since the homography of epipolar lines is unconstrained. The fact that a pair of silhouettes imposes (with certain assumptions) two conditions on the epipolar

geometry, justifies methods that have exploited two silhouettes to recover camera parameters when the ambiguity of camera motion had only two degrees of freedom, namely in the case of pure translation motion [29], or when a visible plane provides a homography between the views [9].

In our opinion, a better understanding of the theory of “silhouette compatibility” can be important for several practical reasons, such as spelling out conditions for when silhouettes may or may not be used to determine camera geometry (and possibly help design better algorithms), or similarly to give conditions for a family of silhouettes alone to provide a unique representation of (the visual hull of) an object. Duality might also prove to be a useful tool for investigating these kinds of questions: much in the same way as in Section 3, we realize that the notion of compatibility of silhouettes can be expressed in terms of the compatibility of planar regions, which need to be “assembled” consistently in order to be feasible sections of a single object.

5. Conclusions

We have analyzed in detail the notion of “geometric consistency” for arbitrary image sets, in a setting that can be seen as an extension of traditional multi-view geometry. In the case of convex silhouettes we have also discussed a “dual” interpretation of consistency, expressing conditions for planar sets to be sections of a single object. These concepts lead to a more general notion of silhouette “compatibility”, that does not require (complete) knowledge of the camera parameters.

We plan to extend this work in various directions. On the practical side, our results need to be revisited for dealing with noisy data, and the theory may be used for comparing different measures of “inconsistency” such as the ones considered in [4, 17, 25]. On the theoretical side, in addition to the questions on “compatibility” discussed above, there remains to gain a better understanding of dual consistency for non-convex silhouettes. In geometric tomography, for example, it is typical to study sections of convex bodies [14]; however, for applications in vision, this assumption is probably restrictive. Solutions to all of these problems would be useful in many practical settings, and would help us clarify the fundamental relationship between two-dimensional projections, and the natural concept of “shape”.

Acknowledgments. This work was supported in part by the ERC grant VideoWorld, the Institut Universitaire de France, the Inria - CMU associate team GAYA, and ONR MURI N000141010934.

References

- [1] K. Åström, R. Cipolla, and P. J. Giblin. Generalised epipolar constraints. In *Computer Vision—ECCV’96*, pages 95–108.

- Springer, 1996. 1, 2, 3
- [2] A. Bottino and A. Laurentini. Introducing a new problem: Shape-from-silhouette when the relative positions of the viewpoints is unknown. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 25(11):1484–1493, 2003. 1, 2, 8
- [3] S. Boyd and L. Vandenberghe. *Convex optimization*. Cambridge university press, 2004. 6
- [4] E. Boyer. On using silhouettes for camera calibration. In *Computer Vision—ACCV 2006*, pages 1–10. Springer, 2006. 2, 4, 5, 8
- [5] N. D. Campbell, G. Vogiatzis, C. Hernández, and R. Cipolla. Automatic 3d object segmentation in multiple views using volumetric graph-cuts. *Image and Vision Computing*, 28(1):14–25, 2010. 2
- [6] K. M. G. Cheung. *Visual hull construction, alignment and refinement for human kinematic modeling, motion tracking and rendering*. PhD thesis, Citeseer, 2003. 2, 4
- [7] H. S. M. Coxeter. *Projective geometry*. Springer Science & Business Media, 2003. 1, 6
- [8] D. Cremers and K. Kolev. Multiview stereo and silhouette consistency via convex functionals over convex domains. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 33(6):1161–1174, 2011. 2
- [9] G. Cross, A. W. Fitzgibbon, and A. Zisserman. Parallax geometry of smooth surfaces in multiple views. In *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on*, volume 1, pages 323–329. IEEE, 1999. 8
- [10] A. Djelouah, J.-S. Franco, E. Boyer, F. Le Clerc, and P. Perez. Multi-view object segmentation in space and time. In *Computer Vision (ICCV), 2013 IEEE International Conference on*, pages 2640–2647. IEEE, 2013. 1
- [11] O. Faugeras, Q.-T. Luong, and T. Papadopoulos. *The geometry of multiple images: the laws that govern the formation of multiple images of a scene and some of their applications*. MIT press, 2004. 1, 3
- [12] Y. Furukawa and J. Ponce. Carved visual hulls for image-based modeling. In *Computer Vision—ECCV 2006*, pages 564–577. Springer, 2006. 1
- [13] Y. Furukawa, A. Sethi, J. Ponce, and D. Kriegman. Structure and motion from images of smooth textureless objects. In *Computer Vision—ECCV 2004*, pages 287–298. Springer, 2004. 2
- [14] R. J. Gardner. *Geometric tomography*, volume 6. Cambridge University Press Cambridge, 1995. 7, 8
- [15] W. Goldman. Projective geometry on manifolds. *Lecture notes, University of Maryland*, 1988. 6
- [16] R. Hartley and A. Zisserman. *Multiple view geometry in computer vision*. Cambridge university press, 2003. 1, 3, 4
- [17] C. Hernández, F. Schmitt, and R. Cipolla. Silhouette coherence for camera calibration under circular motion. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 29(2):343–349, 2007. 2, 8
- [18] V. Howard and M. Reed. *Unbiased stereology: three-dimensional measurement in microscopy*. Garland Science, 2004. 7
- [19] D. Jacobs, P. Belhumeur, and I. Jermyn. Judging whether multiple silhouettes can come from the same object. In *Visual Form 2001*, pages 532–541. Springer, 2001. 2
- [20] T. Joshi, N. Ahuja, and J. Ponce. Structure and motion estimation from dynamic silhouettes under perspective projection. In *Computer Vision, 1995. Proceedings., Fifth International Conference on*, pages 290–295. IEEE, 1995. 2
- [21] J. J. Koenderink. *Solid shape*, volume 2. Cambridge Univ Press, 1990. 7
- [22] A. Laurentini. The visual hull for understanding shapes from contours: a survey. *ISSPA*, 1:25–28, 2003. 5
- [23] S. Lazebnik, Y. Furukawa, and J. Ponce. Projective visual hulls. *International Journal of Computer Vision*, 74(2):137–165, 2007. 5, 7
- [24] S. Lazebnik, A. Sethi, C. Schmid, D. Kriegman, J. Ponce, and M. Hebert. On pencils of tangent planes and the recognition of smooth 3d shapes from silhouettes. In *Computer Vision—ECCV 2002*, pages 651–665. Springer, 2002. 2
- [25] P. R. Mendonça, K.-Y. K. Wong, and R. Cipolla. Epipolar geometry from profiles under circular motion. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 23(6):604–616, 2001. 2, 8
- [26] N. J. Mitra and M. Pauly. Shadow art. In *ACM Transactions on Graphics*, volume 28, pages 156–1, 2009. 2
- [27] J. Ponce and M. Hebert. On image contours of projective shapes. In *European Conference on Computer Vision*, 2014. 1, 2, 7
- [28] J. Rieger. Three-dimensional motion from fixed points of a deforming profile curve. *Optics Letters*, 11(3):123–125, 1986. 2
- [29] J. Sato and R. Cipolla. Affine reconstruction of curved surfaces from uncalibrated views of apparent contours. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 21(11):1188–1198, 1999. 8
- [30] R. Schneider. *Convex bodies: the Brunn–Minkowski theory*. Number 151. Cambridge University Press, 2013. 2, 7
- [31] J. Stolfi. *Oriented projective geometry: A framework for geometric computations*. Academic Press, 2014. 1
- [32] E. A. Tevelev. *Projective duality and homogeneous spaces*, volume 133. Springer Science & Business Media, 2006. 2, 6
- [33] C. D. Toth, J. O’Rourke, and J. E. Goodman. *Handbook of discrete and computational geometry*. CRC press, 2004. 5
- [34] P. L. Van Hove. Silhouette-slice theorems. Technical report, DTIC Document, 1986. 2
- [35] K.-Y. K. Wong and R. Cipolla. Reconstruction of sculpture from its profiles with unknown camera positions. *Image Processing, IEEE Transactions on*, 13(3):381–389, 2004. 2, 3
- [36] L. Yi and D. W. Jacobs. Efficiently determining silhouette consistency. In *Computer Vision and Pattern Recognition, 2007. CVPR’07. IEEE Conference on*, pages 1–8. IEEE, 2007. 2, 5
- [37] C. L. Zitnick, J. Sun, R. Szeliski, and S. Winder. Object instance recognition using triplets of feature symbols. Technical report, Technical report, Microsoft Research, Redmond, WA, USA, 2007. 1