



## When is it Possible to Identify 3D Objects From Single Images Using Class Constraints?

RONEN BASRI AND YAEL MOSES

*Department of Computer Science, The Weizmann Institute of Science, Rehovot, 76100, Israel*

**Abstract.** One approach to recognizing objects seen from arbitrary viewpoint is by extracting invariant properties of the objects from single images. Such properties are found in images of 3D objects only when the objects are constrained to belong to certain classes (e.g., bilaterally symmetric objects). Existing studies that follow this approach propose how to compute invariant representations for a handful of classes of objects. A fundamental question regarding the invariance approach is whether it can be applied to a wide range of classes. To answer this question it is essential to study the set of classes for which invariance exists. This paper introduces a new method for determining the existence of invariant functions for classes of objects together with the set of images from which these invariants can be computed. We develop algebraic tests that determine whether the objects in a given class can be identified from single images. These tests apply to classes of objects undergoing affine projection. In addition, these tests allow us to determine the set of views of the objects which are degenerate. We apply these tests to several classes of objects and determine which of them is identifiable and which of their views are degenerate.

**Keywords:** object recognition, class-based recognition, viewpoint invariance

### 1. Introduction

Inferring the identity of objects despite variations due to changes in viewing position is a fundamental problem in object recognition. One problem that arises when we attempt to recognize 3D objects from single 2D images is that some information about the shape of objects is lost with projection. Consequently, a particular image could be the result of projecting any of infinitely many objects. Determining which of these objects have in fact produced the image is impossible unless further constraints are imposed. Two common approaches to solving this problem are the model-based and class-based invariance approaches to recognition.

Model-based methods (Fischler and Bolles, 1981; Huttenlocher and Ullman, 1990; Lowe, 1985; Ullman and Basri, 1991) approach recognition by storing a finite library of object models in the system memory. Given an image, the identity of objects in the image is determined by comparing the image to the models in the

library. To avoid comparing the image to all the models, indexing tables were proposed (Jacobs, 1992; Lamdan and Wolfson, 1988; Weinshall, 1993). To determine whether a given object can produce the image, the object model must contain 3D information of the object. Therefore, model construction generally requires the acquisition and matching of two or more images of the object.

Invariance offers an alternative to the model-based approach. In this approach objects are recognized by extracting image properties that are invariant to changes in viewing positions. Unfortunately, the invariance approach cannot be used directly as a method for identifying general 3D objects from single 2D images since, as has been shown in (Burns et al., 1993; Clemens and Jacobs, 1991; Moses and Ullman, 1992), single 2D images of general 3D objects exhibit no viewpoint invariance. One possible approach then is to recognize objects from two or more views (Hartley, 1993, 1994; Quan, 1995, 1996; Yan et al., 1997).

Alternatively, we may recognize objects using invariant functions if we limit the method to certain classes of objects. Such invariants will distinguish between objects that belong to a certain class, but in general will not distinguish them from objects outside the class. A system for recognizing objects from single images using invariants typically proceeds in two stages. In the first stage the class of the object is recognized, and in the second stage the invariant properties of the class are used for identification. The advantage of using invariant functions for such classes is twofold. First, invariants are computed from single images; hence also object models can be constructed from single images. Second, invariants can distinguish between objects even if models for these objects are not stored in the system memory. Invariant functions have successfully been used for certain classes of objects, e.g., planar, bilateral symmetric, or polyhedral objects (Forsyth et al., 1991; Jones and Poggio, 1995; Moses and Ullman, 1992; Mundy and Zisserman, 1994; Poggio and Vetter, 1992; Rothwell et al., 1992; Sparr, 1992; Sugimoto, 1994; Weiss, 1988), see a recent review in (Zisserman et al., 1995). Nevertheless, restricting the class of objects is not always sufficient for invariants to exist (e.g., Moses and Ullman, 1992).

It is of interest, therefore, to study whether the invariance approach can be extended to handle a wide range of classes of objects. Existing studies are limited to merely a handful of examples of particular classes. These studies, however, offer no general tools to extend their results to other classes of objects. In this paper we develop a method to identify the set of classes for which invariant functions exist. Our method will identify classes for which there exist invariants that can distinguish between *all* the different objects in a given class (as opposed to invariants that distinguish between subsets of the objects, in which case different objects may give rise to the same invariant values). We call these classes *identifiable classes*. We introduce necessary and sufficient conditions for such invariants to exist. For the identifiable classes we can also use our method to detect the *degenerate views*, views from which the invariants cannot be computed. Our method is general and does not depend on a specific choice of a class. While we supply algebraic tests to determine whether objects in a given class can be distinguished from single images we do not supply a method to construct explicit invariants that will identify the objects. Also, the problem of determining the class of the object is not addressed in this paper.

Our approach is based on the following principles. A system can identify an object unequivocally from a given image if and only if that object is the only object, among the objects of interest, that can produce the given image. (Of course one can design systems that identify objects according to, say, maximal-likelihood principles. We do not deal with such a framework here.) We therefore develop a method for exploring the set of *ambiguous images*, images that can be produced by at least two objects of the same class. Clearly, if no further constraints are imposed an object cannot be identified from an ambiguous image. Therefore, given a class of objects we attempt to determine the set of ambiguous images of the class. If we find that all the images in the class are ambiguous we conclude that the class is not identifiable from any view. If some of the images are non-ambiguous, then the class is identifiable, and the ambiguous images are considered degenerate. Below we develop algebraic tests to determine the set of ambiguous images of a class. These tests depend on the choice of projection model. In this paper we introduce tests for objects (given as 3D point sets) undergoing affine projection (that is, 3D affine transformation followed by an orthographic projection).

The rest of the paper is organized as follows. Section 2 introduces the general framework of our work. In Section 3 we use this framework to develop algebraic tests for classes of objects that undergo affine projection. In Section 4 we apply these tests to a number of classes of objects. We conclude with a discussion of our results in Section 5.

## 2. General Framework

Object recognition (“*naming*”) can be expressed as a mapping from a set of images to a set of object names. In general, we would like any image of a given object to be mapped to the object name. However, it is not always possible to define such a mapping. This is because an image of an object depends not only on the shape of the object, but also on its position, the projection model, and other imaging parameters (e.g., camera parameters). As a result, different objects may produce the same image, and therefore the inverse imaging function, from images to objects, cannot be defined in such cases. Below we limit our discussion to images of objects that vary because of camera position (that is, we do not consider changes of illumination direction, background, etc.). In this section we discuss the conditions under which a mapping from a set of

images to a set of object names can be defined. We first outline conditions which are independent of the projection model. Later, we develop algebraic tests for detecting these conditions for the case of objects undergoing an affine projection.

We call two objects,  $O$  and  $O'$ , *equivalent* if and only if every image of  $O$  is also an image of  $O'$  and vice versa. Clearly, no recognition algorithm can discriminate between two equivalent objects. The choice of projection model determines the set of equivalent objects. For example, under the affine projection model every two objects that are related by a 3D affine transformation are equivalent. When two objects are not equivalent (*different*) they may still share some of their images. In this case too no recognition algorithm can identify an object from an ambiguous image. If two objects share an image, they will also share other images which differ by image plane transformations from each other. Such images are called *equivalent images*. Similar to the equivalent objects, the set of equivalent images is determined by the projection model. For example, for rigid objects undergoing orthographic projection two images are equivalent if they are related by a 2D rigid transformations, whereas under the affine projection model two images are equivalent if they are related by a 2D affine transformation. For a given object, we define a *view* to be the set of all equivalent images. It is clear that if two objects share an image they will also share the set of images that are in the same view. It follows that any recognition function cannot identify an object from any of the images in its set of *ambiguous views*. We will therefore alternately consider in the rest of this paper ambiguous views and ambiguous images.

Note that ambiguous views are related but not identical to *accidental views*. The term “accidental view” often is used to describe a view of an object from which a non-stable image is obtained (see, e.g., Weinshall and Werman, 1997). That is, a view is accidental if a small change in the viewing direction will cause a large change in the appearance of the object. Such instability often is caused as a result of a change of aspect of the object due to self occlusion. Degenerate views, in contrast, are views from which an object appears identical to another object of the same class. Degeneracy, therefore, is a property which depends on the class, while accidentalness is a property of the object irrespective of the class. In many cases accidental views are also degenerate. For instance, consider the side views of a planar objects. These views are accidental, since the object appears as a line. Likewise, within the class of

planar objects these views are degenerate since many different planar objects can be confused from images that were taken from these viewing directions.

The following two simple examples demonstrate the use of ambiguous views for determining whether a given class of objects is identifiable. Consider first the extreme case in which the class of objects contains all the objects described by  $n$ -tuples of 3D points. It is easy to show that, for this class of objects, every image is ambiguous. (There are at least two different objects that can project to the image.) Since it is impossible to identify a given object from its ambiguous views, it is impossible in general to construct a mapping that can discriminate between any two objects. This is a weaker version of the result that for the class of 3D point sets there exists no view invariance (Burns et al., 1993; Clemens and Jacobs, 1991; Moses and Ullman, 1992). It follows that to define a recognition mapping it is necessary to restrict the domain of the naming function. In our work we restrict the domain by imposing two types of constraints. First we constrain the set of objects considered by the system (e.g., the class of bilaterally symmetric objects as opposed to the set of general 3D objects). In addition, we may constrain the set of images considered for each object in the class (e.g., excluding the side viewing direction of planar objects). Note that when we consider all 3D point sets it is not sufficient to restrict only the set of views since for this class of objects every view is ambiguous.

Another extreme case of class constraints is provided by the model-based approach to recognition. In this paradigm only finite sets of objects are considered. For a finite set of objects the set of ambiguous views is finite (at most quadratic in the number of objects in the set; this is a consequence of Proposition 2 below), and so by excluding the ambiguous views we obtain identifiable classes. Thus, every finite class of objects is identifiable. Below we concentrate on infinite classes of objects.

Our task then is to determine, given an infinite class of objects, the set of views from which the objects can be identified. Formally, given a class of objects  $\mathcal{C}$ , let  $\mathcal{I}_{\mathcal{C}}$  be the set of all images of  $\mathcal{C}$ , and let  $\mathcal{N}_{\mathcal{C}}$  denote the set of names of objects in  $\mathcal{C}$ , we seek to determine the set of images  $\bar{\mathcal{I}}_{\mathcal{C}} \subseteq \mathcal{I}_{\mathcal{C}}$  on which the naming mapping

$$N_{\mathcal{C}} : \bar{\mathcal{I}}_{\mathcal{C}} \rightarrow \mathcal{N}_{\mathcal{C}}$$

can be defined. Note that usually  $\mathcal{N}_{\mathcal{C}}$  will be isomorphic to  $\mathcal{C}$ . In this case the existence of this mapping would imply that shape reconstruction of objects of the class

from single images is in principle possible if the class constraints are known.

Since without imposing further assumptions it is impossible to determine the identity of an object from an ambiguous view, the ambiguous views will restrict the domain of  $N_C$ . The set  $\tilde{I}_C$ , therefore, will denote the non-ambiguous images of  $C$ .

Our general idea can be summarized as follows. An object  $O \in C$  is identifiable from a view  $\hat{\mathbf{v}}$  if and only if there exists no different object  $O' \in C$  that shares the view  $\hat{\mathbf{v}}$  of  $O$ .

Below we consider the case of objects that undergo affine projection. For these objects we develop algebraic tests that, given a class of objects  $C$ , determine the set  $\tilde{I}_C$  of its non-ambiguous images. When  $\tilde{I}_C \neq \emptyset$  we say that  $C$  is *identifiable*. In such a case we may also use our tests to determine the set of *degenerate views* of  $C$ , the viewing directions from which ambiguous images are produced. When  $\tilde{I}_C = \emptyset$  we say that  $C$  is not identifiable from any view.

### 3. Identification Under Affine Projection

In this section we develop tests for identification of 3D objects undergoing affine projection. An image is an affine projection of a given object if it is obtained by applying a 3D affine transformation to the object followed by an orthographic projection. (Affine projection is the model of choice in several recognition studies (Jacobs, 1992; Koenderink and van Doorn, 1991; Lamdan and Wolfson, 1988; Tomasi and Kanade, 1992; Ullman and Basri, 1991), see (Jacobs, 1992; Basri, 1996) for more on this transformation). The tests are applied to objects given as ordered 3D point sets. We assume here that all the points are visible in every considered view. It is possible, however, to use our results in cases of partial (or self) occlusion by considering subsets of the points of the objects.

#### 3.1. Notation

Given an object  $O$  consisting of  $n$  ordered points  $\mathbf{p}_1 = (x_1, y_1, z_1)^T, \dots, \mathbf{p}_n = (x_n, y_n, z_n)^T$ , we write  $O$  in a matrix form as a  $3 \times n$  matrix  $O = [\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_n]$ . Further, we denote the rows of  $O$  by  $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$  and by  $\text{row}(O)$  the row space of  $O$  (which is the vector space spanned by  $\mathbf{x}, \mathbf{y}, \mathbf{z}$ ). Throughout the paper we shall assume that the object points,  $\mathbf{p}_1, \dots, \mathbf{p}_n$  are non-coplanar. This implies that the rank of  $\text{span}(\text{row}(O), \mathbf{1})$

is 4, where  $\mathbf{1} \in \mathbb{R}^n$  is a vector whose components are all ones. We shall also assume that  $n$  is sufficiently large to induce invariance whenever such invariance exists. Note that we consider two objects that contain the same set of points but ordered differently as different objects. This assumption can be dropped for some of our results (see Section 5).

We denote an image by a  $2 \times n$  matrix  $I$ , that consists of the location of the object points in the image. Let  $I$  be an image obtained as a result of an affine projection of  $O$ .  $I$  is produced by applying a 3D affine transformation to  $O$  followed by an orthographic projection. Specifically, the image position  $\mathbf{q}_i \in \mathbb{R}^2$  of an object point  $\mathbf{p}_i \in \mathbb{R}^3$  of  $O$  taken from a view  $\hat{\mathbf{v}}$  is given by

$$\mathbf{q}_i = A_{\hat{\mathbf{v}}}\mathbf{p}_i + \mathbf{t},$$

where  $A_{\hat{\mathbf{v}}}$  is a  $2 \times 3$  linear matrix of rank 2 and  $\mathbf{t} \in \mathbb{R}^2$ . The subscript  $\hat{\mathbf{v}}$  represents the viewing direction from which the image is observed. This will be defined below. The result of applying this affine transformation to  $O$  can be written in a matrix form as

$$I = A_{\hat{\mathbf{v}}}O + \mathbf{t}\mathbf{1}^T.$$

Two objects,  $O$  and  $O'$ , share an image  $I$  if there exist two  $2 \times 3$  matrices of rank 2,  $A_{\hat{\mathbf{v}}}$  and  $A_{\hat{\mathbf{u}}}$ , and two vectors,  $\mathbf{t}, \mathbf{s} \in \mathbb{R}^2$ , such that

$$I = A_{\hat{\mathbf{v}}}O + \mathbf{t}\mathbf{1}^T = A_{\hat{\mathbf{u}}}O' + \mathbf{s}\mathbf{1}^T.$$

A view under the affine projection model is the set of images that are 2D affine equivalent (namely, related by a 2D affine transformation). That is, both  $I$  and  $I' = A_{2D}I + \mathbf{t}\mathbf{1}^T$  (where  $A_{2D}$  is a  $2 \times 2$ , non-singular matrix) are affine equivalent and therefore belong to the same view of  $O$ . (Below we use the letter  $A$  to denote a  $3 \times 3$  linear transformation,  $A_{2D}$  to denote a  $2 \times 2$  linear transformation, and  $A_{\hat{\mathbf{v}}}$  to denote a  $2 \times 3$  linear transformation.) Clearly, whenever two objects share an image from a given view, they will also share all images in that view. An affine view  $\hat{\mathbf{v}}$  is defined by the set of matrices  $A_{2D}A_{\hat{\mathbf{v}}}$ , where  $A_{2D}$  is an arbitrary  $2 \times 2$  non-singular matrix. It is not difficult to see that under affine projection a view is determined uniquely by a unit vector  $\hat{\mathbf{v}} \in \mathbb{R}^3$  such that  $A_{\hat{\mathbf{v}}}\hat{\mathbf{v}} = \mathbf{0}$  (or, in other words, by a point on a unit sphere), since  $A_{2D}A_{\hat{\mathbf{v}}}\hat{\mathbf{v}} = \mathbf{0}$  for every 2D affine matrix  $A_{2D}$ . Below we shall use the notation  $A_{\hat{\mathbf{v}}}$  to denote a  $2 \times 3$  matrix of a given view,  $\hat{\mathbf{v}}$ .

Given two objects,  $O$  and  $O'$ , we define their  $7 \times n$  joint matrix  $\tilde{J}(O, O')$  to contain the rows of  $O$  and

$O'$  and the vector  $\mathbf{1}$  stacked on below each other in the form:

$$\bar{J}(O, O') = \begin{bmatrix} O \\ O' \\ \mathbf{1}^T \end{bmatrix}. \quad (1)$$

In Section 3.2 below we show that the rank of  $\bar{J}(O, O')$  can be used to determine whether  $O$  and  $O'$  are affine equivalent, share a view, or are completely disjoint.

### 3.2. Tests for Identification

Two objects,  $O$  and  $O'$ , are called *affine-equivalent* if there exists an affine transformation, a  $3 \times 3$  non-singular matrix  $A$  and a vector  $\mathbf{t} \in \mathbb{R}^3$ , such that  $O = AO' + \mathbf{t}\mathbf{1}^T$ . In this case every image of  $O$  is also an image of  $O'$  and vice versa (see Proposition 1 below). Clearly, under the affine projection, it is impossible to distinguish between objects that are affine-equivalent. Otherwise, if  $O'$  cannot be obtained from  $O$  by applying an affine transformation to  $O$  then the two objects are called *affine-different*. Following is a list of necessary and sufficient conditions for affine equivalence.

**Proposition 1.** *The following conditions are equivalent:*

- (a)  $O = AO' + \mathbf{t}\mathbf{1}^T$  for some  $3 \times 3$  non-singular matrix  $A$  and a vector  $\mathbf{t} \in \mathbb{R}^3$ .
- (b)  $\text{rank}(\bar{J}(O, O')) = 4$ .
- (c) Every image of  $O$  is also an image of  $O'$  and vice versa.
- (d)  $O$  and  $O'$  share more than a single view.

**Proof:** (a)  $\Rightarrow$  (b): Assume that  $O = AO' + \mathbf{t}\mathbf{1}^T$ . Since both  $O$  and  $O'$  are non-planar (both ranks of  $\text{span}(\text{row}(O), \mathbf{1})$  and  $\text{span}(\text{row}(O'), \mathbf{1})$  are 4) it follows that  $\text{rank}(\bar{J}(O, O')) \geq 4$ , and since the rows of  $O$  contain only linear combinations of the row vectors of  $O'$  and  $\mathbf{1}$ , it follows that  $\text{rank}(\bar{J}(O, O'))$  is exactly 4. (b)  $\Rightarrow$  (a): Assume that  $\text{rank}(\bar{J}(O, O')) = 4$ . By our assumption the rank of  $\text{span}(\text{row}(O'), \mathbf{1})$  is 4. It follows that  $\text{row}(O) \subseteq \text{span}(\text{row}(O'), \mathbf{1})$ . In particular, it follows that there exists a  $3 \times 3$  matrix  $A$  and a vector  $\mathbf{t} \in \mathbb{R}^3$  such that  $O = AO' + \mathbf{t}\mathbf{1}^T$ . The matrix  $A$  is clearly non-singular because otherwise it will contradict our assumption that the rank of  $\text{span}(\text{row}(O), \mathbf{1})$  is 4.

(a)  $\Rightarrow$  (c): Assume that  $O = AO' + \mathbf{t}\mathbf{1}^T$ . Let an image of  $O$  taken from a viewing direction  $\hat{\mathbf{v}}$  be

$$I = A_{\hat{\mathbf{v}}}O + \mathbf{t}_1\mathbf{1}^T. \quad (2)$$

Denote  $A_{\hat{\mathbf{u}}} = A_{\hat{\mathbf{v}}}A$  and  $\mathbf{t}_2 = A_{\hat{\mathbf{v}}}\mathbf{t} + \mathbf{t}_1$ . The image of  $O'$  which is identical to  $I$  is given by:

$$A_{\hat{\mathbf{u}}}O' + \mathbf{t}_2\mathbf{1}^T. \quad (3)$$

This can be easily verified since

$$\begin{aligned} I &= A_{\hat{\mathbf{v}}}O + \mathbf{t}_1\mathbf{1}^T \\ &= A_{\hat{\mathbf{v}}}(AO' + \mathbf{t}\mathbf{1}^T) + \mathbf{t}_1\mathbf{1}^T = A_{\hat{\mathbf{u}}}O' + \mathbf{t}_2\mathbf{1}^T. \end{aligned} \quad (4)$$

(c)  $\Rightarrow$  (d): Trivial.

(d)  $\Rightarrow$  (a): Assume that  $O$  and  $O'$  share more than one view. Denote the common views of  $O$  by  $\hat{\mathbf{v}}_1$  and  $\hat{\mathbf{v}}_2$  and of  $O'$  by  $\hat{\mathbf{u}}_1$  and  $\hat{\mathbf{u}}_2$  respectively. Let  $I_1$  and  $I_2$  be two common images obtained with the different views,

$$\begin{aligned} I_1 &= A_{\hat{\mathbf{v}}_1}O = B_{\hat{\mathbf{u}}_1}O' + \mathbf{t}\mathbf{1}^T \\ I_2 &= A_{\hat{\mathbf{v}}_2}O = B_{\hat{\mathbf{u}}_2}O' + \mathbf{s}\mathbf{1}^T, \end{aligned} \quad (5)$$

where

$$A_{\hat{\mathbf{v}}_1} = \begin{bmatrix} \mathbf{a}_{x_1} \\ \mathbf{a}_{y_1} \end{bmatrix}, \quad A_{\hat{\mathbf{v}}_2} = \begin{bmatrix} \mathbf{a}_{x_2} \\ \mathbf{a}_{y_2} \end{bmatrix},$$

and

$$B_{\hat{\mathbf{u}}_1} = \begin{bmatrix} \mathbf{b}_{x_1} \\ \mathbf{b}_{y_1} \end{bmatrix}, \quad B_{\hat{\mathbf{u}}_2} = \begin{bmatrix} \mathbf{b}_{x_2} \\ \mathbf{b}_{y_2} \end{bmatrix},$$

are non-singular matrices ( $\mathbf{a}_{x_1}, \dots, \mathbf{b}_{y_2}$  denote their row vectors), and  $\mathbf{t} = (t_x, t_y)^T$  and  $\mathbf{s} = (s_x, s_y)^T$  are two translation vectors. (Without loss of generality we may translate  $I_1$  and  $I_2$  so that  $A_{\hat{\mathbf{v}}_1}$  and  $A_{\hat{\mathbf{v}}_2}$  will bring  $O$  to the images with no translation.) Since  $I_1$  and  $I_2$  are obtained from different views (and so they are not related by a 2D affine transformation), then without a loss of generality  $\mathbf{a}_{x_2}$  is linearly independent of  $\mathbf{a}_{x_1}$  and  $\mathbf{a}_{y_1}$ . Denote by

$$A = \begin{bmatrix} \mathbf{a}_{x_1} \\ \mathbf{a}_{y_1} \\ \mathbf{a}_{x_2} \end{bmatrix}, \quad B = \begin{bmatrix} \mathbf{b}_{x_1} \\ \mathbf{b}_{y_1} \\ \mathbf{b}_{x_2} \end{bmatrix}, \quad \text{and } \mathbf{t}' = \begin{bmatrix} t_x \\ t_y \\ s_x \end{bmatrix}.$$

From (5) it follows that

$$AO = BO' + \mathbf{t}\mathbf{1}^T. \quad (6)$$

By construction,  $A$  is non-singular. We can therefore choose  $A' = A^{-1}B$  and  $\mathbf{t}' = A^{-1}\mathbf{t}$  to obtain

$$O = A'O' + \mathbf{t}'\mathbf{1}^T. \quad (7)$$

The matrix  $A'$  is clearly non-singular because otherwise it will contradict our assumption that the rank of  $\text{span}(\text{row}(O), \mathbf{1})$  is 4.  $\square$

To determine the set of ambiguous views of a class we need to develop necessary and sufficient conditions for two affine different objects to share a view. These conditions are specified in Proposition 2 below.

**Proposition 2.** *The following conditions are equivalent:*

- (a) *There exist a single direction  $\hat{\mathbf{v}} \in \mathbb{R}^3$  and translation  $\mathbf{t} \in \mathbb{R}^2$  such that  $A_{\hat{\mathbf{v}}}O = A_{\hat{\mathbf{u}}}O' + \mathbf{t}\mathbf{1}^T$ , where  $A_{\hat{\mathbf{v}}}$  and  $A_{\hat{\mathbf{u}}}$  are  $2 \times 3$  matrices of rank 2.*
- (b)  *$\text{rank}(\bar{J}(O, O')) = 5$ .*
- (c)  *$O = AO' + \mathbf{t}\mathbf{1}^T + \hat{\mathbf{v}}\eta^T$ , where  $A$  is a  $3 \times 3$  matrix and  $\text{rank}(A) \geq 2$ ,  $\mathbf{t} \in \mathbb{R}^3$ ,  $\hat{\mathbf{v}} \in \mathbb{R}^3$  is a non-zero vector, and  $\eta \in \mathbb{R}^n$  is orthogonal to  $\text{span}(\text{row}(O'), \mathbf{1})$ . ( $O'$  shares the view  $\hat{\mathbf{v}}$  of  $O$ .)*
- (d)  *$O = AO' + \mathbf{t}\mathbf{1}^T + \hat{\mathbf{v}}\eta^T$ , where  $A$  is a  $3 \times 3$  matrix and  $\text{rank}(A) = 3$ ,  $\hat{\mathbf{v}} \in \mathbb{R}^3$  is a non-zero vector, and  $\eta \in \mathbb{R}^n$  where  $\eta \notin \text{span}(\text{row}(O'), \mathbf{1})$ . ( $O'$  shares the view  $\hat{\mathbf{v}}$  of  $O$ .)*

**Proof:** (a)  $\Rightarrow$  (b): Assume that there exists a single direction  $\hat{\mathbf{v}}$  such that  $A_{\hat{\mathbf{v}}}O = A_{\hat{\mathbf{u}}}O' + \mathbf{t}\mathbf{1}^T$ . Let  $B = [A_{\hat{\mathbf{v}}}, -A_{\hat{\mathbf{u}}}, -\mathbf{t}]$  be a  $2 \times 7$  matrix. We obtain that

$$B\bar{J}(O, O') = 0. \quad (8)$$

Since  $\text{rank}(A_{\hat{\mathbf{v}}}) = \text{rank}(A_{\hat{\mathbf{u}}}) = 2$  it follows that  $\text{rank}(B) = 2$ . This implies that  $\text{rank}(\bar{J}(O, O')) \leq 5$ . Since  $O$  is non-planar (the rank of  $\text{span}(\text{row}(O), \mathbf{1})$  is 4) it follows that  $\bar{J}(O, O') \geq 4$ . It must be different than 4 because  $O$  and  $O'$  share only a single view, and therefore they cannot be affine-equivalent (Proposition 1, (d)  $\Rightarrow$  (b)). Therefore,  $\text{rank}(\bar{J}(O, O')) = 5$ .

(b)  $\Rightarrow$  (c): Assume that  $\text{rank}(\bar{J}(O, O')) = 5$ . At least one of the rows of  $O$  must be independent of the rows of  $O'$  and the vector  $\mathbf{1}$ . Assume without the loss of generality that this row is  $\mathbf{x}_1$ . It follows that there exists

a vector  $\eta$  which is orthogonal to the  $\text{span}(\text{row}(O'), \mathbf{1})$  such that

$$\mathbf{x}_1 = c_x\mathbf{x}_2 + c_y\mathbf{y}_2 + c_z\mathbf{z}_2 + c_1\mathbf{1} + \eta, \quad (9)$$

where  $\mathbf{x}_2$ ,  $\mathbf{y}_2$ , and  $\mathbf{z}_2$  are the row vectors of  $O'$ . Since  $\text{rank}(\bar{J}(O, O')) = 5$ , it follows that  $\mathbf{y}_1, \mathbf{z}_1 \in \text{span}\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{y}_2, \mathbf{z}_2, \mathbf{1}\}$  or equivalently  $\mathbf{y}_1, \mathbf{z}_1 \in \text{span}\{\eta, \mathbf{x}_2, \mathbf{y}_2, \mathbf{z}_2, \mathbf{1}\}$ . Consequently,

$$O = AO' + \mathbf{t}\mathbf{1}^T + \hat{\mathbf{v}}\eta^T. \quad (10)$$

Finally, since  $\mathbf{1} \notin \text{row}(O)$  it follows that  $\text{rank}(O - \mathbf{t}\mathbf{1}^T) = 3$ . The rank of  $\hat{\mathbf{v}}\eta^T$  is 1, therefore,  $\text{rank}(AO') \geq 2$ . Since  $\text{rank}(O') = 3$  it follows that  $\text{rank}(A) \geq 2$ .

We next show that the view of  $O$  shared with  $O'$  is  $\hat{\mathbf{v}}$ . Let  $A_{\hat{\mathbf{v}}}$  be a  $2 \times 3$  matrix such that  $A_{\hat{\mathbf{v}}}\hat{\mathbf{v}} = 0$  then  $O'$  shares the view  $\hat{\mathbf{v}}$  with  $O$ :

$$A_{\hat{\mathbf{v}}}O = A_{\hat{\mathbf{v}}}(AO' + \mathbf{t}\mathbf{1}^T + \hat{\mathbf{v}}\eta^T) = A_{\hat{\mathbf{v}}}AO' + A_{\hat{\mathbf{v}}}\mathbf{t}\mathbf{1}^T. \quad (11)$$

(c)  $\Rightarrow$  (d): Assume that  $O = AO' + \mathbf{t}\mathbf{1}^T + \hat{\mathbf{v}}\eta^T$  where  $\hat{\mathbf{v}} = (v_x, v_y, v_z)^T$ ,  $\text{rank}(A) = 2$  and

$$A = [\mathbf{a}_x, \mathbf{a}_y, \mathbf{a}_z]^T.$$

Assume without a loss of generality that  $\mathbf{a}_z$  depends on  $\mathbf{a}_x$  and  $\mathbf{a}_y$ , and that  $\mathbf{a}_x$  is independent of  $\mathbf{a}_y$ . Let  $\mathbf{a} = \mathbf{a}_x \times \mathbf{a}_y$ , let

$$B = [(\mathbf{a}_x - v_x(\mathbf{a}_z - \mathbf{a})), (\mathbf{a}_y - v_y(\mathbf{a}_z - \mathbf{a})), v_z\mathbf{a}]^T,$$

and let  $\eta' = \eta + (\mathbf{a}_z - \mathbf{a})^T O$ . It can be readily verified that

$$O = AO' + \hat{\mathbf{v}}\eta'^T = BO' + \hat{\mathbf{v}}\eta'^T. \quad (12)$$

By construction  $\text{rank}(B) = 3$  and  $\eta' \notin \text{span}(\text{row}(O), \mathbf{1})$ .

It can be shown as in the previous case that the view of  $O$  shared with  $O'$  is  $\hat{\mathbf{v}}$ .

(d)  $\Rightarrow$  (a): Assume that  $O = AO' + \mathbf{t}_1\mathbf{1}^T + \hat{\mathbf{v}}\eta^T$ . Consider the images of  $O$  taken from a direction  $\hat{\mathbf{v}}$ . Let  $A_{\hat{\mathbf{v}}}$  be a  $2 \times 3$  matrix of rank 2 such that  $A_{\hat{\mathbf{v}}}\hat{\mathbf{v}} = 0$ . It follows that

$$A_{\hat{\mathbf{v}}}O = A_{\hat{\mathbf{v}}}(AO' + \mathbf{t}_1\mathbf{1}^T + \hat{\mathbf{v}}\eta^T) = A_{\hat{\mathbf{v}}}AO' + A_{\hat{\mathbf{v}}}\mathbf{t}_1\mathbf{1}^T. \quad (13)$$

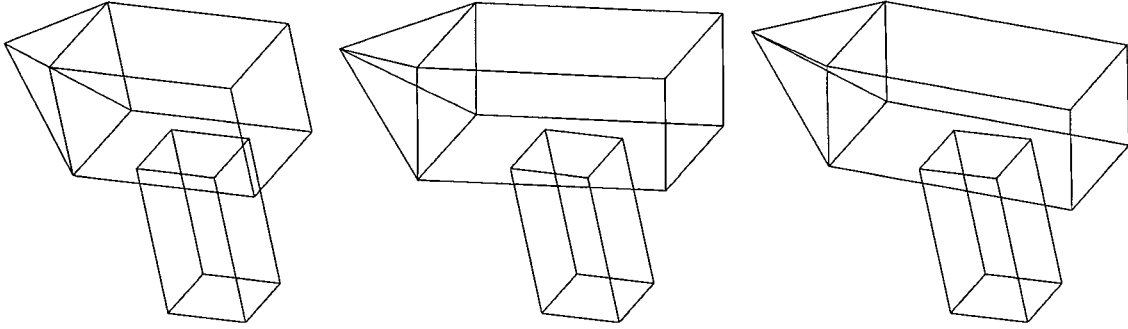


Figure 1. Three hammer-like shapes composed of identical parts but differ in the shapes of their heads by generic affine transformations.

Let  $A_{\hat{\mathbf{u}}} = A_{\hat{\mathbf{v}}}A$  and  $\mathbf{t} = \mathbf{A}_{\hat{\mathbf{v}}}\mathbf{t}_1$ , then we obtain

$$A_{\hat{\mathbf{v}}}O = A_{\hat{\mathbf{u}}}O' + \mathbf{t}\mathbf{1}^T. \quad (14)$$

Since  $\eta \notin \text{span}(\text{row}(O'), \mathbf{1})$  then  $O$  and  $O'$  cannot be affine equivalent. Hence, according to Proposition 1,  $\hat{\mathbf{v}}$  must be unique.  $\square$

#### 4. Examples

In this section we apply the tests developed in Section 3 to several classes of objects. For each class we determine the set of views from which the objects of the class can be identified. Some of these classes are shown to be identifiable from most or all views. Other classes are shown to be non-identifiable from any view. We begin with classes of objects that are composed of the same set of parts, but the relative position and orientation of their parts may vary across objects (Section 4.1). We then proceed to discuss classes of objects which have two identical parts, except that one part may appear at different position and orientation with respect to the second part. In this case the shape of the parts may vary across objects (Section 4.2). Next, we consider classes that contain combinations of sub-structures, that is one part of the object is a function of the other parts (Section 4.3). Finally we consider classes of objects that can be expressed as combinations of prototype objects (Section 4.4).

##### 4.1. Objects With Same Set of Parts

Many classes of interest contain objects all of which are composed of the same set of parts. The identity of an object within the class is defined by the exact position and orientation of one part with respect

to another, and by the relative size or stretch of one part with respect to another. Some examples for such classes are hammers, screwdrivers, cars, and chairs. Figure 1 shows three hammer-like shapes with identical handles and heads which differ by affine transformations. Formally, a class  $\mathcal{C}_P$  is defined by an ordered set of parts  $\{Q_1, \dots, Q_m\}$  where  $Q_i$  are  $3 \times n_i$  matrices for some arbitrary  $n_i > 3$ . An object  $O \in \mathcal{C}_P$  is given by  $O = [P_1, \dots, P_m]$ , where the  $3 \times n_i$  matrices  $P_i = f_i(Q_i)$  describe the parts of  $O$ . The functions  $f_i$  determine the identity of the objects in the class.

Below we limit our discussion to  $\{f_i\}$  that represent affine transformations. In this case, an object  $O \in \mathcal{C}_P$  can be written in a matrix notation as  $O = [P_1, \dots, P_m]$ , where  $P_i = B_i Q_i + \mathbf{s}_i \mathbf{1}^T$  for  $1 \leq i \leq m$ .  $B_i$  are  $3 \times 3$  non-singular matrices and  $\mathbf{s}_i \in \mathcal{R}^3$ ; both vary between the different objects in the class. Below we assume that every part  $Q_i$  is non-planar, that is, the rank of  $\text{span}(\text{row}(Q_i), \mathbf{1})$  is 4.

Note that this representation of a class is non-unique. It is possible to transform every  $Q_i$  by the same affine transformation, and still obtain the same set of objects by compensating for this affine transformation in  $B_i$  and  $\mathbf{s}_i$ . Likewise, the representation of an object is non-unique since it is possible to transform the object by an affine transformation and obtain an equivalent representation of the object. The ambiguity in the representation of the object will be handled by applying Proposition 1. Notice, however, that transforming every part by a different affine transformation would deform the object and will not produce an equivalent representation of the object.

The next proposition establishes that if the shape of the objects in a class may change by applying general affine transformations to their parts then the class is non-identifiable from any view. If however we restrict the transformation applied to the parts to scaling and

stretching along any three fixed directions we obtain classes which are identifiable from almost all views. This implies, in particular, that it is possible in principle to extract the relative scale and stretch of objects' parts from single images.

**Proposition 3.** *Given  $m > 1$  non-planar object parts denoted by  $3 \times n$  matrices  $Q_i (1 \leq i \leq m)$ , the class  $\mathcal{C}_P = \{[P_1, \dots, P_m] \mid P_i = B_i Q_i + \mathbf{s}_i \mathbf{1}^T\}$  is:*

1. *Non-identifiable for general affine transformation, that is, for arbitrary  $3 \times 3$  non-singular matrices  $B_i$  and translations  $\mathbf{s}_i \in \mathbb{R}^3$ .*
2. *Non-identifiable for pure translation, that is, for  $B_i = I$  and arbitrary  $\mathbf{s}_i \in \mathbb{R}^3$ .*
3. *Identifiable from almost all views for scaling and stretching along any three fixed directions. For example, scaling and stretching along the three primary axes correspond to diagonal  $B_i$  and constant  $\mathbf{s} \in \mathbb{R}$ . In this case degenerate views are obtained only for  $\hat{\mathbf{v}} = (v_1, v_2, v_3)^T$  such that  $v_i = 0$  for some  $1 \leq i \leq 3$ .*

**Proof:**

1. It will be sufficient to show that this class is not identifiable even if we restrict the translations  $\mathbf{s}_i$  to be the same across objects. To show this we need to show that for every choice of object  $O \in \mathcal{C}_P$  and every choice of view  $\hat{\mathbf{v}}$  there exists an affine-different object  $O' \in \mathcal{C}_P$  that shares the view  $\hat{\mathbf{v}}$  of  $O$ . Let  $O = [P_1, \dots, P_m]$  where  $P_i = B_i Q_i + \mathbf{s}_i \mathbf{1}^T$  for  $1 \leq i \leq m$ , and let  $O' = [P'_1, \dots, P'_m]$  where  $P'_i = B'_i Q_i + \mathbf{s}_i \mathbf{1}^T$  for another set of  $B'_i$  and for  $1 \leq i \leq m$ . Since the matrices  $B_i$  determine the identity of the object it follows that given  $\{B_i\}$  and  $\hat{\mathbf{v}}$  we need to find matrices  $\{B'_i\}$  and a view  $\hat{\mathbf{u}}$  such that  $O$  and  $O'$  will share an image. We set  $B'_1 = B_1$  and  $B'_i = B_i + \hat{\mathbf{v}} \mathbf{r}_i^T$  for some non-zero vector  $\mathbf{r}_i \in \mathbb{R}^3$  such that  $\text{rank}(B'_i) = 3$  and  $B'_i \neq I$ . It is readily verified that  $O$  and  $O'$  are affine-different (since the first parts of  $O$  and  $O'$  are related by the identity matrix, while the other parts are not). Furthermore, the two objects  $O$  and  $O'$  share an image when both are taken from the viewing direction  $\hat{\mathbf{v}}$ . That is,  $A_{\hat{\mathbf{v}}} O = A_{\hat{\mathbf{v}}} O'$ . To prove it, it is sufficient to show that  $A_{\hat{\mathbf{v}}} P'_i - A_{\hat{\mathbf{v}}} P_i = 0$ . This can be easily verified as follows.

$$A_{\hat{\mathbf{v}}}(P'_i - P_i) = A_{\hat{\mathbf{v}}}(B'_i - B_i)Q_i = A_{\hat{\mathbf{v}}}\hat{\mathbf{v}}\mathbf{r}_i^T Q_i,$$

and since  $A_{\hat{\mathbf{v}}}\hat{\mathbf{v}} = 0$  we obtain that  $A_{\hat{\mathbf{v}}}(P'_i - P_i) = 0$ .

2. To show this we need to show that for every choice of object  $O$  (determined by  $\mathbf{s}_i$ ) and view  $\hat{\mathbf{v}}$  there exists an affine different object  $O'$  (determined by  $\mathbf{s}'_i \neq \mathbf{s}_i$ ) that shares the view  $\hat{\mathbf{v}}$  of  $O$ . Let  $O = [P_1, P_2, \dots, P_m]$  such that  $P_i = Q_i + \mathbf{s}_i \mathbf{1}^T$ , and let  $O' = [P'_1, P'_2, \dots, P'_m]$  such that  $P'_i = Q_i + \mathbf{s}'_i \mathbf{1}^T$ . We set  $\mathbf{s}'_1 = \mathbf{s}_1$  and  $\mathbf{s}'_i = \mathbf{s}_i + \lambda_i \hat{\mathbf{v}}$  for  $2 \leq i \leq m$  for some arbitrary  $\lambda_i \neq 0$ . In this case  $O$  and  $O'$  are affine-different since their first parts are related by a different translation than their second parts, and they share the view  $\hat{\mathbf{v}}$  of  $O$  since

$$A_{\hat{\mathbf{v}}}(P'_i - P_i) = A_{\hat{\mathbf{v}}}(\mathbf{s}'_i - \mathbf{s}_i) = \lambda_i A_{\hat{\mathbf{v}}}\hat{\mathbf{v}} = 0.$$

3. We show this for scaling and stretching along the three primary axes (that is, for diagonal  $B_i$  and constant  $\mathbf{s}$ ). Note that since we can redefine the class by applying a (possibly different) affine transformation to each part this will imply that the class is identifiable for stretching in any three fixed directions. To show that this class is identifiable we need to show that, given an object  $O$  and a view  $\hat{\mathbf{v}}$ , any object  $O'$  that shares the view  $\hat{\mathbf{v}}$  of  $O$  must be affine equivalent to  $O$ . Let  $O = [P_1, \dots, P_m]$  where  $P_i = B_i Q_i + \mathbf{s}_i \mathbf{1}^T$ , and let  $O' = [P'_1, \dots, P'_m]$  where  $P'_i = B'_i Q_i + \mathbf{s}_i \mathbf{1}^T$  for another set of  $B'_i$ . Since  $O$  and  $O'$  share the view  $\hat{\mathbf{v}}$  of  $O$  then there exists a view  $\hat{\mathbf{u}}$  such that

$$A_{\hat{\mathbf{v}}} O = A_{\hat{\mathbf{u}}} O' + \mathbf{t} \mathbf{1}^T.$$

Without loss of generality, we can assume that  $B_1 = B'_1 = I$  and  $\mathbf{s}_1 = \mathbf{s}'_1 = 0$  (since we can choose to consider any of the objects that are affine equivalent to  $O'$ ). This implies that

$$\begin{aligned} A_{\hat{\mathbf{v}}} P_1 &= A_{\hat{\mathbf{u}}} P'_1 + \mathbf{t} \mathbf{1}^T \\ A_{\hat{\mathbf{v}}}(B_i Q_i + \mathbf{s}_i \mathbf{1}^T) &= A_{\hat{\mathbf{u}}}(B'_i Q_i + \mathbf{s}_i \mathbf{1}^T + \mathbf{t} \mathbf{1}^T). \end{aligned}$$

The first equation implies that  $\hat{\mathbf{u}} = \hat{\mathbf{v}}$  and that  $\mathbf{t} = 0$  (since  $P_1 = P'_1$  and  $P_1$  is non-planar), and so the second equation becomes

$$A_{\hat{\mathbf{v}}}(B_i - B'_i)Q_i = 0.$$

Since  $Q_i$  is non-planar we obtain that  $A_{\hat{\mathbf{v}}}(B_i - B'_i) = 0$ . If  $B_i = B'_i$  for  $1 \leq i \leq m$  then  $O$  and  $O'$  are identical. It follows that  $B_i - B'_i = \hat{\mathbf{v}} \mathbf{r}^T$  for some  $\mathbf{r} \neq 0$  (since the only vector  $\mathbf{w}$  for which  $A_{\hat{\mathbf{v}}}\mathbf{w} = 0$  is of the form  $\mathbf{w} = \lambda \hat{\mathbf{v}}$ ). However, since both  $B_i$



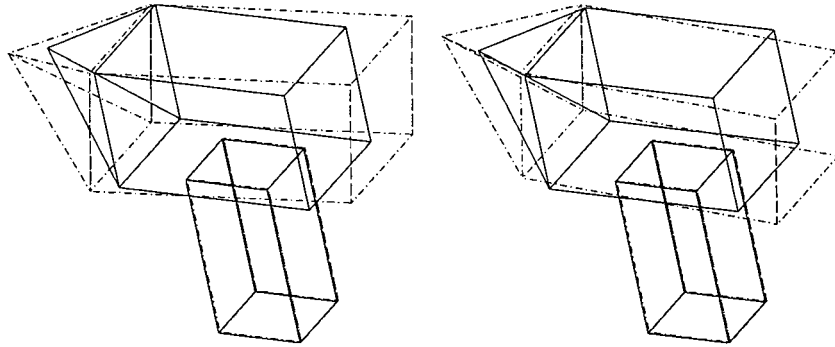


Figure 2. An overlay of the leftmost hammer (solid) with the other two hammers (dashed). Notice that the handles of the hammers are identical whereas their heads differ.

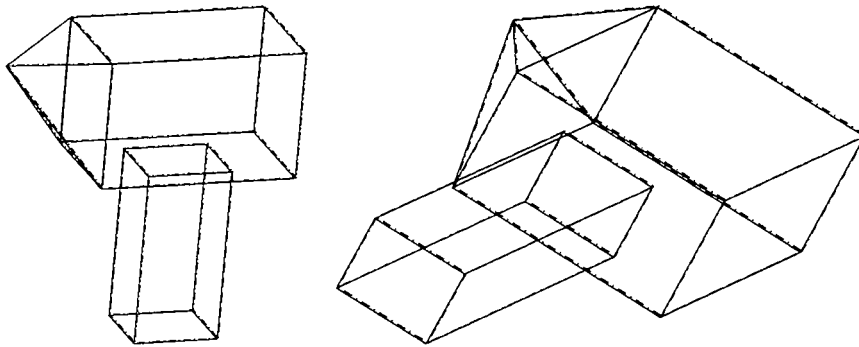


Figure 3. An overlay of the leftmost hammer (solid) and each of the other two hammers (dashed) taken from their common views. The common views are  $(0.29, -0.41, 0.87)$  and  $(0.32, -0.73, 0.60)$ . Notice that for every view of the leftmost hammer we can produce infinitely many different hammer-like shapes that look identical to this hammer.

and  $B'_i$  are diagonal also  $\hat{\mathbf{v}}\mathbf{r}^T$  must be diagonal. But since  $\hat{\mathbf{v}}$  contains no zero components, the only way to obtain zeros in all non-diagonal components is by setting  $\mathbf{r} = 0$ . This, however, will imply again that  $B_i = B'_i$  for  $1 \leq i \leq m$ , and so  $O$  and  $O'$  are identical.  $\square$

An example to these results is shown in Figs. 1–3. In these figures we show three hammer-like objects with identical handles, but their heads differ by an arbitrary linear transformation. We constructed these hammers as follows. We first constructed the leftmost hammer. Then, we arbitrarily selected two views and modified the head of the hammer according to the proof of Proposition 3(1) so as to obtain two different hammers that will share those views with our original hammer. Note that according to the proposition we could do this to any desired view, and that at every view we could find infinitely many different hammers (determined by the choice of  $\mathbf{r}$ ) that share this view with our original

hammer. It follows that the class of objects with the same parts which may differ by an arbitrary affine transformation is non-identifiable. Notice that according to Proposition 3(3) if we allow the parts to only scale and stretch in three fixed directions we would not be able to find two hammers with a common view except for a small set of degenerate views.

A special case of this proposition is the case that the class composed of objects of two identical parts ( $Q_1 = Q_2$ ). Notice that the actual shape of the parts was not used in the proof, so even the negative results extend to classes such that all the parts of a given object are identical. In Section 4.2 below we further explore the case of objects such that the parts of a given object are identical, but the parts may differ across the objects in the class.

We can conclude that a set of objects that consist of identical set of parts cannot be identified if the objects differ by arbitrary affine transformation or translation of their parts. However, if the objects in the class differ

by scaling or stretching of the parts, then they can be identified from almost all of their images.

#### 4.2. Repeated Structures

The next set of classes that we consider consists of objects each of which contains two identical, non-planar parts except that these parts are related by an affine transformation (so one part is translated, rotated, stretched, or reflected with respect to the other). Real objects that contain symmetries (e.g., faces, chairs) often belong to this class. An object from the class  $\mathcal{C}_{B,s}$  has the following form:

$$O = [P, BP + \mathbf{s}\mathbf{1}^T],$$

with  $B$   $3 \times 3$  non-singular matrix and  $\mathbf{s} \in \mathbb{R}^3$ . In contrast to the classes discussed in Section 4.1 now the shape of the parts (denoted by  $P$ ) may also vary across objects. We consider below two cases. First we consider the case that the affine transformation relating the two parts may vary across objects. Then we consider the case that this affine transformation is the same for all the objects in the class. A special case for this latter class is the class of bilaterally symmetric objects, in which the identical parts are related by a reflection.

Note that for repeated structures the two parts in an object are identical up to an affine transformation. However, we assume that in the image we know which is the first part and which is the second part. In practice, when we compute invariance for such objects we may need to try both orderings of the parts in the image.

Invariants for objects with repeated structures undergoing 3D-to-2D projective transformations were introduced in (Mundy and Zisserman, 1994). In addition, invariants for bilaterally symmetric objects and objects composed of planar repeated structure under both affine and projective transformations were presented in (Fawcett et al., 1994; Mitsumoto et al., 1992; Moses and Ullman, 1992; Poggio and Vetter, 1992; Sugimoto, 1994). These studies showed that the class of repeated structures induces invariance on the set of images. The basic intuition is that, when an image of an object with repeated structures is given, this image will in general contain two copies of the same structure. Thus, under the appropriate conditions, the shape of the repeated part can be recovered up to an affine (or projective) transformation by simply using a stereo algorithm (Faugeras, 1992; Koenderink and van Doorn, 1991). Nonetheless, these studies do not determine if

in general the identity of the objects can be determined from single images. Two objects with repeated structures may have exactly the same parts, but these parts may differ in their relative location or size across the objects. These studies determine the shape of the parts, but leave the relative position and size of the parts unknown. Below we show that if we permit the parts to be related by an arbitrary transformation then the class will not be identifiable (and thus objects that have the same parts will be confused). This implies that the relationships between the parts cannot be uniquely recovered from single images, which is equivalent to saying that the problem of calibrating an affine camera from two images is inherently ambiguous. If however we consider a class of objects with repeated structures in which the transformation relating the parts is the same across objects (such as in the case of bilaterally symmetric objects) then for most choices of transformation the class is identifiable from almost all views.

It is straightforward at this point to show that the class of objects with repeated structures whose parts are related by an arbitrary affine transformation is not identifiable from any view. This can be seen by applying Proposition 3(1) with  $Q_1 = Q_2$ , which tells us that a subset of this class, the objects that have the same parts, cannot be identified from all their views. Consequently, no invariants can distinguish between all the objects in this class.

Next we consider the case that the same affine transformation relates the two parts in every object. We show that for most choices of an affine transformation the class determined by this transformation is identifiable from almost all views. The degenerate views in this case lie along at most three great circles on the viewing sphere. We list these views in Proposition 4.

**Proposition 4.** *Given a  $3 \times 3$  non-singular matrix  $B$  and a vector  $\mathbf{s} \in \mathbb{R}^3$ , the class  $\mathcal{C}_{B,s}$  is identifiable from all views  $\hat{\mathbf{v}} \in \mathbb{R}^3$  unless  $\hat{\mathbf{v}}$  is an eigenvector of  $B$  or  $\hat{\mathbf{v}}$  is located on at most three great circles on the viewing sphere which depend on  $B$ . (The additional degenerate views are listed in Table 1.)*

The proof is given in Appendix A. The additional degenerate views listed in Table 1 depend on the matrix  $B$ , and, in particular, on the number of different eigenvectors and different eigenvalues of  $B$ . The degenerate views include all the eigenvectors of  $B$ . For convenience, we list the degenerate views according to the Jordan form of the matrix  $B$ ,  $\bar{B} = A^{-1}BA$ . The views are given in terms of  $\mathbf{v}'$ , where  $\mathbf{v}' = A^{-1}\mathbf{v}$ . Note

Table 1. List of degenerate views for the class  $\mathcal{C}_{B,s}$ , in addition to the eigenvectors of  $B$ , listed according to the number of different eigenvalues and different eigenvectors of  $B$ . The degenerate views are given in terms of  $\mathbf{v}'$ , where  $\mathbf{v}'$  is an eigenvector of the Jordan form of  $B$ .

Case	No. of eigenvectors	No. of eigenvalues	Constraints on $\mathbf{v}'$
(a)	3	3	$v'_i = 0$ for some $1 \leq i \leq 3$ .
(b)	3	2	$v'_1 = 0$ .
(c)	3	1	no identification from any view.
(d)	2	2	$v'_3 = 0$ or $v'_2 = 0$ .
(e)	2	1	only eigenvectors.
(f)	1	1	$v'_3 = 0$ or $v'_1 = v'_2 = 0$ .

that given a vector  $\mathbf{v}'$  it is straightforward to recover the corresponding view  $\hat{\mathbf{v}} = A\mathbf{v}'$ . Furthermore,  $\hat{\mathbf{v}}$  is an eigenvector of  $B$  if and only if  $\mathbf{v}'$  is an eigenvector of the Jordan form  $\bar{B}$ . Note also that in certain cases the actual list of degenerate views is smaller than what is specified in Table 1 since the list of vectors  $\mathbf{v}'$  includes vectors that correspond to complex views  $\hat{\mathbf{v}}$ . Evidently, only real view vectors are geometrically feasible.

As an example consider the class of bilaterally symmetric objects. In this case the matrix  $B$  representing reflection about a plane is given by

$$B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

and  $\mathbf{s} = 0$ .  $B$  has three eigenvectors and two different eigenvalues, corresponding to the second row of the table. Since  $B$  is already in its Jordan form then  $\mathbf{v}' = \hat{\mathbf{v}}$ . Accordingly, bilaterally symmetric objects are identifiable from all views except for those which coincide with the symmetry plan,  $\hat{\mathbf{v}} = \alpha[1, 0, 0] + \beta[0, 1, 0]$ , and the direction perpendicular to the symmetry plan,  $\hat{\mathbf{v}} = [0, 0, 1]$ .

The following example demonstrates that bilaterally symmetric objects are identifiable from single images, whereas objects with repeated structures are not identifiable when the kind of repetition is unknown (that is, with arbitrary  $B$  and  $\mathbf{s}$ ). To show this we consider an image of a bilaterally symmetric object (a chair). We show that the 3D structure of the chair can be recovered correctly from this image if we restrict the class to bilaterally symmetric objects, whereas if we allow arbitrary affine repetitions the image gives rise to ambiguous reconstruction. Figure 4 shows an image of

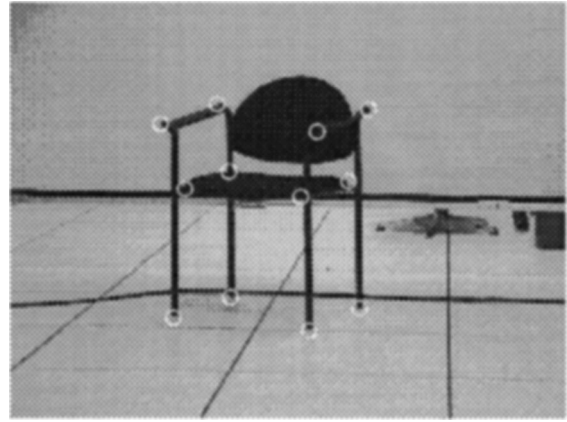


Figure 4. An image of a chair. The circles mark the positions of feature points in the image.

a chair, and feature points extracted manually. Using the assumption that the chair is bilaterally symmetric we have recovered its affine shape. To demonstrate the accuracy of this reconstruction we aligned the obtained shape with two images of the same chair taken from two different viewing positions (Fig. 5). As can be seen, the transformed feature points of the recovered shape align well with the corresponding image features, indicating that our affine reconstruction is accurate.

Next, we demonstrate that there exist different repetitions (that is, different choices of  $B$  and  $\mathbf{s}$ ) that are consistent with Fig. 4, but lead to quite different reconstructions. This implies that unless  $B$  and  $\mathbf{s}$  are determined the object is not identifiable from a single image. To show this we have derived the shape of the chair assuming a false repetition (by replacing  $B$  with a rotation matrix of  $60^\circ$  about the vertical axis of the chair and keeping  $\mathbf{s} = 0$ ). This reconstruction is consistent with Fig. 4, but is inconsistent with the actual 3D shape of the chair. To demonstrate this we have aligned the obtained shape with the same two images of the chair by minimizing the least-squares error in the position of the feature points (Fig. 5). As can be seen the transformed feature points poorly align with the corresponding image features, indicating that the reconstructed shape in this case is inconsistent with the 3D shape of the chair.

#### 4.3. Combinations of Sub-structures

Next we consider classes of objects in which one part can be expressed as a linear combination of the other

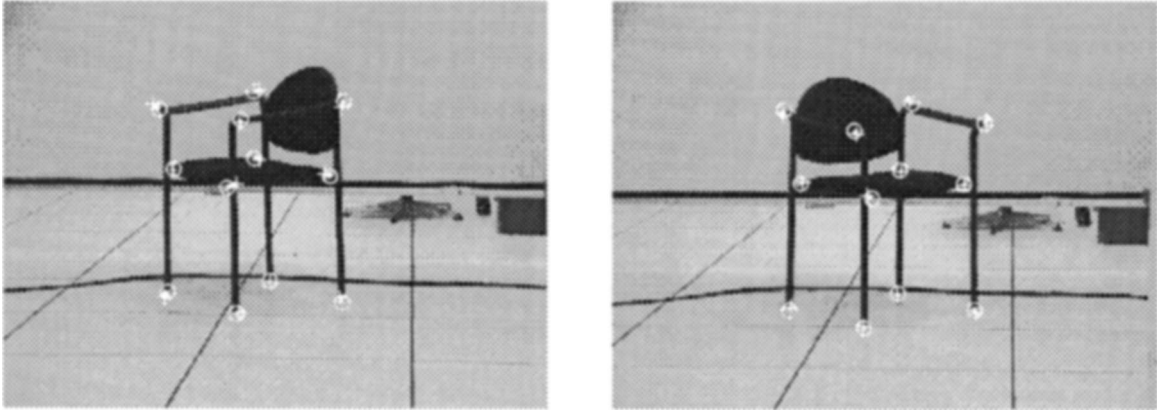


Figure 5. Aligning the affine shape of the chair derived assuming it is bilaterally symmetric to two other images of the same chair. The plus signs mark the position of the transformed model points and the circles mark the position of the corresponding image points.

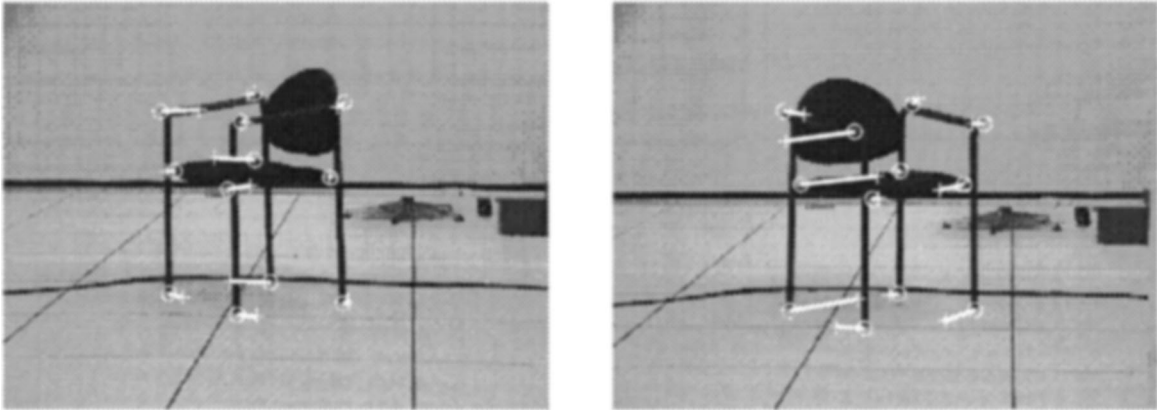


Figure 6. Aligning the affine shape of the chair derived assuming a false repetition to two other images of the same chair. The plus signs mark the position of the transformed model points and the circles mark the position of the corresponding image points. The lines connecting pairs of corresponding points illustrate the displacements obtained.

parts. Below we consider only classes of objects with three or more sub-structures. When the number of parts is two we obtain the case of repeated structures, which was discussed in Section 4.2. Intuitively, objects that belong to these classes contain one part whose shape is the average of the rest of the parts after these parts are aligned (that is, rotated and translated to a common coordinate frame). While it is difficult to find such classes in practice, we are interested in them since it is possible to build such classes such that their objects will have the same degrees of freedom as planar objects do ( $2n$ , where  $n$  is the number of points on the object). Yet, as is shown below, unlike the class of planar objects these classes are not identifiable from any view.

In a class that contains a combination of sub-structures an object is divided into  $m \geq 2$  parts. The

location of every point on the  $m$ 'th parts can be expressed as a linear function of the  $m - 1$  corresponding points in the other parts, where the same linear function is applied to all the points. An object from the class  $\mathcal{C}_{B_1, \dots, B_{m-1}, s}$  has the following form:

$$O = [P_1, P_2, \dots, P_m],$$

where  $P_i$  are  $3 \times n$  matrices defining the shape of the  $i$ 'th part ( $1 \leq i < m$ ),

$$P_m = B_1 P_1 + B_2 P_2 + \dots + B_{m-1} P_{m-1} + s \mathbf{1}^T,$$

$B_i$  are  $3 \times 3$  non-singular matrices, and  $s \in \mathbb{R}^3$ . Notice that the case that  $P_m$  is a linear combination of the rest of the parts in the ordinary sense (that is,

$P_m = \sum_{i=1}^{m-1} \beta_i P_i$ ) is included as a special case of this definition obtained by setting  $B_i = \beta_i I$  and  $\mathbf{s} = 0$ . The results below will hold for this special case as well.

In Proposition 5 below we show that for  $m > 2$ , identification is impossible from all views even when all the matrices  $B_i$  and  $\mathbf{s}$  are the same for all objects in the class.

**Proposition 5.** *Given  $3 \times 3$  non-singular matrices  $B_i$  where  $1 \leq i \leq m-1$ ,  $m \geq 3$ , and given  $\mathbf{s} \in \mathbb{R}^3$ , the class  $\mathcal{C}_{B_1, \dots, B_{m-1}, \mathbf{s}}$  is not identifiable from any view.*

**Proof:** Here we prove the proposition for  $m > 3$ . The proof for  $m = 3$  is given in Appendix B. To show this we need to show that for every choice of object  $O$  and view  $\hat{\mathbf{v}}$  there exists an affine different object  $O'$  and a view  $\hat{\mathbf{u}}$  so that  $O$  and  $O'$  share the view  $\hat{\mathbf{v}}$  and  $\hat{\mathbf{u}}$  respectively. According to Proposition 2(d) it is sufficient to show that there exists an object  $O'$  in the class such that  $O = A O' + \mathbf{t} \mathbf{1}^T + \hat{\mathbf{v}} \eta^T$  for  $\eta \notin \text{span}(\text{row}(O'), \mathbf{1})$  and for some non-singular  $A$  and  $\mathbf{t}$ . In particular, it will be sufficient to show this for  $A = I$  and  $\mathbf{t} = 0$ , so that  $O = O' + \hat{\mathbf{v}} \eta^T$  for  $\eta \perp \text{span}(\text{row}(O'), \mathbf{1})$ .

Let  $O = [P_1, P_2, \dots, P_m]$ , with  $P_m = \sum_{i=1}^{m-1} B_i P_i + \mathbf{s} \mathbf{1}^T$ . We set  $O' = [P'_1, P'_2, \dots, P'_m]$  such that

$$P'_i = P_i + \hat{\mathbf{v}} \eta_i^T \quad \text{for } (1 \leq i \leq m), \quad (15)$$

where  $\eta^T = [\eta_1^T, \eta_2^T, \dots, \eta_m^T]$ . To belong to the class the object  $O'$  must satisfy the following equation:

$$P'_m = \sum_{i=1}^{m-1} B_i P'_i + \mathbf{s} \mathbf{1}^T. \quad (16)$$

Combining (15) (the case  $i = m$ ) and (16) we obtain that

$$P'_m = P_m + \hat{\mathbf{v}} \eta_m^T = \sum_{i=1}^{m-1} B_i P'_i + \mathbf{s} \mathbf{1}^T. \quad (17)$$

Plugging in the first  $m-1$  equations in (15) into (17) and rearranging we obtain

$$\sum_{i=1}^{m-1} B_i \hat{\mathbf{v}} \eta_i^T - \hat{\mathbf{v}} \eta_m^T = 0. \quad (18)$$

It is left to show that for every view  $\hat{\mathbf{v}}$  there exists a non-trivial vector  $\eta^T = [\eta_1^T, \dots, \eta_m^T]$  that solves the above equations subject to the constraint  $O' \eta = 0$  and

$\mathbf{1}^T \eta = 0$ . The last two equations and Eq. (18) contain  $3n+4$  homogeneous equations in  $mn$  unknowns. Since  $m \geq 4$ ,  $mn > 3n+4$  for  $n > 3$ , and so for any choice of  $\hat{\mathbf{v}}$  a non-trivial solution will exist. It follows that for every  $\hat{\mathbf{v}}$  there exists an object given by  $O' = O - \hat{\mathbf{v}} \eta^T$  that shares the view  $\hat{\mathbf{v}}$  with  $O$ . Hence, the class is not identifiable.  $\square$

We next discuss the counting argument for the case  $m = 3$ . In this case, an object has the form  $O = [P_1, P_2, P_3]$  where  $P_3 = B P_1 + C P_2 + \mathbf{s} \mathbf{1}^T$ ,  $B$  and  $C$  are  $3 \times 3$  non-singular matrices, and  $\mathbf{s} \in \mathbb{R}^3$ . Below we assume in addition that  $\text{rank}(J(P_1, P_2)) = 7$ .<sup>1</sup> If we rely only on counting arguments we may be misled to believe that the class  $\mathcal{C}_{B,C,\mathbf{s}}$  is identifiable. Consider for example the counting arguments given in (Werman and Shashua, 1995). Given an object  $O \in \mathcal{C}_{B,C,\mathbf{s}}$ , suppose  $O$  includes  $n$  points. An image of  $O$  gives  $2n$  measurements. Together with  $n$  class constraints (the linear relations between the first two parts and the third part) they give rise to  $3n$  equations. The number of variables in these equations are as follows. The shape of  $O$  is defined by  $3n$  variables. The projection parameters (the parameters of a 3D-to-2D affine transformation) are eight. A 3D affine reference frame can be obtained by picking the position of four points, hence such a frame involves setting 12 parameters. Therefore, the total number of variables is  $3n + 8 - 12 = 3n - 4$ . According to a counting argument the number of equations obtained,  $3n$ , is sufficiently large to determine the values of the  $3n - 4$  unknowns, and so in theory using elimination we should be able to recover the shape of the object from the equations. However, the underlying assumption in counting arguments is that the counted equations are all independent and consistent. In classes which contain combinations of sub-structures this assumption is violated, and so the identity of the objects cannot be recovered from single images, as was shown above in Proposition 5.

#### 4.4. Combinations of Prototypes

Finally, we consider classes of objects that can be expressed as linear combinations of some prototype objects. That is, let  $\{O_1, O_2, \dots, O_k\}$  be a set of prototype objects, all with the same number of points, and assume further that the row spaces of all  $O_i$  ( $1 \leq i \leq k$ ) are linearly independent (that is, the union of the row spaces is of rank  $3k$ ), and that  $\mathbf{1}$  does not belong to these spaces. The objects in the class can be described

by combining the  $k$  prototypes, that is,  $O \in C$  if and only if it can be written as

$$O = \sum_{i=1}^k \alpha_i O_i, \quad \text{for } \alpha_i \in \mathbb{R}.$$

Poggio and Vetter (1992) (see also (Jones and Poggio, 1995)) already showed that this class is identifiable. Our tools provide a short and elegant proof for this case and establish that identification is possible from *all* views.

**Proposition 6.** *The objects in the class  $C$  are identifiable from all their views.*

**Proof:** To show this we need to show that, given an object  $O$  and a view  $\hat{\mathbf{v}}$ , any object  $O'$  that shares the view  $\hat{\mathbf{v}}$  of  $O$  must be affine equivalent to  $O$ . Let  $O = \sum_{i=1}^k \alpha_i O_i$  and let  $O' = \sum_{i=1}^k \beta_i O_i$  share the view  $\hat{\mathbf{v}}$  of  $O$ . It follows that,

$$A_{\hat{\mathbf{v}}}\left(\sum_{i=1}^k \alpha_i O_i\right) = A_{\hat{\mathbf{u}}}\left(\sum_{i=1}^k \beta_i O_i\right) + \mathbf{t}\mathbf{1}^T. \quad (19)$$

This can be written as

$$\sum_{i=1}^k (\alpha_i A_{\hat{\mathbf{v}}} - \beta_i A_{\hat{\mathbf{u}}}) O_i - \mathbf{t}\mathbf{1}^T = 0. \quad (20)$$

Since the row spaces of all the prototype objects are linearly independent, and since  $\mathbf{1}$  does not belong to these spaces, we obtain that

$$\alpha_i A_{\hat{\mathbf{v}}} - \beta_i A_{\hat{\mathbf{u}}} = 0 \quad (21)$$

and that  $\mathbf{t} = 0$ . Thus,  $A_{\hat{\mathbf{v}}}$  and  $A_{\hat{\mathbf{u}}}$  are related by the scale factor  $c_i = \beta_i/\alpha_i$  for all  $1 \leq i \leq k$ , and obviously all these scale factors  $c_i$  ( $1 \leq i \leq k$ ) must be identical. (Note also that neither  $\alpha_i$  and  $\beta_i$  can be zero because the rank of both  $A_{\hat{\mathbf{v}}}$  and  $A_{\hat{\mathbf{u}}}$  is 2.) This implies that  $\hat{\mathbf{v}} = \hat{\mathbf{u}}$  and  $O$  and  $O'$  are affine equivalent (since  $O' = cO$ ).  $\square$

The reason this class is of interest is the following. Suppose that the set of prototype objects is composed of a few similar objects that belong to a single perceptual category, say, two chairs, and suppose that “reasonable” correspondences between feature points on the objects can be assigned (by applying form and function considerations). Then, if we construct new objects by

taking averages of the prototype objects, these new objects would tend to look similar to the prototype objects and assume the same perceptual category. If indeed for this kind of classes identification is possible then one may be able to distinguish between different exemplars of such a category even if the specific exemplar is being seen for the first time.

## 5. Summary and Discussion

A fundamental question regarding the invariance approach is whether it can be applied to a wide range of classes. To answer this question it is essential to study the set of classes for which invariance exists. In this paper we investigated the invariant representations that discriminate between all the objects of a given class. We addressed the problem of determining, given a class of objects, the set of images from which the objects can be identified. Our approach is based on exploring the set of ambiguous images. We developed a number of algebraic tests to determine the ambiguous images under affine projection, and applied these tests to a number of classes of objects.

We now consider a number of assumptions we have made and discuss how they should be relaxed in future work.

*Projection Model:* Our tests were developed for objects which may undergo affine projection. We intend in the future to develop similar tests for other, more realistic projection models. We would like to note, however, that some of our results (e.g., Proposition 3) apply also to rigid objects undergoing weak-perspective projection. Similarly, some of the results in Proposition 3 and 4 can be extended quite easily to the case of objects undergoing perspective projection.

*Constructive Tests:* Although our tests can determine which classes are identifiable and from what views, the tests at their present form are not constructive. That is, the tests cannot be used to derive the invariants for the objects. Consequently, our approach may serve only as a first step in deriving invariance for new classes of objects. In particular, it can be used to avoid seeking invariance for classes for which invariance does not exist. In addition, for identifiable classes our method can be used to determine the set of views from which objects are identifiable and exclude the ambiguous views.

*Dealing with Noise:* In our tests we did not take into account the effect of noise on identifiable classes. Obviously, if we allow for noise, more images may become ambiguous. One possible way to detect sensitivity to

noise in classes of objects is by looking at the singular values of the matrix  $\bar{J}(O_1, O_2)$ , which was introduced in Section 3.2. However, those singular values depend on the specific representation chosen for the objects  $O_1$  and  $O_2$ , so we must somehow “normalize” the objects first. We can do this by replacing the rows of  $O_1$  and  $O_2$  by orthonormal bases (that are also normal to the vector  $\mathbf{1}$ ). Denote the new representation by  $\tilde{O}_1$  and  $\tilde{O}_2$ , it can be shown that the singular values of  $\bar{J}(\tilde{O}_1, \tilde{O}_2)$  are related to the angles between the two row spaces of  $O_1$  and  $O_2$ . Specifically, the rank of  $\bar{J}(O_1, O_2)$  is exactly 7 minus the number of zero angles between these spaces. In the presence of noise objects in the class may be confused when these angles are small. Thus, by verifying that the angles between the row spaces exceed a certain threshold one can guarantee that the objects will not be confused due to noise.

*The objects:* In our tests we assumed that the objects are given as ordered point sets. In particular we regarded two objects that contain the same set of points ordered differently as two different objects. Several of our results are not affected by this assumption. For example, the class of objects with identical parts and the class of objects with repeated structures are closed under permutation of the objects’ points. (One only has to find correspondence between the different parts of the same object.) Furthermore, classes which were shown to be non-identifiable remain non-identifiable even if we relax this requirement. We intend in the future to extend our tests to other objects that consists of non-ordered points. In addition, it is clearly of interest to develop tests for contour images, and gray-level images.

*Invariance:* In this paper we concentrated on analyzing whether given classes are identifiable. For classes that are not identifiable it may still be possible to extract invariants from single images. These invariants will not suffice to discriminate between all the objects of the class, but will distinguish only between subsets of the objects. Developing tests for such classes is left for future research.

*Classifying the objects:* The first step of applying class-based invariance to images involves classifying the object in the image. Unfortunately, it is impossible to both classify the object and recover its specific identity using invariance, since this will contradict the result that there exists no view invariance that can discriminate between all 3D objects. In practical systems it may be the case that many of the objects belong to a small number of classes, in which case one may enumerate all these classes, or find some properties which

distinguish these classes from one another. This problem too is beyond the scope of this paper.

## Appendix A: Repeated Structures

In this appendix we prove Proposition 4.

**Proof:** We will show that the class  $\mathcal{C}_{B,s}$  is identifiable from all views  $\hat{\mathbf{v}} \in \mathbb{R}^3$  except for the degenerate views which are the views that are eigenvalues of  $B$  and the views listed in Table 1. To show this we will show that given an object  $O \in \mathcal{C}_{B,s}$ , and a view  $\hat{\mathbf{v}}$  there exists an affine different object  $O' \in \mathcal{C}_{B,s}$  that shares the view  $\hat{\mathbf{v}}$  of  $O$  if and only if  $\hat{\mathbf{v}}$  is a degenerate view.

Let  $O$  and  $O'$  be two affine different objects in  $\mathcal{C}_{B,s}$  given by

$$O = [P_1, P_2] \in \mathcal{C}_{B,s}, \quad (\text{A1})$$

where  $P_2 = BP_1 + \mathbf{s}\mathbf{1}^T$ , and

$$O' = [P'_1, P'_2] \in \mathcal{C}_{B,s}, \quad (\text{A2})$$

where  $P'_2 = BP'_1 + \mathbf{s}\mathbf{1}^T$ .

Suppose that  $O'$  shares only the view  $\hat{\mathbf{v}}$  with  $O$ . It follows that there exist two  $2 \times 3$  non-singular matrices  $A_{\hat{\mathbf{v}}}$  and  $A_{\hat{\mathbf{q}}}$  and a vectors  $\mathbf{t}_1 \in \mathbb{R}^2$  such that

$$A_{\hat{\mathbf{v}}}O = A_{\hat{\mathbf{q}}}O' + \mathbf{t}_1\mathbf{1}^T. \quad (\text{A3})$$

In particular, since  $O$  and  $O'$  share a view, also their parts share the same view. That is,

$$A_{\hat{\mathbf{v}}}P_1 = A_{\hat{\mathbf{q}}}P'_1 + \mathbf{t}_1\mathbf{1}^T \quad (\text{A4})$$

$$A_{\hat{\mathbf{v}}}(BP_1 + \mathbf{s}\mathbf{1}^T) = A_{\hat{\mathbf{q}}}(BP'_1 + \mathbf{s}\mathbf{1}^T) + \mathbf{t}_1\mathbf{1}^T. \quad (\text{A5})$$

Note that although we assumed that  $O$  and  $O'$  are affine different, it might still be the case that their parts are affine equivalent. Therefore, the rank of  $\bar{J}(P_1, P'_1) \leq 5$ . In addition, we assumed that both parts are non-planar, that is, both ranks of  $\text{span}(\text{row}(P_1), \mathbf{1})$  and  $\text{span}(\text{row}(P'_1), \mathbf{1})$  are 4. This implies that  $\text{rank}(\bar{J}(P_1, P'_1)) \geq 4$ . Below we analyze each of the two cases,  $\text{rank}(\bar{J}(P_1, P'_1)) = 5$  and  $\text{rank}(\bar{J}(P_1, P'_1)) = 4$ , separately.

*Case I.* Suppose that  $\text{rank}(\bar{J}(P_1, P'_1)) = 5$ . According to Proposition 2(a), the two parts,  $P_1$  and  $P'_1$  have

exactly one view in common,  $\hat{\mathbf{v}}$ . From (A4) it follows that the projection matrix  $A_{\hat{\mathbf{v}}}$  belongs to the view  $\hat{\mathbf{v}}$ , and similarly from (A5) it follows that the projection matrix  $A_{\hat{\mathbf{v}}}B$  belongs to the view  $\hat{\mathbf{v}}$ . Therefore,  $A_{\hat{\mathbf{v}}}\hat{\mathbf{v}} = 0$  and also  $A_{\hat{\mathbf{v}}}B\hat{\mathbf{v}} = 0$ . By definition of  $A_{\hat{\mathbf{v}}}$ , the only vector  $\mathbf{w}$  that satisfies  $A_{\hat{\mathbf{v}}}\mathbf{w} = 0$  is of the form  $\mathbf{w} = \lambda\hat{\mathbf{v}}$ . In particular it follows that  $B\hat{\mathbf{v}} = \lambda\hat{\mathbf{v}}$ . Consequently,  $O'$  shares the view  $\hat{\mathbf{v}}$  of  $O$  when  $\text{rank}(\bar{J}(P_1, P'_1)) = 5$  if and only if  $\hat{\mathbf{v}}$  is an eigenvector of  $B$ , which implies that  $\hat{\mathbf{v}}$  is a degenerate view.

*Case II.* Suppose that  $\text{rank}(\bar{J}(P_1, P'_1)) = 4$ . According to Proposition 1, the parts  $P_1$  and  $P'_1$  are affine equivalent, and so there exists a  $3 \times 3$  non-singular matrix  $D$  and a vector  $\mathbf{t} \in \mathbb{R}^3$  such that

$$P_1 = DP'_1 + \mathbf{t}\mathbf{1}^T. \quad (\text{A6})$$

We will use the following two claims to prove that  $O'$  shares a single view  $\hat{\mathbf{v}}$  with the object  $O$  only if  $\hat{\mathbf{v}}$  is an eigenvector of  $B$  or  $\hat{\mathbf{v}}$  is listed in Table 1.

**Claim 1.**  *$O'$  shares the view  $\hat{\mathbf{v}}$  of  $O$  if and only if there exists a vector  $\mathbf{r} \in \mathbb{R}^3$  such that  $\mathbf{r} \neq 0$ , and the commutator matrix  $[B, D] = BD - DB$  satisfies*

$$[B, D] = \hat{\mathbf{v}}\mathbf{r}^T. \quad (\text{A7})$$

**Claim 2.** *For every  $3 \times 3$  non-singular matrices  $B$  and  $D$ , if*

$$[B, D] = \hat{\mathbf{v}}\mathbf{r}^T \neq 0$$

*for some  $\mathbf{r} \in \mathbb{R}^3$  then  $\hat{\mathbf{v}}$  is an eigenvector of  $B$  or  $\hat{\mathbf{v}}$  is listed in Table 1.*

These two claims imply that  $O'$  shares the view  $\hat{\mathbf{v}}$  of  $O$  when  $\text{rank}(\bar{J}(P_1, P'_1)) = 4$  only if  $\hat{\mathbf{v}}$  is an eigenvector of  $B$  or  $\hat{\mathbf{v}}$  is listed in Table 1.  $\square$

We now turn to proving these two claims.

**Proof of Claim 1:** If  $O'$  shares the view  $\hat{\mathbf{v}}$  of  $O$  then (A4) and (A5) must be satisfied. Since we also assume here that their parts are affine equivalent it follows that (A6) must be satisfied as well. Plugging (A6) into (A4) we obtain that

$$A_{\hat{\mathbf{v}}}(DP'_1 + \mathbf{t}\mathbf{1}^T) = A_{\hat{\mathbf{u}}}P'_1 + \mathbf{t}_1\mathbf{1}^T. \quad (\text{A8})$$

Rearranging, we obtain

$$(A_{\hat{\mathbf{v}}}D - A_{\hat{\mathbf{u}}})P'_1 + (A_{\hat{\mathbf{v}}}\mathbf{t} - \mathbf{t}_1)\mathbf{1}^T = 0. \quad (\text{A9})$$

Since we assume that  $P'_1$  is non-planar (the rank of  $\text{span}(\text{row}(P'_1), \mathbf{1})$  is 4) the coefficients of  $P'_1$  and  $\mathbf{1}$  in the last equation must vanish, namely,

$$\begin{aligned} A_{\hat{\mathbf{v}}}D - A_{\hat{\mathbf{u}}} &= 0 \\ A_{\hat{\mathbf{v}}}\mathbf{t} - \mathbf{t}_1 &= 0. \end{aligned} \quad (\text{A10})$$

Consider now Eq. (A5). Replacing  $A_{\hat{\mathbf{u}}}$  by  $A_{\hat{\mathbf{v}}}D$  and  $P_1$  by  $DP'_1 + \mathbf{t}\mathbf{1}^T$  we obtain

$$A_{\hat{\mathbf{v}}}B(DP'_1 + \mathbf{t}\mathbf{1}^T) = A_{\hat{\mathbf{v}}}DBP'_1 + \mathbf{t}_2\mathbf{1}^T, \quad (\text{A11})$$

where  $\mathbf{t}_2 = (A_{\hat{\mathbf{v}}}D - A_{\hat{\mathbf{v}}})\mathbf{s} + \mathbf{t}_1$ . Rearranging, we get

$$A_{\hat{\mathbf{v}}}(BD - DB)P'_1 + (A_{\hat{\mathbf{v}}}B\mathbf{t} - \mathbf{t}_2)\mathbf{1}^T = 0. \quad (\text{A12})$$

Again, since  $O'$  is non-planar the coefficients must vanish, namely,

$$A_{\hat{\mathbf{v}}}(BD - DB) = 0 \quad (\text{A13})$$

$$A_{\hat{\mathbf{v}}}B\mathbf{t} - \mathbf{t}_2 = 0. \quad (\text{A14})$$

It is immediate to see that there always exists  $\mathbf{t}$  that satisfies (A14) for any viewing directions  $\hat{\mathbf{v}}$  (since  $\mathbf{t}$  should satisfy a system of two independent linear equations in three unknowns). Since  $\text{rank}(A_{\hat{\mathbf{v}}}) = 2$  it follows that there are two cases for which (A13) can be satisfied. The first is if  $BD - DB = 0$ . In this case (A13) will vanish for all viewing directions  $\hat{\mathbf{v}}$ . In particular it follows that in this case  $O$  and  $O'$  will share all their views (that is,  $O$  and  $O'$  are affine equivalent). We are therefore left with the second case where  $BD - DB = \hat{\mathbf{v}}\mathbf{r}^T$  for some non zero vector  $\mathbf{r} \in \mathbb{R}^3$ .  $\square$

**Proof of Claim 2:** To prove this claim we will use the Jordan form of the matrix  $B$ . The Jordan form,  $\bar{B}$ , is obtained from  $B$  by a similarity transformation, that is,  $\bar{B} = A^{-1}BA$  for some  $3 \times 3$  non singular matrix  $A$ . If  $D$  and  $\mathbf{r}$  satisfy (A7) for a given matrix  $B$  and a vector  $\mathbf{v}$ , then  $D' = A^{-1}DA$  and  $\mathbf{r}'^T = \mathbf{r}^T A$  satisfy the same equation for  $\bar{B}$  and  $\mathbf{v}' = A^{-1}\mathbf{v}$ . Thus, rather than showing that for every two non-singular matrices  $B$  and  $D$  and a view  $\hat{\mathbf{v}}$  there exists a non-zero vector  $\mathbf{r}$  that satisfies (A7) only if  $\hat{\mathbf{v}}$  is degenerate, we may instead show that for every matrix  $\bar{B}$  in a Jordan form, a matrix  $D'$ , and a vector  $\hat{\mathbf{v}}'$  there exists a non-zero vector  $\mathbf{r}'$  only if  $\hat{\mathbf{v}}'$  is as specified in Table 1. Note however that the matrices  $\bar{B}$  and  $D'$  and the vectors  $\mathbf{r}'$  and  $\mathbf{v}'$  are defined over the complex field. Nevertheless, this will not affect our proof, since by proving the claim for



Table 2. The shape of the commutator  $[\bar{B}, F]$  as a function of the number of eigenvectors and eigenvalues of  $B$ .  $\bar{B}$  is the Jordan form of  $B$ ,  $\lambda_i$  ( $1 \leq i \leq 3$ ) are the eigenvalues of  $B$  and  $f_{ij}$  ( $1 \leq i, j \leq 3$ ) are the components of  $F$ .

Case	No. of eigenvectors	No. of eigenvalues	$\bar{B}$	$G = [\bar{B}, \bar{F}]$
(a)	3	3	$\begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$	$\begin{pmatrix} 0 & (\lambda_1 - \lambda_2)f_{12} & (\lambda_1 - \lambda_3)f_{13} \\ (\lambda_2 - \lambda_1)f_{21} & 0 & (\lambda_2 - \lambda_3)f_{23} \\ (\lambda_3 - \lambda_1)f_{31} & (\lambda_3 - \lambda_2)f_{32} & 0 \end{pmatrix}$
(b)	3	2	$\begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 & (\lambda_1 - \lambda_3)f_{13} \\ 0 & 0 & (\lambda_1 - \lambda_3)f_{23} \\ (\lambda_3 - \lambda_1)f_{31} & (\lambda_3 - \lambda_1)f_{32} & 0 \end{pmatrix}$
(c)	3	1	$\begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_1 \end{pmatrix}$	0
(d)	2	2	$\begin{pmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$	$\begin{pmatrix} f_{21} & -f_{11} + f_{22} & (\lambda_1 - \lambda_3)f_{13} + f_{23} \\ 0 & -f_{21} & (\lambda_1 - \lambda_3)f_{23} \\ (\lambda_3 - \lambda_1)f_{31} & -f_{31} + (\lambda_3 - \lambda_1)f_{32} & 0 \end{pmatrix}$
(e)	2	1	$\begin{pmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_1 \end{pmatrix}$	$\begin{pmatrix} f_{21} & -f_{11} + f_{22} & f_{23} \\ 0 & -f_{21} & 0 \\ 0 & -f_{31} & 0 \end{pmatrix}$
(f)	1	1	$\begin{pmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 1 \\ 0 & 0 & \lambda_1 \end{pmatrix}$	$\begin{pmatrix} f_{21} & -f_{11} + f_{22} & -f_{12} + f_{23} \\ f_{31} & -f_{21} + f_{32} & -f_{22} + f_{33} \\ 0 & -f_{31} & -f_{32} \end{pmatrix}$

every complex matrix  $D'$  we prove in particular that every real matrix  $D$  satisfies the claim.

To derive the six cases listed in Table 1, let us first list the forms that a commutator  $G = [\bar{B}, F]$  of a Jordan form matrix  $\bar{B}$  takes according to the number of independent eigenvectors and eigenvalues of  $\bar{B}$ . These forms are listed in Table 2. In our case,  $F = D'$ . Note that  $[\bar{B}, D'] = \mathbf{v}'\mathbf{r}'^T \neq 0$  if and only if  $G_{ij} = v_i r_j$  and  $G \neq 0$ . We next show that there exists a vector  $\mathbf{r}' \neq 0$  that satisfies  $G_{ij} = v_i r_j$  only if  $\mathbf{v}'$  is an eigenvector of  $\bar{B}$  or  $\mathbf{v}'$  is listed in Table 1.

- (a)  $B$  has 3 independent eigenvectors and 3 different eigenvalues. In this case  $v'_i r'_i = 0$ , for  $1 \leq i \leq 3$ . Since  $G \neq 0$  it follows that at least one of the matrix entry is non-zero. In particular it follows that for some  $i \neq j$ ,  $(\lambda_j - \lambda_i)D_{ij} \neq 0$ . That is  $v'_i r'_j \neq 0$ . Since  $v'_i r'_i = 0$ , it follows that  $v'_i = 0$ .

The three eigenvectors of  $\bar{B}$  are of the form  $(v'_1, 0, 0)^T$ ,  $(0, v'_2, 0)^T$ , and  $(0, 0, v'_3)^T$ , note that  $v'_i \neq 0$  for all  $1 \leq i \leq 3$  if and only if  $\hat{\mathbf{v}}$  cannot be expressed as a linear combination of any two of the eigenvectors of  $B$ . If the eigenvectors of  $B$  are all real we exclude by this from the viewing sphere exactly three great circles through all pairs

of eigenvectors. If some of the eigenvectors of  $B$  are not real we exclude from the viewing sphere even less views.

- (b)  $B$  has 3 independent eigenvectors and 2 different eigenvalues. In this case  $v'_1 = 0$ . We next show that if  $v'_1 \neq 0$  then  $v'_3 = 0$  (in this case,  $\mathbf{v}'$  is an eigenvector of  $\bar{B}$ ). Assume that both  $v'_1 \neq 0$  and  $v'_3 \neq 0$ . In this case  $\mathbf{r} = 0$  since  $v'_1 r'_2 = v'_1 r'_1 = v'_3 r'_3 = 0$ .
- (c)  $B$  has 3 independent eigenvectors and only 1 eigenvalue. This case has been handled in Section 4.2. In this case every  $\mathbf{v}'$  is an eigenvector.
- (d)  $B$  has 2 independent eigenvectors and 2 different eigenvalues. In this case  $v'_3 = 0$  or  $v'_2 = 0$ . We next show that if both  $v'_2 \neq 0$  and  $v'_3 \neq 0$  then  $\mathbf{r}' = 0$  contradicting our assumption. Since  $v'_3 r'_3 = 0$  then  $r'_3 = 0$ . Further, since  $v'_2 r'_1 = 0$  it follows that  $r'_1 = 0$ . This implies that the entire first column of  $G$  is zero. Therefore,  $v'_1 r'_1 = d'_{21} = 0$ . But now also  $v'_2 r'_2 = -d'_{21} = 0$ .
- (e)  $B$  has 2 independent eigenvectors and only 1 eigenvalue. In this case  $v'_2 = 0$ , and therefore,  $\mathbf{v}'$  is an eigenvector. We next show that if  $v'_2 \neq 0$  then  $\mathbf{r}' = 0$  contradicting our assumption. Since  $v'_2 r'_1 = v'_2 r'_3 = 0$  it follows that  $r'_1 = r'_3 = 0$ .

As a result the entire first and third columns of  $G$  vanish. It follows that  $d'_{21} = 0$  and therefore also  $v'_2 r'_2 = -d'_{21} = 0$ . This implies that  $r'_2 = 0$ .

- (f)  $B$  has only 1 independent eigenvectors and 1 eigenvalues. In this case  $v'_3 = 0$ . We next show that if  $v'_3 \neq 0$  then  $\mathbf{r}' = 0$  contradicting our assumption. Since  $v'_3 r'_1 = 0$  it follows that  $r'_1 = 0$ . The entire first column of  $G$  therefore vanishes, and so  $v'_2 r'_1 = d'_{31} = 0$ . This implies also that  $v'_3 r'_2 = -d'_{31} = 0$ , and since  $v'_3 \neq 0$  we obtain that  $r'_2 = 0$ . Thus, also the entire second column of  $G$  vanishes, implying in particular that  $v'_2 r'_2 = -d'_{21} + d'_{32} = 0$ . Since  $v'_1 r'_1 = d'_{21} = 0$ , we can conclude also that  $d'_{32} = 0$ . But  $v'_3 r'_3 = -d'_{32} = 0$ , and again, since  $v'_3 \neq 0$ ,  $r'_3 = 0$ .  $\square$

## Appendix B: Combinations of Sub-structures

**Proposition 7.** *Given  $3 \times 3$  non-singular matrices  $B$  and  $C$ , and given  $\mathbf{s} \in \mathbb{R}^3$ , the class  $\mathcal{C}_{B,C,\mathbf{s}}$  is not identifiable from any view.*

**Proof:** We can prove this by showing that for every object  $O \in \mathcal{C}_{B,C,\mathbf{s}}$  and every view  $\hat{\mathbf{v}} \in \mathbb{R}^3$ , there exists an object  $O' \in \mathcal{C}_{B,C,\mathbf{s}}$  that shares the view  $\hat{\mathbf{v}}$  with  $O$  (and no other view).

According to Proposition 2(d), if  $O'$  shares a single view  $\hat{\mathbf{v}}$  with  $O$  then there exist a  $3 \times 3$  non singular matrix  $D$ , a vector  $\mathbf{t} \in \mathbb{R}^3$ , and a vector  $\eta \in \mathbb{R}^{3n}$  such that

$$O = DO' + \hat{\mathbf{v}}\eta^T + \mathbf{t}\mathbf{1}^T \quad (\text{B1})$$

and

$$\eta \notin \text{span}(\text{row}(O'), \mathbf{1}). \quad (\text{B2})$$

Since  $O$  and  $O'$  are objects in the class  $\mathcal{C}_{B,C,\mathbf{s}}$ , they take the following form:

$$O = [P_1, P_2, P_3], \quad O' = [P'_1, P'_2, P'_3], \quad (\text{B3})$$

where  $P_3 = BP_1 + CP_2 + \mathbf{s}\mathbf{1}^T$ , and  $P'_3 = BP'_1 + CP'_2 + \mathbf{s}\mathbf{1}^T$ .

It is therefore sufficient to show that for every object  $O$  and a view  $\hat{\mathbf{v}}$  there exist a matrix  $D$  and vectors  $\mathbf{t} \in \mathbb{R}^3$  and  $\eta \in \mathbb{R}^{3n}$  that satisfy Eqs. (B1)–(B3).

Plugging in the constraints specified in (B3) into (B1) results in a set of equations that is specified in the following claim:  $\square$

**Claim 3.** *For every  $\hat{\mathbf{v}}$ , there exist a matrix  $D$  and vectors  $\mathbf{t} \in \mathbb{R}^3$  and  $\eta \in \mathbb{R}^{3n}$  that satisfy Eqs. (B1) and (B3) if and only if they satisfy the following equation:*

$$B\hat{\mathbf{v}}\eta_1^T + C\hat{\mathbf{v}}\eta_2^T - \hat{\mathbf{v}}\eta_3^T = [D, B]P'_1 + [D, C]P'_2 + \mathbf{t}'\mathbf{1}^T \quad (\text{B4})$$

where  $[D, B] = DB - BD$ ,  $[D, C] = DC - CD$  and,  $\mathbf{t}' = (I - B - C)\mathbf{t} + (D - I)\mathbf{s}$ .

It follows that to show that there exists an object  $O'$  that shares the view  $\hat{\mathbf{v}}$  with  $O$  it suffices to show that there exist a non-singular matrix  $D$  and vectors  $\mathbf{t} \in \mathbb{R}^2$  and  $\eta \in \mathbb{R}^{3n}$  that satisfy (B4) subject to the constraint that  $\eta \notin \text{span}(\text{row}(O'), \mathbf{1})$ .

**Proof of Claim 3:** (B1) together with (B3) imply that

$$\begin{aligned} P_1 &= DP'_1 + \hat{\mathbf{v}}\eta_1^T + \mathbf{t}\mathbf{1}^T \\ P_2 &= DP'_2 + \hat{\mathbf{v}}\eta_2^T + \mathbf{t}\mathbf{1}^T \\ P_3 &= DP'_3 + \hat{\mathbf{v}}\eta_3^T + \mathbf{t}\mathbf{1}^T, \end{aligned} \quad (\text{B5})$$

where  $\eta^T = [\eta_1, \eta_2, \eta_3]^T$ . Using (B3), the third of these equations can be written in terms of the first two parts of the objects as follows:

$$\begin{aligned} BP_1 + CP_2 + \mathbf{s}\mathbf{1}^T \\ = D(BP'_1 + CP'_2 + \mathbf{s}\mathbf{1}^T) + \hat{\mathbf{v}}\eta_3^T + \mathbf{t}\mathbf{1}^T. \end{aligned} \quad (\text{B6})$$

Replacing  $P_1$  and  $P_2$  by the right hand side of the first two equations in (B5) we obtain

$$\begin{aligned} B(DP'_1 + \hat{\mathbf{v}}\eta_1^T + \mathbf{t}\mathbf{1}^T) + C(DP'_2 + \hat{\mathbf{v}}\eta_2^T + \mathbf{t}\mathbf{1}^T) \\ + \mathbf{s}\mathbf{1}^T = D(BP'_1 + CP'_2 + \mathbf{s}\mathbf{1}^T) + \hat{\mathbf{v}}\eta_3^T + \mathbf{t}\mathbf{1}^T. \end{aligned} \quad (\text{B7})$$

Denote  $\mathbf{t}' = (I - B - C)\mathbf{t} + (D - I)\mathbf{s}$ , by rearranging we get

$$\begin{aligned} B\hat{\mathbf{v}}\eta_1^T + C\hat{\mathbf{v}}\eta_2^T - \hat{\mathbf{v}}\eta_3^T \\ = (DB - BD)P'_1 + (DC - CD)P'_2 + \mathbf{t}'\mathbf{1}^T \end{aligned} \quad (\text{B8})$$

which is identical to (B4).

Note that for a fixed matrix  $D$  and a vector  $\mathbf{t}$  Eq. (B4) contains  $3n$  linear equations in  $3n$  unknowns, the  $3n$  components of  $\eta$ . In the rest of the proof we will show

how given an object  $O$  and a vector  $\hat{\mathbf{v}}$  it is possible to select  $D$  and  $\mathbf{t}$  so that a solution to (B4) that satisfies  $\eta \notin \text{span}(\text{row}(O'), \mathbf{1})$  will exist.  $\square$

We will consider two cases according to the rank of the linear system in (B4). To do so, we first show that the rank of the linear equations in (B4) is either equal to  $3n$ , or it is smaller or equal to  $2n$ .

**Claim 4.** *Let*

$$W = [B\hat{\mathbf{v}}, C\hat{\mathbf{v}}, -\hat{\mathbf{v}}].$$

*Given  $D$  and a vector  $\mathbf{t}$ , the rank of Eq. (B4) is  $kn$ , where  $k$  is the rank of  $W$ .*

**Proof of Claim 4:** Let  $\eta_{1i}$ ,  $\eta_{2i}$  and  $\eta_{3i}$  be the  $i$ 'th components of the three vectors  $\eta_1$ ,  $\eta_2$  and  $\eta_3$ , respectively, and let  $\mathbf{p}'_{1i}$  and  $\mathbf{p}'_{2i}$  be the  $i$ 'th points of  $P'_1$  and  $P'_2$ , respectively. (B4) can be written as  $n$  sets of the following equations ( $1 \leq i \leq n$ ):

$$\begin{aligned} [B\hat{\mathbf{v}}, C\hat{\mathbf{v}}, -\hat{\mathbf{v}}] \begin{pmatrix} \eta_{1i} \\ \eta_{2i} \\ \eta_{3i} \end{pmatrix} \\ = (DB - BD)\mathbf{p}'_{1i} + (DC - CD)\mathbf{p}'_{2i} + \mathbf{t}'. \end{aligned} \quad (\text{B9})$$

It can be readily verified that the rank of the system given in (B4) is  $kn$ , where  $k$  is the rank of  $W$ .

Below we consider two cases according to the rank of  $W$  (or the rank of (B4)).

*Case I.* Suppose that  $W$  is singular, that is,  $\text{rank}(W) \leq 2$ . In this case the rank of (B4) is at most  $2n$ . We select  $D = I$  and  $\mathbf{t} = 0$ . This implies that  $[D, B] = [D, C] = 0$ , and  $\mathbf{t}' = 0$ . (B4) now simplifies to a set of homogeneous linear equations:

$$B\hat{\mathbf{v}}\eta_1^T + C\hat{\mathbf{v}}\eta_2^T - \hat{\mathbf{v}}\eta_3^T = 0. \quad (\text{B10})$$

To satisfy (B2) we will choose  $\eta$  that is perpendicular to the union of the row space of  $O'$  and  $\{\mathbf{1}\}$  (and so, in particular, it will not belong to this space). This requirement results in four more homogeneous equations. Thus, we obtain  $3n + 4$  homogeneous equations of rank  $2n + 4$  or less in  $3n$  unknowns. Consequently, there will always exist a non-trivial solution for  $\eta$  that will satisfy both Eqs. (B4) and (B2). It follows that

reconstruction is impossible from any direction  $\hat{\mathbf{v}}$  for which  $\text{rank}(W) < 3$ .

*Case II.* Suppose that  $\text{rank}(W) = 3$ . In this case for every choice of  $D$  and  $\mathbf{t}$  (B4) is a system of  $3n$  linear equations of rank  $3n$  with  $3n$  unknowns, the components of  $\eta$ . This system has a non-trivial solution if and only if it is non-homogeneous. Furthermore, in this case the solution is unique. We therefore need to find in this case  $D$  and  $\mathbf{t}$  for which (B4) is non-homogeneous and such that the solution to this equation will also satisfy  $\eta \notin \text{span}(\text{row}(O'), \mathbf{1})$ . We show this by using the following two claims:

**Claim 5.** *Let  $\text{rank}(W) = 3$  and denote  $E = D + \hat{\mathbf{v}}\mathbf{r}^T$ . If there exists a non-singular matrix  $D$  such that for every vector  $\mathbf{r}$  the commutator  $[B, E] \neq 0$  then*

1. (B4) is non-homogeneous.
2. The solution,  $\eta$ , to (B4) satisfies (B2).

**Claim 6.** *For any view  $\hat{\mathbf{v}}$  for which  $\text{rank}(W) = 3$  there exists a matrix  $D$  that satisfies the conditions of claim 5.*

The same arguments can be made for the matrix  $C$ .  $\square$

**Proof of Claim 5:**

1. We first show that when  $[B, E] \neq 0$  for every vector  $\mathbf{r}$  then (B4) is necessarily non-homogeneous. (B4) is non-homogeneous when

$$[D, B]P'_1 + [D, C]P'_2 + \mathbf{t}'\mathbf{1}^T \neq 0. \quad (\text{B11})$$

Since  $\text{rank}(\bar{J}(P'_1, P'_2)) = 7$  this is true if (and only if) either one of the following conditions holds:

$$[D, B] \neq 0, \quad [D, C] \neq 0, \quad \text{or} \quad \mathbf{t}' \neq 0.$$

Now  $[D, B] \neq 0$ , since for  $\mathbf{r} = 0$   $[B, D] = [B, E] \neq 0$ .

2. Given  $D$  such that for every  $\mathbf{r}$ ,  $[B, E] \neq 0$ , we will show that any  $\eta \in \text{span}(\text{row}(O'), \mathbf{1})$  will violate (B4). Consequently, any solution to (B4) will satisfy the constraint that  $\eta \notin \text{span}(\text{row}(O'), \mathbf{1})$ .

Assume, by way of contradiction, that  $\eta \in \text{span}(\text{row}(O'), \mathbf{1})$ . In this case  $\eta$  can be expressed as a linear combination of the rows of  $O'$  and of  $\mathbf{1}$ , namely

$$\eta^T = \mathbf{r}^T O' + \alpha \mathbf{1}^T, \quad (\text{B12})$$

for some  $\mathbf{r} \in \mathbb{R}^3$  and  $\alpha \in \mathbb{R}$ . Considering each part of the object separately we obtain

$$\begin{aligned}\eta_1^T &= \mathbf{r}^T P'_1 + \alpha \mathbf{1}^T \\ \eta_2^T &= \mathbf{r}^T P'_2 + \alpha \mathbf{1}^T \\ \eta_3^T &= \mathbf{r}^T (B P'_1 + C P'_2 + \mathbf{s} \mathbf{1}^T) + \alpha \mathbf{1}^T.\end{aligned}\quad (\text{B13})$$

Replacing this in (B4) we obtain

$$\begin{aligned}B \hat{\mathbf{v}}(\mathbf{r}^T P'_1 + \alpha \mathbf{1}^T) + C \hat{\mathbf{v}}(\mathbf{r}^T P'_2 + \alpha \mathbf{1}^T) \\ - \hat{\mathbf{v}}(\mathbf{r}^T (B P'_1 + C P'_2 + \mathbf{s} \mathbf{1}^T) + \alpha \mathbf{1}^T) \\ = [D, B] P'_1 + [D, C] P'_2 + \mathbf{t}' \mathbf{1}^T.\end{aligned}\quad (\text{B14})$$

Denote  $\mathbf{t}'' = \alpha(B + C - I)\hat{\mathbf{v}} + \hat{\mathbf{v}}\mathbf{r}^T \mathbf{s} + \mathbf{t}'$ . Rearranging, we get

$$\begin{aligned}(B(D + \hat{\mathbf{v}}\mathbf{r}^T) - (D + \hat{\mathbf{v}}\mathbf{r}^T)B)P'_1 \\ + (C(D + \hat{\mathbf{v}}\mathbf{r}^T) - (D + \hat{\mathbf{v}}\mathbf{r}^T)C)P'_2 \\ + \mathbf{t}'' \mathbf{1}^T = 0.\end{aligned}\quad (\text{B15})$$

Substituting  $E = D + \hat{\mathbf{v}}\mathbf{r}^T$  we obtain

$$[B, E]P'_1 + [C, E]P'_2 + \mathbf{t}'' \mathbf{1}^T = 0.\quad (\text{B16})$$

Since  $\text{rank}(\bar{J}(P'_1, P'_2)) = 7$  it follows that the above equation holds only if all the three following equations hold:

$$[B, E] = 0, \quad [C, E] = 0, \quad \text{and} \quad \mathbf{t}'' = 0.\quad (\text{B17})$$

However, by our assumption  $[B, E] \neq 0$  for all  $\mathbf{r}$ . It follows that (B17) does not hold, contradicting our assumption that  $\eta \in \text{span}(\text{row}(O'), \mathbf{1})$ .  $\square$

**Proof of Claim 6:** As in Section 4.2, we show this by first bringing  $B$  to a Jordan form. Let  $\bar{B} = A^{-1}BA$  and let  $E' = A^{-1}EA$  then  $[B, E] = 0$  if and only if  $[\bar{B}, E'] = 0$ . Since  $E = D + \hat{\mathbf{v}}\mathbf{r}^T$  we denote also  $D' = A^{-1}DA$ ,  $\mathbf{v}' = A^{-1}\hat{\mathbf{v}}$  and  $\mathbf{r}'^T = \mathbf{r}^T A$  (so that  $E' = D' + \mathbf{v}'\mathbf{r}'^T$ ). Note that  $\bar{B}$ ,  $D'$ ,  $\mathbf{v}'$  and  $\mathbf{r}'$  may be complex.

We begin by showing that if  $D'$  causes  $[\bar{B}, E'] \neq 0$  for all complex vectors  $\mathbf{r}'$  then there exists a real matrix  $D$  such that for every real vector  $\mathbf{r}$ ,  $[B, E] \neq 0$ . Consequently, it will suffice to show that for every  $\mathbf{v}'$

there exists a complex matrix  $D'$  such that for every complex  $\mathbf{r}$ ,  $[\bar{B}, D' + \mathbf{v}'\mathbf{r}'] \neq 0$ . Consider  $D'$  that satisfies this requirement. Denote  $AD'A^{-1} = D_1 + iD_2$ , where  $D_1$  and  $D_2$  are some  $3 \times 3$  real matrices, and denote  $A^{-T}\mathbf{r}' = \mathbf{r}_1 + i\mathbf{r}_2$ , where  $\mathbf{r}_1, \mathbf{r}_2 \in \mathbb{R}^3$  and  $A^{-T}$  denotes the inverse of  $A^T$ . If either  $D_1 = 0$  or  $D_2 = 0$  then we are done. Suppose that  $[B, D_1 + iD_2 + \hat{\mathbf{v}}(\mathbf{r}_1 + i\mathbf{r}_2)] \neq 0$  for every  $\mathbf{r}_1 + i\mathbf{r}_2$ . It follows that either

$$[B, D_1 + \hat{\mathbf{v}}\mathbf{r}_1] \neq 0 \quad (\text{B18})$$

or

$$[B, D_2 + \hat{\mathbf{v}}\mathbf{r}_2] \neq 0.\quad (\text{B19})$$

If (B18) is satisfied for every  $\mathbf{r} \in \mathcal{R}$ , then  $D_1$  is the sought real matrix. Similarly, if (B19) is satisfied for every  $\mathbf{r}$ , then  $D_2$  is the sought real matrix. One of these cases must hold since otherwise, assume that (B18) is not satisfied for some vector  $\mathbf{r}_1$  and that (B19) is not satisfied for some vector  $\mathbf{r}_2$ , then consider the complex vector  $\mathbf{r} = \mathbf{r}_1 + i\mathbf{r}_2$ , this commutator  $[B, E]$  will vanish for  $\mathbf{r}$  contradicting the assumption that  $[B, D_1 + iD_2 + \mathbf{v}'(\mathbf{r}_1 + i\mathbf{r}_2)] \neq 0$  for every  $\mathbf{r}_1 + i\mathbf{r}_2$ . Consequently, it is sufficient to show that for every  $\mathbf{v}'$  there exists a complex matrix  $D'$  such that for every complex vector  $\mathbf{r}'$ ,  $[\bar{B}, D' + \mathbf{v}'\mathbf{r}'^T] \neq 0$ .

We now turn to showing that for every vector  $\mathbf{v}'$  for which  $\text{rank}(W) = 3$  there exists a non-singular matrix  $D'$  such that for every  $\mathbf{r}'$ ,  $G = [\bar{B}, E'] \neq 0$ . We show this by looking at the shape of  $G = [\bar{B}, E']$  for every possible form of  $\bar{B}$  and deriving constraints on  $D'$  that guarantee that  $G = [\bar{B}, E'] \neq 0$  for any choice of  $\mathbf{r}'$ . Notice that by requiring that  $\text{rank}(W) = 3$  we exclude those views  $\hat{\mathbf{v}}$  which are eigenvectors of  $B$ . Below we consider six cases according to the number of linearly independent eigenvectors and different eigenvalues of  $B$  (see Table 2). Notice that  $f_{ij}$  in this table corresponds here to the components of  $E'$  (denoted as  $e'_{ij}$ ), and since  $E' = D' + \mathbf{v}'\mathbf{r}'^T$  these components are given by  $e'_{ij} = d'_{ij} + v'_i r'_j$ .

(a) As can be seen in Table 2, when  $B$  has three different eigenvalues  $G = 0$  if and only if all the six non-diagonal elements of  $E'$  are non-zero. In particular, consider the second row of  $G$ ,  $e'_{12} = e'_{32} = 0$  implies that  $d'_{12} = -v'_1 r'_2$  and  $d'_{32} = -v'_3 r'_2$ . Therefore, given a view  $\hat{\mathbf{v}}$  if  $v'_1 = 0$  we can choose any non-singular matrix  $D$  such that  $d'_{12} = 1$ , and if  $v'_1 \neq 0$  we can choose any  $D$  such that  $d'_{12} = 0$

and  $d'_{32} = 1$ . Note that  $v'_1$  and  $v'_2$  cannot vanish simultaneously since that would imply that  $\hat{v}$  is an eigenvector of  $B$ .

- (b) In this case  $G = 0$  implies in particular that  $d'_{13} = -v'_1 r'_3$  and  $d'_{23} = -v'_2 r'_3$ . If  $v'_1 = 0$  we can choose  $D$  such that  $d'_{13} = 1$ , and if  $v'_1 \neq 0$  we can choose  $D$  such that  $d'_{13} = 0$  and  $d'_{23} = 1$ .
- (c) In this case all vectors are eigenvectors of  $B$ , and so  $W$  is necessarily singular.
- (d)–(f) In these three cases,  $G = 0$  implies in particular that  $d'_{21} = -v'_2 r'_1$  and  $d'_{31} = -v'_3 r'_1$ . If  $v'_2 = 0$  we can choose  $D$  such that  $d'_{21} = 1$ , and if  $v'_2 \neq 0$  we can choose  $D$  such that  $d'_{21} = 0$  and  $d'_{31} = 1$ .  $\square$

## Acknowledgments

The authors thank Amitai Regev for his assistance in proving Proposition 7. This research was supported by a grant from the Israel Science Foundation No. 148/96. The vision group at the Weizmann Inst. is supported in part by the Israeli Ministry of Science, Grant No. 8504. Ronen Basri is an incumbent of Arye Dissentshik Career Development Chair at the Weizmann Institute.

## Note

- Note that dropping the assumption that  $\text{rank}(\bar{J}(P_1, P_2)) = 7$  will not change our final result that  $\mathcal{C}_{B,C,s}$  is not identifiable since extending the class with more objects cannot reduce the set of ambiguous views.

## References

- Basri, R. 1996. Paraperspective  $\equiv$  affine. *International Journal of Computer Vision*, 19(2):169–180.
- Burns J.B., Weiss R.S., and Riseman, E.M. 1993. View variation of point-set and line-segment features. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 15(1):51–68.
- Clemens, D. and Jacobs, D. 1991. Space and time bounds on model indexing. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 13(10):1007–1018.
- Fischler, M.A. and Bolles, R.C. 1981. Random sample consensus: a paradigm for model fitting with application to image analysis and automated cartography. *Com. of the A.C.M.*, 24(6):381–395.
- Faugeras, O.D. 1992. What can be seen in three dimensions with an uncalibrated stereo rig? In *Second European Conf. on Computer Vision*, pp. 563–378.
- Fawcett, R., Zisserman, A., and Brady, J.M. 1994. Extracting structure from an affine view of a 3D point set with one or two bilateral symmetries. *Image and Vision Computing*, 12(9):615–622.
- Forsyth, D., Mundy, J.L., Zisserman, A., Coelho, C., Heller, A., and Rothwell, C. 1991. Invariant descriptors for 3-D object recognition and pose. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 13(10):971–991.
- Hartley, R.I. 1993. Invariants of lines in spaces, In *Proc. of DARPA Image Understanding Workshop*, pp. 737–744.
- Hartley, R.I. 1994. Projective reconstruction and invariants from multiple images. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 16(10):1036–1041.
- Huttenlocher, D.P. and Ullman, S. 1990 “Recognizing Solid Objects by Alignment with an Image, *Int. J. Computer Vision*, 5(2):195–212.
- Jacobs, D. 1992. Space efficient 3D model indexing. In *IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 439–444.
- Jones, M.J. and Poggio, T. 1995. Model-Based Matching of Line Drawings by Linear Combinations of Prototypes. In *International Conference on Computer Vision*, pp. 531–536
- Koenderink, J. and van Doorn, A. 1991. Affine structure from motion. *Journal of the Optical Society of America*, 8(2):377–385.
- Lamdan, Y. and Wolfson, H.J. 1988. Geometric Hashing: A General and Efficient Model-Based Recognition Scheme. In *Second International Conference on Computer Vision*, pp. 238–249.
- Lowe, D. 1985. *Perceptual Organization and Visual Recognition*. Kluwer Academic Publishers: The Netherlands.
- Mitsumoto, H., Tamura, S., Okazaki, K., Kajimi, N., and Fukui, Y. 1992. 3-D reconstruction using mirror images based on a plane symmetry recovering method. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 14(9):941–946.
- Moses, Y. and Ullman, S. 1992. Limitations of non model-based recognition schemes. In *Second European Conf. on Computer Vision*, pp. 820–828.
- Mundy, J.L. and Zisserman, A. 1994. Repeated structures: image correspondence constraints and 3D structure recovery. In *Applications of Invariance in Computer Vision*, J.L. Mundy, A. Zisserman and D.A. Forsyth (Eds.), Lecture Notes in Computer Science, Vol. 825, Springer-Verlag.
- Poggio, T. and Vetter, T. 1992. Recognition and structure from one 2D model view: observations on prototypes, object classes and symmetries. *M.I.T., A.I. Memo No. 1347*.
- Quan, L. 1995. “Invariants of six points and projective reconstruction from three uncalibrated images. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 17(1):34–46.
- Quan, L. 1996. Conic reconstruction and correspondence from two views.” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 18(2):151–160.
- Rothwell, C.A. 1995. *Object Recognition through Invariant Indexing*. Oxford University Press, Oxford, UK.
- Rothwell, C.A., Zisserman, A., Forsyth, D.A., and Mundy, J.L. 1992. Canonical Frames for Planar Object Recognition. In *Proc. of 2nd European Conf. on Computer Vision*, pp. 757–772.
- Sparr, G. 1992. Depth Computations from Polyhedral Images. *Image and Vision Computing*, 10:683–688.
- Sugimoto, A. 1994. Geometric invariant of noncoplanar lines in a single view. In *Proc. of Int. Conference on Pattern Recognition*, 1:190–195.
- Thompson, D.W. and Mundy, J.L. 1987. Three dimensional model matching from an unconstrained viewpoint. In *Proc. of IEEE Int. Conf. on robotics and Automation*, pp. 208–220.
- Tomasi, C. and Kanade, T. 1992. Shape and Motion from Image Streams under Orthography: a Factorization Method. *International Journal of Computer Vision*, 9(2):137–154.
- Ullman, S. and Basri, R. 1991. Recognition by linear combinations

- of models. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 13(10):992–1006.
- Weiss, I. 1988. Projective invariants of shape. *DARPA Image Understanding Workshop*, pp. 1125–1134.
- Weinshall D. 1993. Model-based invariants for 3D vision,” *International Journal of Computer Vision*, 10(1):27–42.
- Weinshall, D. and Werman, M. 1997. On view likelihood and stability. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 19(2):97–108.
- Werman, M. and Shashua, A. 1995. The study of 3D-from-2D using elimination. *Fifth International Conference Computer Vision*, pp. 473–479.
- Yan, X., Peng, J.X., Ding, M.Y., and Xue, D.H. 1997. The unique solution of projective invariants of 6 points from 4 uncalibrated images. *Pattern Recognition*, 30(3):513–517.
- Zisserman, A., Forsyth, D., Mundy, J., Rothwell, C., Liu, J., and Pillow, N. 1995. 3D object recognition using invariance. *Artificial Intelligence*, 78:239–288.