A Framework for Reconstruction based Recognition of Partially Occluded Repeated Objects

Ragini Choudhury\textsuperscript{a}  Santanu Chaudhury\textsuperscript{b}  J. B. Srivastava \textsuperscript{c}

\textsuperscript{a} : INRIA Rhone-Alpes, ZIRST- 655 avenue de l’Europe, 38330 Montbonnot St. Martin, France\textsuperscript{1}
Ragini.Choudhury@inrialpes.fr

\textsuperscript{b} : Dept. of Elec. Engg., Indian Institute of Technology, Delhi, Hauz Khas New Delhi 110016, India
santanuc@cse.iitd.ernet.in

\textsuperscript{c} : J. B. Srivastava, Dept. of Mathematics, Indian Institute of Technology, Delhi, Hauz Khas,
New Delhi 110016, India
jbsrivas@maths.iitd.ernet.in

\textbf{Running title} : Recognition : Partially Occluded Repeated Objects

\textbf{Running author} : Choudhury et. al.

\footnotesize{\textsuperscript{1}The work was done while the author was at the Indian Institute of Technology, Delhi (India) with the Dept. of Electrical Engg. and the Dept. of Mathematics}
Abstract

In this paper we propose a reconstruction based recognition scheme for objects with repeated components, using a single image of such a configuration, in which one of the repeated components may be partially occluded. In our strategy we reconstruct each of the components with respect to the same frame and use these to compute invariants. We propose a new mathematical framework for the projective reconstruction of affinely repeated objects. This uses the repetition explicitly and hence is able to handle substantial occlusion of one of the components. We then apply this framework to the reconstruction of a pair of repeated quadrics. The image information required for the reconstruction are the outline conic of one of the quadrics and correspondence between any four points which are images of points in general position on the quadric and its repetition. Projective invariants computed using the reconstructed quadrics have been used for recognition. The recognition strategy has been applied to images of monuments with multi-dome architecture. Experiments have established the discriminatory ability of the invariants.

1 Introduction

In this paper we present an invariant based recognition scheme for a special class of 3D objects having repeated components. The recognition scheme uses a single image of such a configuration, and permits partial occlusion of one of the repeated components. In the proposed strategy, we explicitly reconstruct these repeated components and use the reconstructed subparts to compute the invariants. We propose a mathematical framework for reconstructing the components which characterizes the 3D constraints between the components by exploiting the repetition explicitly. The reconstruction strategy is adapted to the special case when the repeated components are quadrics. Quadrics are 3D shapes like ellipsoids, hyperboloids etc. In real 3D scenes, they can occur as individual objects or as parts of a more complex object. Historical monuments with multi-dome architecture are examples of such configurations.

The tremendous appeal of invariant based recognition schemes for 3D objects arises from the
fact that invariants provide viewpoint independent descriptors of these objects. However, unlike planar objects [9, 15, 17] it is not possible to formulate invariants for the general class of 3D objects [2]. Invariants of only specific classes of 3D objects can be computed from a single image, by exploiting certain class specific geometric constraints [21]. In this paper, we explore a special class of repeated objects, namely objects with repeated quadric components. Such reconstruction based recognition schemes, as proposed in this paper, in which the reconstructed components are used to compute invariants, for recognition, are few in literature. The added novelty of our strategy is that it can handle the partial occlusion of one of the components.

We first develop a new framework for the projective (relative affine) reconstruction of repeated objects from a single uncalibrated image by converting it to its equivalent stereo image framework where the second camera is a transform of the original camera. This new mathematical framework proposed transforms a pair of uncalibrated cameras such that the first camera gets aligned. The second camera matrix involves a homography between two image planes and a normalized second epipole, making its form computable from image information. Although relative affine structure has been obtained in the past [12, 18] using two images of an object, the novelty in the contribution of our paper lies in developing the theory which handles reconstruction of repeated objects from a single image, even when one of the components is partially occluded. We use the image information available from a single image and exploit the repetition explicitly, to reconstruct both the components with respect to the same frame, unlike earlier schemes which reconstructed a single object from two images. Our reconstruction scheme is, in fact, general and applicable to different classes of repeated objects. It requires correspondences of four points which are images of points in general position, on the object and its repetition, and the choice of four points is not critical to the strategy. When each of the components is a quadric, we require the additional image information of the outline conic of the first quadric only. This permits the algorithm to handle partial occlusion of the repeated quadric.

Work on repeated objects, in the past, has concentrated on their reconstruction, handled by converting a single image to the equivalent multiple view of the single instance. This has led to the convergence of the single view [13, 14] and multiple view based approaches [6, 16]. The affine structure for translationally repeated objects has been obtained by Moons [13] using vanishing
points and five point correspondences between the two views. Mundy and Zisserman [14] have studied affine structure as a projective ambiguity matrix. Shashua [18, 19] has handled affine structure as a special case of relative affine structure. By exploiting this affine ambiguity, Liu et. al. [10] obtained 3D affine invariants for recognizing 3D translationally repeated objects. But affine structure does not suffice for an extension to reconstruction scheme for translationally repeated quadrics. Therefore the technique of [10] is not applicable to quadric configurations, in our context and hence in our strategy we use the relative affine structure [19].

Work on quadrics has also been limited to their reconstruction [4, 20]. Although Shashua et. al. [20] and Cross et. al. [4] have reconstructed quadrics using two views, they do not address the problem of repeated objects. Shashua and Toelg [20] reconstructed a quadric reference surface using the knowledge of outline conic and four point correspondences in the two images. Cross and Zisserman [4] do a quadric reconstruction using dual space geometry. Repeated applications of these strategies, either by taking both the components together or by reversing the camera set up, will lead to components which are reconstructed with respect to different frames and hence cannot be used for computing joint invariants essential for recognition. Our reconstruction framework overcomes this difficulty by incorporating the repetitions in the same frame.

In our recognition strategy, the values of projective invariants computed using these reconstructed quadrics have been used to recognize images of such scenes. We have formulated projective invariants for a pair of proper quadrics. The reconstruction scheme requires the visibility of any four corresponding points in general position on both the quadrics. Since an alternate choice of four points leads to the reconstruction of a projectively equivalent configuration, the values of projective invariants are the same for the original and the reconstructed configuration. Therefore the proposed recognition strategy can identify scenes with substantial occlusion of one of the quadrics. In [3], the recognition strategy was designed for repeated objects in which both the components needed to be visible. The strategy proposed in our paper relaxes this constraint. We need the outline conic of the first quadric, knowledge that a repeated quadric exists and four point correspondences on the object and its repetition. This makes the proposed recognition framework robust to occlusion of the repeated component. The proposed recognition strategy is applied to images of monuments which are characterized by repeated domes each of which can be
modeled as a quadric. The proposed strategy is applicable for these scenes because the possibility of partial occlusion of one of the domes in the images of these monuments is, in general, high. Experiments have established the discriminatory ability of the invariants.

The paper is organized as follows. Section 2 gives the mathematical background required for the paper, and defines the projective ambiguity matrix used in reconstruction. Section 3 gives a projective reconstruction of affinely repeated objects, in general. In section 4, we have reconstructed a pair of affinely repeated quadrics. Section 5 deals with the computation of joint projective invariants for an arbitrary pair of proper quadrics. These invariants are then used for the purpose of recognition of quadric configurations. Section 6 contains experimental results regarding discriminatory power and the stability of the invariants. Implementation has been done on synthetic and real images. The applications of the theory developed, have been carried out for the special case of famous historical monuments, which contain translationally repeated domes. Finally, some conclusions are given at the end.

2 Problem Definition and Mathematical Background

In this paper, the aim is to develop a recognition framework for 3D objects with repeated components, which takes into account the repetition explicitly. The repetition can be affine, rigid or a translation. Objects in this paper contain components $S$ and $S'$ such that $S' = A(S)$ where $A = \begin{pmatrix} A & t \\ 0 & 1 \end{pmatrix}$, $A = 3 \times 3$ non-singular matrix, $t = (t_1 \ t_2 \ t_3)^t$, denotes an affine transformation. Such objects are called Affinely Repeated Objects. If $A = R = \text{Rotation matrix}$, then the objects are said to be Rigidly repeated. When $A = I_3$, the transformation is a translation and such objects are called Translationally repeated objects. We work in the projective space and use homogeneous coordinates. Also all results are proved for affine repetition. The case of translation is discussed as special case of this. In this section we review the preliminaries of epipolar geometry, homography between image planes with respect to a reference plane $\Pi$ and the projective ambiguity matrix.

A camera here means a perspective uncalibrated camera represented by a $3 \times 4$ matrix of the form $[P \ p]$ with $\det(P) \neq 0$. In this situation, the centre of perspectivity (COP) in homogeneous
coordinates given by \( \text{COP} \approx \begin{pmatrix} -P^{-1}p \\ 1 \end{pmatrix} \), where \( \approx \) denotes the projective equivalence. The problem of reconstruction and recognition of repeated objects from a single image is best handled by converting the single image framework to its equivalent stereo image framework. Let \( \tilde{P} = [P \ p] \) and \( \tilde{P}' = [P' \ p'] \) represent two cameras in the stereo framework. Let their respective centre of perspectivities be given by \( \text{COP}_1 = \begin{pmatrix} -P^{-1}p \\ 1 \end{pmatrix} \) and \( \text{COP}_2 = \begin{pmatrix} -P'^{-1}p' \\ 1 \end{pmatrix} \) and respective image planes \( \mathcal{R}_1 \) and \( \mathcal{R}_2 \).

Let \( M = (X \ Y \ Z \ 1)^t \) represent an arbitrary 3D point in homogeneous coordinates. Define
\[
\bar{m} = [P \ p]M = P \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} + p = \begin{pmatrix} x_M \\ y_M \\ z_M \end{pmatrix} = \lambda m, \ \lambda = z_M \neq 0 \text{ and } m = (x \ y \ 1)^t \text{ is observable in the image. Similarly, } \bar{m}' = [P' \ p']M = \lambda' m', \ \lambda' \neq 0, \ m' = (x' \ y' \ 1)^t \text{ where } m' \text{ is observable in the image. By definition, } m \text{ and } m' \text{ or } \bar{m} \text{ and } \bar{m}' \text{ are said to be corresponding points.}
\]

The epipoles are defined as follows:
\[
\bar{e} = [P \ p] \begin{pmatrix} -P'^{-1}p' \\ 1 \end{pmatrix} = \lambda_e \begin{pmatrix} e_1 \\ e_2 \\ 1 \end{pmatrix} = \lambda_{ee} e, \ e = \begin{pmatrix} e_1 \\ e_2 \\ 1 \end{pmatrix}
\]
\[
\bar{e}' = [P' \ p'] \begin{pmatrix} -P^{-1}p \\ 1 \end{pmatrix} = \lambda_{e'} \begin{pmatrix} e'_1 \\ e'_2 \\ 1 \end{pmatrix} = \lambda_{e'e'} e', \ e' = \begin{pmatrix} e'_1 \\ e'_2 \\ 1 \end{pmatrix}
\]

\( H_{\Pi} \) is the homography from image plane \( \mathcal{R}_1 \) to the image plane \( \mathcal{R}_2 \) via the plane \( \Pi \). It is represented by a \( 3 \times 3 \) non-singular matrix up to a scale, thus given 4 image point correspondences of points in general position, the homography \( H_{\Pi} \) can be computed up to a scale [19]. When \( \Pi = \Pi_{\infty} \), the plane at infinity, then \( H_{\Pi_{\infty}} = H_{\infty} \), called the Infinite Homography. Also \( H_{\infty} = P'P^{-1} \) [12]. The relationship between \( H_{\Pi} \) and \( H_{\infty} \) given below can be found in Shashua and Navab (Shashua et al 1994) and Luong and Vieville (Luong et al 1996).

**Theorem 2.1** Let \( \tilde{P} = K[R_w \ t_w] \) and \( \tilde{P}' = K'[R'_w \ t'_w] \) be the two cameras. Let \( \Pi \) be the reference plane. Then
\[
H_{\Pi} = H_{\infty} + e'_N v^t_{\Pi_x} \tag{1}
\]
where \( e'_N = \frac{e'}{\|e'\|} \), the normalized epipole, \( H_{\Pi} \) is the homography between the two image planes, with respect to the plane \( \Pi \), and \( \nu^t_{\Pi N} \) is expressed in terms of the coordinates of the plane \( \Pi \) and the internal camera parameters.

Now, \( \tilde{e'} = [P' p'] \begin{pmatrix} -P^{-1} p \\ 1 \end{pmatrix} = -P' P^{-1} p + p' = p' - H_{\infty} p \). Following a similar kind of reasoning, we have \( H_{\infty} \tilde{e'} = P' P^{-1} [P p] \begin{pmatrix} -P^{-1} p' \\ 1 \end{pmatrix} = -\tilde{e'} \). Using (1) we get, \( \nu^t_{\Pi N} = e'^t_{N}(H_{\Pi} - H_{\infty}) \). Thus, \( H_{\Pi} \tilde{e'} = H_{\infty} \tilde{e'} + e'_N \nu^t_{\Pi N} \tilde{e'} = (-1 + \frac{\lambda e'_N}{\|e'_N\|}) \tilde{e'} = \gamma \tilde{e'} \). Therefore, we have been able to prove the following simple and useful results: (i) \( H_{\Pi} \tilde{e'} = \gamma \tilde{e'} \) for some \( \gamma \in \mathbb{R} \), \( \gamma \neq 0 \). (ii) \( \tilde{e'} = p' - H_{\infty} p \) (iii) \( H_{\infty} \tilde{e'} = -\tilde{e'} \).

We now define a projective reconstruction matrix which aligns the first camera \([P p]\). Define, \( \mathcal{H} = \begin{bmatrix} P \\ L^t_N \\ \nu_N \end{bmatrix} \), \( L^t_N = -\nu^t_{\Pi N} P, \quad \nu_N = -\nu^t_{\Pi N} p + \lambda e' ||e'||, \quad \nu^t_{\Pi N} = e'^t_{N}(H_{\Pi} - H_{\infty}) \), and \( e'_N = \frac{e'}{\|e'\|} \).

\( e' \) is the second epipole and \( H_{\Pi} \) the homography with respect to a plane \( \Pi \). It can be easily seen that \( \mathcal{H} \) is invertible. Straightforward computations show that \([P p] = [I_3 \ 0] \mathcal{H}, \quad [P' p'] = [H_{\Pi} e'_N] \mathcal{H} \). Let \( M = (X \ Y \ Z \ 1)^t \) represent an arbitrary 3D point in homogeneous coordinates. Then its images in the two cameras are \( \bar{m} = [P p] M = [I_3 \ 0] \mathcal{H} M, \quad \bar{m}' = [P' p'] M = [H_{\Pi} e'_N] \mathcal{H} M \).

Thus, the transformed cameras \([I_3 \ 0]\) and \([H_{\Pi} e'_N]\) map \( \mathcal{H} M \) to the same points as original cameras \([P p]\) and \([P' p']\) map \( M \). It is this specific choice of \( \mathcal{H} \) which makes the form of the transformed cameras image computable. The reconstructed point is \( \bar{M} = \mathcal{H} M \).

In the next section, we discuss the reconstruction framework for repeated objects using the above reconstruction matrix.

3 Reconstruction of Repeated Objects

In this section, we projectively reconstruct affinely repeated objects from a single image. The proposed reconstruction framework is general and applicable to all repeated objects. The reconstruction problem for repeated objects from a single image is best handled by converting the problem from a single image framework to the equivalent stereo image framework where the
second camera is a transform of the original camera. Let \([P\ p]\), \([P'\ p']\) be the two cameras in stereo framework. Then \([P'\ p'] = [P\ p]A = [PA\ Pt + p]\). Therefore

\[
P' = PA, \quad p' = Pt + p \quad \text{and} \quad H_\infty = PA P^{-1}
\]

When \(A\) is a translation, \(A = I_3\) and the infinite homography \(H_\infty = PP^{-1} = I\). Hence from section 2, \(H_\infty\hat{e} = \hat{e} = -\hat{e}'.\)

The relative affine structure for objects using two images has been obtained [19, 12]. But computing the relative affine structure for a repeated object is a new and significant contribution of the paper. Also both the components are reconstructed with respect to the same frame, which is essential in order to be able to use these reconstructed components to compute joint invariants. By using the relative affine structure [12, 19] repeatedly, the components are reconstructed with respect to different frames, which cannot be used to compute joint invariants. Thus the reconstruction of the repeated component with respect to the same frame as the first component is a very significant contribution of our reconstruction framework. The proposed reconstruction scheme makes explicit use of the repetition. We now present a framework for the reconstruction of general affinely repeated objects by establishing a relationship between the reconstructed point on the first object and the affinely repeated reconstructed point on the affine repetition. This is then extended to obtain a reconstruction framework for affinely repeated quadrics.

The point \(M\) on the first component \(S\) is reconstructed as \(\mathcal{HM} = \tilde{M}\) as in section 3.1. This is repeated for each point on the first component and hence it is completely reconstructed. Then in theorem 3.1, a relationship is established between \(\tilde{M}\) and its reconstructed affinely repeated point \(\tilde{M}' = \mathcal{HM}'\), which is computable from image information, upto the knowledge of infinite homography. This is used to reconstruct the repeated component.

### 3.1 Reconstruction of \(M\) projectively as \(\mathcal{HM}\)

Let \(M = (X\ Y\ Z\ 1)^t = (\tilde{M}\ 1)^t\) where \(\tilde{M} = (X\ Y\ Z)^t\) be a point on the first component and \(\mathcal{HM} = (A\ B\ C\ D)^t\) be its reconstructed point. We know that \(\bar{m} = [P\ p]M = P\tilde{M} + p\).

Also \(\mathcal{HM} = \begin{pmatrix} \bar{m} \\ L_N^t\tilde{M} + \nu_N \end{pmatrix}\). This gives \(\bar{m} = (A\ B\ C)^t\) and \(\mathcal{HM} = \begin{pmatrix} \bar{m} \\ D \end{pmatrix} = \begin{pmatrix} \bar{m} \\ \lambda k \end{pmatrix}\) where \(\lambda\) is
\[ k = \frac{p}{x}. \quad \bar{m}' = [P' p']M = [H_{II} e'_{N}]\mathcal{H}M = H_{II}\bar{m} + \lambda ke'_{N}, \text{ i.e., } \lambda' m' = \lambda(H_{II}m + ke'_{N}). \]  

Taking cross product on both sides by \( m' \) and simplifying, we get \( k(m' \times e'_{N}) = H_{II}m \times m' \). Therefore \( k = \frac{(m' \times e'_{N})(H_{II}m \times m')}{||m' \times e'_{N}||^2} \) and \( \mathcal{H}M = \lambda(m'k)^{t} = \lambda(x y 1 k)^{t} \approx (x y 1 k)^{t} \), as desired.

As can be seen \( k \) is completely image computable. In case of translational repetition, the epipoles can be computed using a pair of image point correspondences. In case of general repetition, eight or more point correspondences are required [8, 12]. \( k \) denotes the relative affine structure [12, 20] and can be computed for all points on the first component assuming their point correspondences are known. Hence we have been able to reconstruct the first component.

### 3.2 Reconstruction of the Repeated Component

In the next theorem, we establish a relationship between the original reconstructed point \( \bar{M} \) and the translated reconstructed point \( \bar{M}' \). Thus \( \bar{M}' \) is implicitly reconstructed without computing \( k' \) explicitly. This is done for each point on the second component and hence the second component has been reconstructed pointwise.

**Theorem 3.1 (Relationship between \( \bar{M} \) and \( \bar{M}' \))** Let \( \bar{M} \) be the original reconstructed point and \( \bar{M}' \) be the reconstructed affinely repeated point \( M' = \mathcal{A}(M) \) of \( M \). Then

\[
\bar{M}' = (\mathcal{H}\mathcal{A}\mathcal{H}^{-1})\bar{M}
\]  

where

\[
\mathcal{H}\mathcal{A}\mathcal{H}^{-1} = \begin{pmatrix}
H_{II} & e'_{N} \\
\nu'_{II, N}(I - H_{II}) & 1 - \nu'_{II, N}e'_{N}
\end{pmatrix}
\]  

and \( \bar{M} \approx (x y 1 k)^{t} \) as reconstructed in section 3.1.

**Proof**: Let \( M \) be the original point in 3D homogeneous coordinates. Also let \( \bar{M} = \mathcal{H}M \) be the reconstructed point of \( M \). Therefore \( \mathcal{H}^{-1} \bar{M} = M \). Now \( M' = \mathcal{A}M \) is the affinely repeated point where \( \mathcal{A} = \begin{pmatrix} A & t \\ 0 & 1 \end{pmatrix} \) is the affine transformation. Also \( \mathcal{H}M' = \bar{M}' \) is the reconstructed affinely repeated point. Then

\[
\bar{M}' = \mathcal{H}M' = \mathcal{H}\mathcal{A}M = (\mathcal{H}\mathcal{A}\mathcal{H}^{-1})\bar{M}
\]
Computation of $\mathcal{H}A\mathcal{H}^{-1}$:

It can be easily shown that

$$\mathcal{H}^{-1} = \begin{pmatrix} D & b \\ c^t & \mu \end{pmatrix}$$

where $D = P^{-1}(I - \frac{\nu_{1_N}^t}{\lambda_{c'}||e'||})$, $b = -\frac{\nu_{1_N}^t p}{\lambda_{c'}||e'||}$, $c^t = \frac{\nu_{1_N}^t}{\lambda_{c'}||e'||}$ and $\mu = \frac{1}{\lambda_{c'}||e'||}$.

Now

$$\mathcal{H}A\mathcal{H}^{-1} = \begin{pmatrix} P & p \\ L_N^t & \nu_N \end{pmatrix} \begin{pmatrix} A & t \\ 0 & 1 \end{pmatrix} \begin{pmatrix} D & b \\ c^t & \mu \end{pmatrix}$$

$$= \begin{pmatrix} PAP^{-1}(I_N - \frac{\nu_{1_N}^t}{\lambda_{c'}||e'||}) + (Pt + p) \frac{\nu_{1_N}^t}{\lambda_{c'}||e'||} \\ L_N^t PAB + \mu (Pt + p) \\ L_N^t AD + (L_N^t t + \nu_N) c^t \\ L_N^t Ab + \mu (L_N^t t + \nu_N) \end{pmatrix}$$

(5)

We will compute each component of $5$ separately.

$$PAP^{-1}(I_N - \frac{\nu_{1_N}^t}{\lambda_{c'}||e'||}) + (Pt + p) \frac{\nu_{1_N}^t}{\lambda_{c'}||e'||} = H_{\infty} - \frac{H_{\infty} \nu_{1_N}^t}{\lambda_{c'}||e'||} + \frac{p' \nu_{1_N}^t}{\lambda_{c'}||e'||} \left( \begin{array}{c} \text{because } Pt + p = p' \text{ and } PAP^{-1} = H_{\infty} \text{ from } 2 \\ \text{from Theorem 2.1} \end{array} \right)$$

(6)

$$PAB + \mu (Pt + p) = PA(-\frac{\nu_{1_N}^t p}{\lambda_{c'}||e'||}) + \frac{1}{\lambda_{c'}||e'||}(Pt + p)$$

$$= \frac{-\nu_{1_N}^t P + p'}{\lambda_{c'}||e'||} \left( \begin{array}{c} \text{because } Pt + p = p' \text{ using } 2 \\ \text{using results in section } 2 \end{array} \right)$$

(7)

Also

$$L_N^t AD + (L_N^t t + \nu_N) c^t = -\nu_{1_N}^t PAP^{-1} - \nu_{1_N}^t Ptc^t - \nu_{1_N}^t pc^t + \lambda_{c'}||e'||c^t$$

$$= \nu_{1_N}^t (I - H_{\Pi}) \left( \begin{array}{c} \text{using equation } 6 \end{array} \right)$$

(8)

Finally

$$L_N^t Ab + \mu (L_N^t t + \nu_N) = -\nu_{1_N}^t PAb + \mu(-\nu_{1_N}^t Pt - \nu_{1_N}^t p + \lambda_{c'}||e'||)$$

$$= -\nu_{1_N}^t e_N^t + 1 \left( \begin{array}{c} \text{from equation } 7 \end{array} \right)$$

(9)
Thus using equations 6, 7, 8, 9 with 5, we have been able to obtain the desired form of $\mathcal{HAH}^{-1}$.

$\Box$

Now $\nu^I_{1N} = e^I_N(H_1 - H_\infty)$. Here $H_1$ and $e^I_N$ are image computable. If $H_\infty$ can be computed from image information, then $\mathcal{HAH}^{-1}$ is image computable. Having reconstructed $\tilde{M}$ in section 3.1, we can use the image computable matrix $\mathcal{HAH}^{-1}$ to obtain $\tilde{M}'$. As will be seen, the application of this method to the reconstruction of configurations with affinely repeated quadric components, reduces the requirement of the image information. When $A = T = \text{translation}$, all the components of the matrix $\mathcal{HTH}^{-1}$ are computable from available image information. The components involve $H_1$, which is computable from image information of 4 image point correspondences, $e^I_N$ which can be computed using at least 2 image point correspondences, exploiting the translating nature of the camera (auto-epipolarity) and $\nu^I_{1N} = e^I_N(H_1 - H_\infty)$. This is image computable because $H_\infty = I_3$ for a translating camera. Next we use this strategy, with additional image information, to reconstruct objects in which the repeated components are quadrics.

### 4 Reconstruction of Quadrics

We now apply the reconstruction framework to the specific case when each of the components is a quadric. The objects we now consider contain quadrics $Q$ and $Q'$ such that $Q' = A(Q)$ where $A = \begin{pmatrix} A & t \\ 0 & 1 \end{pmatrix}$, $A = 3 \times 3$ non-singular matrix, $t = (t_1, t_2, t_3)^t$, denotes an affine transformation. Such objects are called *Affinely Repeated Quadrics*. When $A = I_3$, the transformation is a translation and such objects are called *Translationally repeated quadrics*.

A *quadric surface* is an algebraic surface consisting of all points in $\mathbb{R}^3$ in homogeneous coordinates satisfying a homogeneous polynomial of degree 2 in four variables $X_1, X_2, X_3, X_4$ over the field $\mathbb{R}$ of real numbers. General equation of the quadric surface is

$$S = S(X_1, X_2, X_3, X_4) = \sum_{i=1}^{4} \sum_{j=1}^{4} q_{ij} X_i X_j = 0$$

i.e., $S(X) = X^t Q X = 0$, where $Q = (q_{ij})$ is a $4 \times 4$ real symmetric matrix and $X = (X_1, X_2, X_3, X_4)^t$ is a homogeneous $4 \times 1$ vector. The quadric surface defined above is said to be proper if $\det(Q) \neq 0$, i.e., $Q$ is non-singular and has 9 independent parameters. The image information used should be sufficient to recover the nine parameters.
Let $Q = \begin{pmatrix} Q_{33} & q \\ q^t & q_{44} \end{pmatrix}$, $Q_{33}$ a $3 \times 3$ matrix, $q$ a $3 \times 1$ vector, $q_{44}$ a scalar, be a quadric and let $\tilde{Q} = \mathcal{H}(Q) = \begin{pmatrix} \tilde{Q}_{33} & \tilde{q} \\ \tilde{q}^t & \tilde{q}_{44} \end{pmatrix}$ be the quadric to be reconstructed. It is not difficult to see that 

$\tilde{Q} = (\mathcal{H}^{-1})^t Q \mathcal{H}^{-1}$ (upto a scale). It can also be seen that the outline conic $\tilde{C}$ (image of the rim) of $\tilde{Q}$ is determined by $\tilde{C} = \tilde{q} \tilde{q}^t - \tilde{q}_{44} \tilde{Q}_{33}$ and is the same for the original and the reconstructed quadric.

In Theorem 4.1, we obtain the framework for reconstructing $\tilde{Q}$ and the affinely repeated quadric $\tilde{Q}'$. In order to reconstruct the quadrics $\tilde{Q}, \tilde{Q}'$ we need to recover 9 parameters for each. The image information used are image point correspondences of four points on the quadric $Q$ and the corresponding points on the affinely repeated quadric $Q'$, together with the outline conic of $Q$. (The knowledge of epipoles is also assumed.) The point correspondences are used to reconstruct the points. Thus we get $k_i', i = 1, 2, 3, 4$, one with respect to each point correspondence pair. These 4 parameters together with the 5 parameters of the outline conic, are sufficient to solve for the 9 parameters of the reconstructed quadrics. Since we exploit the constraints imposed by the type of repetition between the components, no additional information is required to reconstruct the repeated component.

Affine repetition has been explicitly taken into account in this framework, by establishing a relationship between the original reconstructed quadric and the affinely reconstructed quadric. In Theorem 4.1, we establish this relationship, then reconstruct $\tilde{Q}$ and use the derived relationship to reconstruct the repeated quadric $\tilde{Q}'$. The theorem is significant from the viewpoint of recognition, because both the quadrics are reconstructed with respect to the same frame.

**Theorem 4.1 (Reconstruction of $\tilde{Q}$ and $\tilde{Q}'$):** A single uncalibrated perspective image of a pair of proper quadrics $Q$ and $Q'$ is given where $Q'$ is a repetition of $Q$. Further image point correspondences of 4 identifiable non-coplanar points $M_i$, $i = 1, 2, 3, 4$ on $Q$ and corresponding points $M_i'$, $i = 1, 2, 3, 4$ on $Q'$ are given. Also known in the image is the outline conic $C$ of first quadric component $Q$. Then the quadrics $\tilde{Q}$ can be reconstructed projectively as $\tilde{Q}$. Let $\tilde{Q}'$ be the reconstructed affinely repeated quadric, then

$$\tilde{Q}' = (\mathcal{H} \mathcal{H}^{-1})^{-1} \tilde{Q} (\mathcal{H} \mathcal{H}^{-1})^{-1}$$

(10)
where \( \mathcal{H}_A \mathcal{H}^{-1} = \left( \begin{array}{cc} H_{ii} & e' \\ \nu_{ii} (I - H_{ii}) & 1 - \nu_{ii} e' \\ \end{array} \right) \)

**Proof** Let \( Q \) be the quadric in 3D whose reconstruction is being considered. Then for points \( X \in Q \), \( X^t Q X = 0 \). Let \( \bar{X} = \mathcal{H} X \) be the reconstructed point and \( \bar{Q} \) be the reconstructed quadric. Then \( \bar{X}^t \bar{Q} \bar{X} = 0 \). \( \bar{X} \in \bar{Q} \) if and only if \( X \in Q \). Substituting for \( X \), we get

\[
(\mathcal{H}^{-1} \bar{X})^t Q (\mathcal{H}^{-1} \bar{X}) = 0 \Rightarrow \bar{X}^t (\mathcal{H}^{-1} Q \mathcal{H}^{-1}) \bar{X} = 0
\]  (11)

Comparing we have

\[
\bar{Q} = \mathcal{H}^{-1} Q \mathcal{H}^{-1} \quad \text{upto a scale}
\]  (12)

The affinely repeated quadric is \( Q' \), then \( X'^t Q' X' = 0 \) where \( X' = A X \). Now \( X' = A X \Rightarrow X = A^{-1} X', X' \in Q' \) if and only if \( X \in Q \). Therefore we have

\[
(A^{-1} X')^t Q (A^{-1} X') = 0 \quad \text{for all } X' \in Q'
\]  (13)

On comparing we get,

\[
Q' = A^{-1} Q A^{-1} \quad \text{upto a scale}
\]  (14)

Let \( \bar{X}' = \mathcal{H} X' \) be the translated reconstructed point and \( \bar{Q}' \) be the translated reconstructed quadric. Then, \( \bar{X}'^t \bar{Q}' \bar{X}' = 0 \). Substituting for \( X' \) we get \( (\mathcal{H}^{-1} \bar{X}')^t Q' (\mathcal{H}^{-1} \bar{X}') = 0 \), for all \( \bar{X}' \in \bar{Q}' \). Thus

\[
\bar{X}'^t (\mathcal{H}^{-1} Q' \mathcal{H}^{-1}) \bar{X}' = 0 \quad \text{for all } \bar{X}' \in \bar{Q}'
\]  (15)

Comparing we get

\[
\bar{Q}' = \mathcal{H}^{-1} Q' \mathcal{H}^{-1}
\]

This on further simplification using equations 14 and 12 gives

\[
\bar{Q}' = \mathcal{H}^{-1} (A^{-1} Q A^{-1}) \mathcal{H}^{-1} = (\mathcal{H} A \mathcal{H}^{-1})^{-1} \bar{Q} (\mathcal{H} A \mathcal{H}^{-1})^{-1}
\]
which is the desired relation 10. Also, it can be shown that \( \mathcal{H}^{-1} = \begin{pmatrix} P^{-1}(I_3 - \frac{p_{11}}{x_i l_i^2 l_i^2}) & \frac{P^{-1} p}{x_i l_i^2 l_i^2} \\ \frac{p_{11}}{x_i l_i^2 l_i^2} & \frac{1}{x_i l_i^2 l_i^2} \end{pmatrix} \).

Then by a series of computations shown in Theorem 3.1, we get

\[
\mathcal{H}A\mathcal{H}^{-1} = \begin{pmatrix}
H_{11} & e_N^t \\
\nu^t_{11N}(I - H_{11}) & 1 - \nu_{11N}^t e_N^t
\end{pmatrix}
\]

(16)

We now reconstruct the first quadric.

**Reconstruction of \( \tilde{Q} = \mathcal{H}(Q) \)**

Define the plane \( \Pi = \angle M_1, M_2, M_3 \). Let \( m_i = (x_i \ y_i \ 1)^t \) and \( m_i^t = (x_i' \ y_i') (1)^t \ i = 1, 2, 3, 4 \) be the corresponding image points of the distinguished points \( M_i, M_i' \ i = 1, 2, 3, 4 \).

Let \( M \) be any point on \( Q \) having an image \( m \). In Theorem 3.1, the reconstructed point \( \mathcal{H}M \approx (x \ y \ 1 \ k)^t \) has been computed. Since \( \tilde{Q} = \mathcal{H}Q \), we have by Theorem 4.1 \( \tilde{C} = \tilde{q}q^t - \tilde{q}_{44}\tilde{q}_{33} \). As commented before, the given outline conic \( C \) of \( Q \) is the outline conic \( \tilde{C} \) of \( \tilde{Q} \). Also \( \tilde{q}_{44} \neq 0 \). Thus \( \tilde{q}_{33} = \frac{\tilde{q}_{33} - \tilde{C}}{\tilde{q}_{44}} = \frac{\tilde{q}_{33} - \tilde{C}}{\tilde{q}_{44}} \) and \( \tilde{Q} = \begin{pmatrix} \tilde{q}_{33} & \tilde{q} \\
\tilde{q}^t & \tilde{q}_{44} \end{pmatrix} \), is known if we can compute \( \tilde{q} \) and \( \tilde{q}_{44} \) consisting of 3 unknowns from \( \tilde{q} \) and 1 from \( \tilde{q}_{44} \). We have \( (m^t k)\tilde{Q}(m^t k)^t = (x \ y \ 1 \ k)^t \tilde{Q}(x \ y \ 1 \ k)^t = 0 \), which on simplification gives the equation \( (m^t \tilde{q} + k\tilde{q}_{44})^2 = m^t C m \). This gives

\[
m^t \tilde{q} + k\tilde{q}_{44} = \pm \sqrt{m^t C m}
\]

(17)

The sign of the right hand side can be chosen unambiguously to be positive (see Remark 4.3 given after the proof). Now \( \tilde{q} \) and \( \tilde{q}_{44} \) can be computed by solving set of 4 equations obtained from (17) by substituting \( m = m_i = (x_i \ y_i \ 1)^t \) and \( k_i \) as computed in Theorem 3.1, \( i = 1, 2, 3, 4 \).

Thus \( \tilde{Q} \) has been reconstructed.

The second quadric \( \tilde{Q}' \) can be reconstructed by exploiting the relationship shown in 10 above.

\( \square \)

In the case of translationally repeated quadrics, as explained in the discussion following Theorem 3.1, \( \mathcal{H}A\mathcal{H}^{-1} \) is image computable. Therefore having reconstructed \( \tilde{Q} \) from the image information, \( \mathcal{H}A\mathcal{H}^{-1} \) being image computable, \( \tilde{Q}' \) is image computable too.
Remark 4.3: The two signs in (17) would give two values of $k$, i.e., $k_1$ and $k_2$. The two points $(x \ y \ 1 \ k_1)^t$ and $(x \ y \ 1 \ k_2)^t$ are the two points of intersection of the ray $< \mathcal{H}_O \ \mathcal{H}_M >$ with the quadric $\tilde{Q}$ where $m = (x \ y \ 1)^t \approx [P \ p]M$. The problem is to pick $k$ unambiguously so as to choose the unoccluded point, which in our case would be the one closer to the transformed $COP_1 = \mathcal{H}_O \approx (0 \ 0 \ 0 \ 1)^t$. By referring to closeness, we are not trying to define the distance between $\mathcal{H}_M$ and $\mathcal{H}_O$, since these are projective points and the concept of distance is not defined between them. Each transformed 3D point is equivalent to $(x \ y \ 1 \ k)^t$, thus we have been able to associate a number $k$ with each 3D point of the ray $< \mathcal{H}_O \ \mathcal{H}_M >$. Since $k$ is the relative affine structure, it can be expressed in terms of actual distances from the plane $\Pi$ by $k = \frac{d_2}{z_4}$ where $d, d_4$ are perpendicular distances of $M$ and $M_4$ (point lying outside the plane $\Pi$) respectively from the plane $\Pi$ and $z, z_4$ are the depths of these points from the origin (proved by Shashua [19] and independently by us). $M_4$ is the point lying outside the plane $\Pi$ formed by the points $M_1, M_2, M_3$. It is the point used to fix the scale of $H_\Pi$ and to normalize $k$. Of the two points $(x, y, 1, k_1)$ and $(x, y, 1, k_2)$, the one to be chosen is to be the one closer to the transformed origin which is characterized by the point with the maximum value of $k$. In order to explain this we need to reinterpret some results of Shashua [20]. Depending on the position of $M_4$ with respect to the COP and $\Pi$, choice of $k$ is made. If $M_4$ lies between COP and $\Pi$, $k = \max(k_1, k_2)$ and if $M_4$ and COP are on opposite sides of $\Pi$ then $k = \min(k_1, k_2)$. We can assume without loss of generality that $\tilde{q}_{44} > 0$, thus $k$ takes the maximum value when $\sqrt{m^tCm}$ is assigned negative sign. By adjusting the scale of $H_\Pi$ it can be shown that all the points on the ray that the points on the ray $< \mathcal{H}_O \ \mathcal{H}_M >$ are such that the value of $k$ decreases as we move away from the origin, with $k = 0$ on the plane $\Pi$ and the point with maximum $k$ is the one closest to the origin. Thus the first point on the quadric which $< \mathcal{H}_O \ \mathcal{H}_M >$ meets, is the one with maximum $k$. Therefore, it can be inferred that a positive sign should be assigned to $\sqrt{m^tCm}$ in the equation (17). Hence the choice of $k$ has been made unambiguous.

As can be seen in Theorem 3.1 and Theorem 4.1, the knowledge of four point correspondences of points in general position is essential for the reconstruction. Therefore it is important to discuss the effect a change of points chosen has on the reconstruction and recognition strategy. By changing the four points, we are changing the plane $\Pi$ and this manifests as a change in $\mathcal{H}$.
In Theorem 4.2, we show that a change in points leads to a pair of reconstructed quadrics which are projectively equivalent to the original pair. Since we use these quadrics to compute joint projective invariants, the values of these invariants remains unchanged for projectively equivalent configurations and hence the recognition strategy remains unaffected.

**Theorem 4.2** Let $Q$ be a quadric and $Q'$ its translate. $M_1, M_2, M_3, M_4$ are the four distinguished points on $Q$ and $M'_1, M'_2, M'_3, M'_4$ are their correspondences on the quadric $Q'$. The reconstructed quadrics are $\tilde{Q}$ and $\tilde{Q}'$ respectively. By changing the distinguished points to $\tilde{M}_1, \tilde{M}_2, \tilde{M}_3, \tilde{M}_4$, the reconstructed quadrics are $\tilde{Q}$ and $\tilde{Q}'$, respectively. These are projectively equivalent to the original pair, that is, there exists a projective transformation $\tilde{H}$ such that $\tilde{Q} = \tilde{H} \tilde{Q} \tilde{H}^{-1}$ and $\tilde{Q}' = \tilde{H} \tilde{Q}' \tilde{H}^{-1}$.

**Proof**: $Q$ and $Q'$ are the translationally repeated quadrics which are to be reconstructed. $M_1, M_2, M_3$ determine the plane $\Pi$ and the image point correspondences together with the knowledge of the epipoles determines the homography $H_{\Pi}$. $M_4$ is used to fix scale of $H_{\Pi}$. The reference plane $\Pi$ is fixed for one reconstruction. Let $M$ be any point on the quadric $Q$, then $\tilde{M} = HM$ is the reconstructed point which lies on the reconstructed quadric $\tilde{Q}$. We have shown that

$$\tilde{Q} = \tilde{H}^{-1}Q\tilde{H}^{-1}$$

Similarly

$$\tilde{Q}' = \tilde{H}^{-1}Q'\tilde{H}^{-1}$$

When the distinguished points are changed to $\tilde{M}_1, \tilde{M}_2, \tilde{M}_3, \tilde{M}_4$, then the homography changes to $\tilde{H}_{\Pi}$ because the plane $\Pi$ changes. Hence the projective reconstruction matrix changes to $\tilde{H}$. The reconstructed quadrics are

$$\tilde{Q} = \tilde{H}^{-1}Q\tilde{H}^{-1}$$

and

$$\tilde{Q}' = \tilde{H}^{-1}Q'\tilde{H}^{-1}$$

Now using equations 19 and 20 and simplifying, we get

$$\tilde{Q} = \tilde{H}^{-1}(\tilde{H}'\tilde{Q}\tilde{H})\tilde{H}^{-1} = (\tilde{H}\tilde{H}^{-1})^t\tilde{Q}(\tilde{H}\tilde{H}^{-1})$$

15
Define a projective transformation \( \tilde{H} = (H \tilde{H}^{-1}) \)
Substituting into relation 22 we get \( \tilde{Q} = \tilde{H} \tilde{Q} \tilde{H} \)
Similarly we can prove \( \tilde{Q}' = \tilde{H} \tilde{Q}' \tilde{H} \)

Therefore any four points in general position can be used for reconstructing projectively equivalent quadrics. This makes the reconstruction scheme independent of the choice of these four points. Having reconstructed the quadrics, we will use these to compute projective invariants, the values of which will be used for recognition. The invariants are computed in the next section.

5 Recognition of Repeated Quadric Configurations

In this section, we propose a recognition strategy for affinely repeated objects. This is applied to the specific case in which each of the components is a quadric. We first compute joint projective invariants for a pair of proper quadrics and use these for recognition of quadric configurations. The nature of the invariants is such that the repeated quadrics with different Euclidean repetitions as well as different nature of quadrics, will be recognized as different configurations. Finally we present the recognition strategy.

5.1 Invariants of a pair of proper quadrics

Let \( Q_1 \) and \( Q_2 \) be a pair of proper quadrics defined up to a scale by \( 4 \times 4 \) non-singular symmetric matrices \( A \) and \( B \) respectively. Let \( T \in PGL_4(\mathbb{R}) \), the group of all \( 4 \times 4 \) non-singular symmetric matrices, be any projective transformation. Define transformed quadrics \( \tilde{Q}_1 = T(Q_1) = \{ TM | M = (X Y Z 1)^t \in Q_1 \} = \{ TM | M^t AM = 0 \} \) and \( \tilde{Q}_2 = T(Q_2) = \{ TM | M^t BM = 0 \} \).

By the Counting Argument [7], the number of projective invariants of the configuration space consisting of a pair of quadrics \( Q_1 \) and \( Q_2 \) are 3. Now \( Q_1 \) and \( Q_2 \) are determined by matrices \( A \) and \( B \) up to a scale, so if the projective invariants are computed with the help of matrices \( A \) and \( B \), then these should be independent of all scalars \( \alpha, \beta \neq 0 \) and all \( T \in PGL_4(\mathbb{R}) \), which is shown in the following theorem.

**Theorem 5.1** If \( Q_1 \) and \( Q_2 \) are the two proper quadrics defined by symmetric non-singular matrices \( A \) and \( B \) respectively up to a scale, then \( \{ \frac{F_1}{F_3}, \frac{F_1 F_2}{F_3}, \frac{F_2 F_3}{F_1} \} \) define a desired set of joint
projective invariants where $I_k, \ k = 1, 2, 3, 4, 5$ are defined by

$$|\lambda A + \mu B| = \lambda^3 I_1 + \lambda^3 \mu I_2 + \lambda^2 \mu^2 I_3 + \lambda \mu^3 I_4 + \mu^4 I_5$$

As a consequence, we have $I_1 = |A|, I_5 = |B|, I_2 = \Sigma_{i=1}^4 a_{ij} B_{ij}, A_{ij}$ is the cofactor of $a_{ij}$ in $|A|, I_4 = \Sigma_{i=1}^4 a_{ij} B_{ij}, B_{ij}$ is the cofactor of $b_{ij}$ in $|B|$, and $I_3$ is the sum of six determinants of order 4 in which any two columns are from $A$ and any other two columns are from $B$.

Proof : Suppose we apply a projective transformation $T \in PGL_4(\mathbb{R})$ and get the transformed quadrics $\overline{Q}_1$ and $\overline{Q}_2$. These are completely determined upto a scale by the matrices $\overline{A} = (T^{-1})^t A T^{-1}$ and $\overline{B} = (T^{-1})^t B T^{-1}$, we get $|\lambda \overline{A} + \mu \overline{B}| = |T^{-1}|^2 |\lambda A + \mu B|$. Thus if $|\lambda \overline{A} + \mu \overline{B}| = \lambda^4 \overline{T}_1 + \lambda^3 \mu \overline{T}_2 + \lambda^2 \mu^2 \overline{T}_3 + \lambda \mu^3 \overline{T}_4 + \mu^4 \overline{T}_5$, then $\overline{T}_k = |T^{-1}|^2 I_k$ for $k = 1, 2, 3, 4, 5$.

So $\frac{\overline{T}_1}{I_1} = \frac{I_1}{I_1}, \frac{\overline{T}_2}{I_2} = \frac{I_2}{I_2}, \frac{\overline{T}_3}{I_3} = \frac{I_3}{I_3}, \frac{\overline{T}_4}{I_4} = \frac{I_4}{I_4}, \frac{\overline{T}_5}{I_5} = \frac{I_5}{I_5}$.

Also if we replace $A$ by $\alpha A$ and $B$ by $\beta B$, then

$$|\lambda \alpha A + \mu \beta B| = \lambda^4 (\alpha^4 I_1) + \lambda^3 \mu (\alpha^3 \beta I_2) + \lambda^2 \mu^2 (\alpha^2 \beta^2 I_3) + \lambda \mu^3 (\alpha \beta^3 I_4) + \mu^4 (\beta^4 I_5)$$

Now if we replace $I_1$ by $\alpha^4 I_1, I_2$ by $\alpha^3 \beta I_2, I_3$ by $\alpha^2 \beta^2 I_3, I_4$ by $\alpha \beta^3 I_4$ and $I_5$ by $\beta^4 I_5$, then also the proposed set of invariants remain unchanged. This proves the theorem. ☐

As $I_j, j = 1, 2, 3, 4, 5$ are defined using the matrices of the reconstructed quadrics, in the case when the quadrics are image computable, the invariants are also image computable. This is true for the case of translational repetition. We now present an overview of the recognition strategy.

### 5.2 Recognition Framework

In this section, we present the recognition framework for configurations containing a pair of repeated quadrics. The framework is general and applicable to all objects with repeated components, for which invariants can be computed. We are given a single image of a configuration with translationally repeated quadric components. The outline conic has been fitted to a quadric and it is known that its repetition exists in the image.

Now we pick four point correspondences on the quadric and its repetition. Using four point correspondences we first compute the homography $H_{\Pi}$ [20]. The epipoles are computed using a pair of point correspondences, in the case of translational repetition. In the case of a general repetition, it is computed through the fundamental matrix which requires eight or more point
correspondences [11, 12]. The four points are reconstructed, that is, their affine structure \( k \) is computed as in section 3.1. Then using the reconstructed points and the outline conic, we reconstruct the first quadric \( \tilde{Q} \) (Theorem 4.1). By exploiting the relationship between the repeated components, derived in Theorem 3.1, we reconstruct the repeated quadric component \( \tilde{Q}' \) (Theorem 4.1). Thus the components have been reconstructed using image information alone, with respect to the same frame. Having identified the outline conic corresponding to the first quadric in the image, the only additional information required for reconstruction, is that a repetition is present in the image.

The components of the matrices of the reconstructed quadrics are used to compute the invariants \( \{ \frac{I_k}{I_5}, \frac{I_4}{I_5}, \frac{I_2}{I_5} \} \) (Theorem 5.1). The \( I_k, k = 1, 2, 3, 4, 5 \) are computed using the components of the quadrics. Any other set of independent invariants may also be used. In Theorem 4.2, we have shown that an alternate choice of four points leads to a projectively equivalent configuration. The values of invariants do not change for projectively equivalent configurations and this helps in generalizing the strategy.

The joint projective invariants computed above are applicable to all pairs of quadrics irrespective of the repetition. These values are used for recognition purposes and are stored for image of each model scene. The same procedure is then applied to a new test image. The values of the invariants corresponding to the test image are used to index into the model base. The distance between the invariants of the test image and those stored in the model base are computed. The test image is identified as that of a model scene if its distance from its nearest neighbour in the model base is less than a threshold.

This recognition strategy is experimentally investigated using real and experimental scenes which are discussed in the next section.

6 Experimental Analysis

The experiments have been carried out on real life images of monuments with multi-dome architecture (Fig. 6 - 8) and experimental scenes (Fig. 2 - 5) of objects with translationally repeated quadric components. In this case, as has been explained the infinite homography is the identity
matrix and hence the reconstruction is completely image computable. The experiments were undertaken in order to study the applicability of these invariants to the recognition problem. The discriminatory power and the stability of these invariants have also been investigated. Results of these experiments establish the effectiveness of the joint projective invariants computed, for the recognition of quadric configurations consisting of a pair of translationally repeated quadrics.

For all images, the conics were fitted interactively by using Bookstein’s algorithm [1]. The fitted conic for some of the scenes are shown in Fig. 1. The correspondences between known points on the translationally repeated objects were used. The processing routines were developed using MATLAB. In the following, we present the results of experiments performed for studying discriminatory power and stability of the invariants.

6.0.1 Experimental Scenes

Now the framework of quadric reconstruction has been applied to all the images Fig 2 - 4. These reconstructed quadrics have been used to compute three independent invariants \( \{ \frac{I_1 I_2}{I_3^2}, \frac{I_1 I_3}{I_2^2}, \frac{I_2^2}{I_3^2} \} \) for each of the images. The values have been given in Table 1. In order to compare values of invariants for all pairs of different images \( im_i, im_j \), we use the distance [15] given by \( \Sigma_{k=1}^{3} (V_k[im_i] - V_k[im_j])^2 \) where \( V_k[im_i] \) is the value of the kth invariant of image i. It can be

![Figures 1: Conics fitted to repeated domes of monuments. Information of one of the conics is sufficient for reconstruction. For our experiments we pick the left conic](image)
observed that the distances between invariants of Table 1 shown in Table 2, indicate the discriminatory power of the invariants.

The distances between the projectively equivalent configurations of vases in Fig. 2 (a), 2 (b) is 0.1148. This is much less as compared to the distance of Fig. 2 (a) from the other quadric configurations of, say, the bottles (Fig. 4 (a), 4 (b)) which are 5.4498e3 and 5.0647e3 respectively.

The invariants also have the capability of distinguishing between configurations on the basis of translation between them. Configuration in Fig. 4 (c) differs from those in Fig. 4 (a), 4 (b) in terms of the translation between the bottles, indicated by the distances 5.0776e3 between Fig. 4 (a), 4 (c), 4.7061e3 between Fig. 4 (b), 4 (c) and 7.0543 between projectively equivalent configurations in Fig. 4 (a), 4 (b). It is also evident that the recognition strategy can handle occlusion. Inspite of the second component being occluded in Fig. 5 (b), (c), the distance between Fig. 5 (a), 5 (b) (see Table 2) is much less compared to the distance between Fig. 5 (b)
Table 1: Values of invariants of experimental scenes

<table>
<thead>
<tr>
<th>Fig. No.</th>
<th>$I_1 I_5 / I_2^2$</th>
<th>$I_1 I_5 / I_2 I_4$</th>
<th>$I_2^2 / I_3 I_5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 (a)</td>
<td>9.4947e-7</td>
<td>3.5618e-6</td>
<td>-5.5926e3</td>
</tr>
<tr>
<td>2 (b)</td>
<td>9.0259e-7</td>
<td>2.1719e-6</td>
<td>-2.2044e3</td>
</tr>
<tr>
<td>3 (a)</td>
<td>2.2409e-10</td>
<td>1.3849e-9</td>
<td>-3.3533e6</td>
</tr>
<tr>
<td>3 (b)</td>
<td>1.3409e-10</td>
<td>1.4025e-9</td>
<td>-3.0445e6</td>
</tr>
<tr>
<td>4 (a)</td>
<td>7.6457e-9</td>
<td>3.4706e-8</td>
<td>-4.4566e4</td>
</tr>
<tr>
<td>4 (b)</td>
<td>3.7948e-9</td>
<td>1.5599e-8</td>
<td>-3.2749e4</td>
</tr>
<tr>
<td>4 (c)</td>
<td>6.8598e-9</td>
<td>4.9711e-8</td>
<td>-3.1247e5</td>
</tr>
<tr>
<td>5 (a)</td>
<td>7.121e-9</td>
<td>5.112e-8</td>
<td>-4.6454e5</td>
</tr>
<tr>
<td>5 (b)</td>
<td>1.3235e-9</td>
<td>2.9669e-8</td>
<td>-4.3797e5</td>
</tr>
<tr>
<td>5 (c)</td>
<td>6.0086e-9</td>
<td>4.9281e-8</td>
<td>-3.4687e5</td>
</tr>
</tbody>
</table>

Table 2: Distance (scaled) between sets of invariants in Table 1

<table>
<thead>
<tr>
<th>Fig. No.</th>
<th>2 (a)</th>
<th>2 (b)</th>
<th>3 (a)</th>
<th>3 (b)</th>
<th>4 (a)</th>
<th>4 (b)</th>
<th>4 (c)</th>
<th>2 (a)</th>
<th>5 (b)</th>
<th>5 (c)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 (a)</td>
<td>0</td>
<td>0.0011</td>
<td>1.1207e3</td>
<td>9.25.4953</td>
<td>0.519</td>
<td>0.6279</td>
<td>9.4174</td>
<td>21.0635</td>
<td>11.6470</td>
<td>18.4950</td>
</tr>
<tr>
<td>2 (b)</td>
<td>0.0011</td>
<td>0</td>
<td>1.12386</td>
<td>9.25.5103</td>
<td>0.705</td>
<td>0.6033</td>
<td>9.9676</td>
<td>21.3714</td>
<td>18.9922</td>
<td>11.8794</td>
</tr>
<tr>
<td>3 (a)</td>
<td>1.20763</td>
<td>1.12363</td>
<td>0</td>
<td>9.5357</td>
<td>1.0948e3</td>
<td>1.02993</td>
<td>9.25.6947</td>
<td>9.25.8034</td>
<td>18.9819</td>
<td>9.9821</td>
</tr>
<tr>
<td>4 (a)</td>
<td>0.25.89</td>
<td>0.1706</td>
<td>1.0948e3</td>
<td>9.99.9604</td>
<td>0</td>
<td>0.04.10</td>
<td>9.773</td>
<td>15.4757</td>
<td>9.1388</td>
<td></td>
</tr>
<tr>
<td>4 (b)</td>
<td>0.0737</td>
<td>0.0033</td>
<td>1.02993</td>
<td>9.97.644</td>
<td>0.0140</td>
<td>0</td>
<td>7.8244</td>
<td>18.6443</td>
<td>16.8294</td>
<td>9.9872</td>
</tr>
<tr>
<td>5 (a)</td>
<td>21.0635</td>
<td>21.3714</td>
<td>8.34.694</td>
<td>9.65.694</td>
<td>17.6378</td>
<td>18.6443</td>
<td>2.325</td>
<td>0</td>
<td>0.0706</td>
<td>1.3846</td>
</tr>
<tr>
<td>5 (b)</td>
<td>11.6470</td>
<td>18.9819</td>
<td>8.48.9149</td>
<td>6.79.3999</td>
<td>15.4767</td>
<td>16.4294</td>
<td>1.575</td>
<td>0.0706</td>
<td>0</td>
<td>0.8299</td>
</tr>
<tr>
<td>5 (c)</td>
<td>16.0950</td>
<td>11.4714</td>
<td>9.98.2021</td>
<td>7.72.008</td>
<td>9.3848</td>
<td>9.8672</td>
<td>0.1183</td>
<td>1.3846</td>
<td>0.8299</td>
<td>0</td>
</tr>
</tbody>
</table>
Figure 4: (a,b) Two views of translationally repeated bottles (c) Bottles separated by a different translation

Figure 5: (a) View of translationally repeated bottles (b,c) one bottle occluded and Fig. 2 (a), which are not projectively equivalent.

6.1 Application: Images of Monuments

As an application, the experiments were carried out for recognition of images of monuments with multi-dome architecture. This example indicates applicability of the approach for content based image retrieval.

The framework is then applied to monuments in Fig. 6 - 8, which can be characterized by translationally repeated domes. Each of these domes is modeled as a quadric. Due to the distance from which the photograph has to be taken in order to get the two domes in view, the projective distortion is not too much. As a result, the outline conics seen in the image do not vary much, which aid in recognition. Also these monuments have parapets around the domes, or engravings on the dome, which aid in the choice of corresponding points. The perfect symmetry of these
domes allow a wide range of views to be acceptable.

In each of the images of the monuments, the outline conics were fitted to the domes, as shown in Fig. 1. By using any four point correspondences of points in general position on the domes and the outline conic of the first dome, the quadric with respect to the first dome is reconstructed. Then using Theorem 4.1, the second (translated) dome is reconstructed. These two quadrics are used to compute the set of invariants given above. The values of invariants for the images of the monuments are given in Table 3. The distances between the sets of invariants with respect to the images are given in the Table 4. The distance between the set of invariants between two images of Jama Masjid (Fig. 7 (a), (b)) 0.0053, which is much less than the distance between the image of Taj Mahal (Fig. 6 (a)) and Jama Masjid (Fig. 7 (b)) which is 107.7481. This indicates that the procedure is able to distinguish between the images of Jama Masjid (Fig. 7) from the images of Taj Mahal (Fig. 6), even when one of the domes of the Jama Masjid is occluded. Also, the two
Figure 8: Views of the (a) Red Fort, Delhi and (b) Birbal’s Tomb, Agra

Table 3: Values of invariants for images of Monuments

<table>
<thead>
<tr>
<th>Fig. No.</th>
<th>$I_1I_5/I_3^2$</th>
<th>$I_1I_5/I_2I_4$</th>
<th>$I_2^2/I_3I_5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>6 (a)</td>
<td>2.1838e-12</td>
<td>-1.2059e-8</td>
<td>-1.7507e3</td>
</tr>
<tr>
<td>6 (b)</td>
<td>7.8844e-12</td>
<td>-6.773e-9</td>
<td>-1.6525e3</td>
</tr>
<tr>
<td>6 (c)</td>
<td>3.9838e-12</td>
<td>-1.0648e-8</td>
<td>-1.9093e3</td>
</tr>
<tr>
<td>6 (d)</td>
<td>1.3312e-12</td>
<td>-3.5861e-8</td>
<td>-2.01e3</td>
</tr>
<tr>
<td>7 (a)</td>
<td>4.189e-10</td>
<td>1.5877e-8</td>
<td>-5.0102e3</td>
</tr>
<tr>
<td>7 (b)</td>
<td>4.0497e-10</td>
<td>1.3656e-7</td>
<td>-5.0332e3</td>
</tr>
<tr>
<td>8 (a)</td>
<td>4.3603e-10</td>
<td>2.4327e-8</td>
<td>-5.0332e3</td>
</tr>
<tr>
<td>8 (b)</td>
<td>1.8401e-6</td>
<td>6.92e-7</td>
<td>2.2265e4</td>
</tr>
</tbody>
</table>

images of Jama Maśjid are identified as being two images of the same monument, inspite of the occlusion. This highlights the robustness to occlusion of this framework. This is an important characteristic of the approach because none of the projective invariants proposed for 3D objects are tolerant to occlusion.

Next we study the stability of the recognition framework, by doing a Principal component analysis.
Table 4: Distance (scaled) between invariants in Table 3

<table>
<thead>
<tr>
<th>Fig. No.</th>
<th>6 (a)</th>
<th>6 (b)</th>
<th>6 (c)</th>
<th>6 (d)</th>
</tr>
</thead>
<tbody>
<tr>
<td>6 (a)</td>
<td>0</td>
<td>0.0962</td>
<td>0.2515</td>
<td>0.6724</td>
</tr>
<tr>
<td>6 (b)</td>
<td>0.0962</td>
<td>0</td>
<td>0.6595</td>
<td>1.2781</td>
</tr>
<tr>
<td>6 (c)</td>
<td>0.2515</td>
<td>0.6595</td>
<td>0</td>
<td>0.1014</td>
</tr>
<tr>
<td>6 (d)</td>
<td>0.6724</td>
<td>1.2781</td>
<td>0.1014</td>
<td>0</td>
</tr>
<tr>
<td>7 (a)</td>
<td>106.2434</td>
<td>112.7415</td>
<td>96.1558</td>
<td>90.0120</td>
</tr>
<tr>
<td>7 (b)</td>
<td>107.7481</td>
<td>114.2913</td>
<td>97.5875</td>
<td>91.3974</td>
</tr>
<tr>
<td>8 (a)</td>
<td>2.6312e4</td>
<td>2.6413e4</td>
<td>2.615e4</td>
<td>9.1397e4</td>
</tr>
<tr>
<td>8 (b)</td>
<td>5.7675e3</td>
<td>5.7205e3</td>
<td>5.844e3</td>
<td>5.8928e3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Fig. No.</th>
<th>7 (a)</th>
<th>7 (b)</th>
<th>8 (a)</th>
<th>8 (b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>7 (a)</td>
<td>0</td>
<td>0.0053</td>
<td>2.3074e4</td>
<td>2.3052e4</td>
</tr>
<tr>
<td>7 (b)</td>
<td>0.0053</td>
<td>0</td>
<td>2.3052e4</td>
<td>7.4519e3</td>
</tr>
<tr>
<td>8 (a)</td>
<td>2.3074e4</td>
<td>2.3052e4</td>
<td>0</td>
<td>5.6717e4</td>
</tr>
<tr>
<td>8 (b)</td>
<td>2.3052e4</td>
<td>7.4519e3</td>
<td>5.6717e4</td>
<td>0</td>
</tr>
</tbody>
</table>

6.1.1 Principal Component Analysis

We consider multiple views of three configurations each containing a pair of translationally repeated quadrics. Fig. 9 (a) - (d), 10 (a) - (d), 11 (a) - (d) are the three test sequences. For each of these, we reconstruct the pair of quadrics as explained above and compute the values of 7 invariants \( \{ \frac{h_{13}}{r_3^2}, \frac{h_{14}}{r_3^2}, \frac{h_{24}}{r_2^2}, \frac{h_{15}}{r_3^4}, \frac{h_{16}}{r_3^4}, \frac{h_{25}}{r_2^4}, \frac{h_{26}}{r_2^4} \} \). We get twelve \( 7 \times 1 \) vectors denoted by \( X_i, \ i = 1, \ldots, 12 \) and is called an invariant vector or vector of invariants, one vector with respect to each image (Fig. 9 - 11). We then carry out a Principal Component Analysis on these Invariant vectors. As the property of the analysis, eigenvectors with larger eigen values provide more information on the invariants variation than those with smaller eigen values. As has been shown, a configuration consisting two translationally repeated quadrics has three independent invariants. Thus we choose the three largest eigen values (the others can be
discarded in comparison to these). We categorize the images into three classes on the basis of their projective equivalence. As can be seen, the classes are clearly distinguished with images 9 (a) - (d) forming Class I, Fig. 10 (a) - (d) forming the class II and Fig. 11 (a) - (d) forming the Class III. In order to maximise the class separation, we use Mahalanobis distance [5]. The distances of each of the class members from their respective classes, as well as from the other classes is shown in Table 5.

We consider the two test images of Fig. 12 and study their categorization into classes when the Principal Component Analysis is applied. The distance of the reduced image vector pertaining to Fig. 12 (a) from class I is 1.3895e13, from Class II is 45.7922 and from Class III is 6.242e7, which correctly classifies image Fig. 12 (a) as belonging to class II. Similarly the distance of the reduced image vector of Fig. 12 (b) from Class I is 9.1642e5, from Class II is 9.3026e7 and from 18.2887 from Class III. Therefore the image 12 (b) is correctly classified to be that of a configuration corresponding to Class III.

These results show that we can organize the object models in terms of multiple exemplars forming individual classes corresponding to each model and use PCA for the representation of the model objects. For a new image, the invariants can be computed and the reduced image vector

---

Figure 9: Views of Scene 1 for stability analysis

Figure 10: Views of Scene 2 for stability analysis
can be used to obtain its Mahalanobis distance from the classes, for the purpose of recognition. It is clear from the results presented here that despite variations in the invariant value, object models can be subdivided into well separated clusters in the eigen space and hence we can obtain stable recognition results by obtaining Mahalanobis distance with respect to individual classes.

7 Conclusions

In this paper, we have proposed an invariant based recognition strategy for repeated objects, which takes into account the repetition explicitly. The projective invariants proposed by us are computed using these reconstructed components. The framework can handle partial occlusion of one of the repeated components. As in the case of translationally repeated quadrics, the only image information required are four point correspondences between the repeated components and the outline conic of the first quadric only, hence it permits partial occlusion of the repeated component. The choice of four points is not crucial to the recognition strategy as an alternate
choice leads to a projectively equivalent reconstruction for which the values of the projective invariants are the same. The recognition strategy is general and applicable to all objects for which invariants can be computed. The discriminatory ability of the invariants have been established experimentally. The stability of the recognition strategy has been illustrated through a principal component analysis. This recognition strategy can be applied for recognition of images of historical monuments with multi-dome structure.

References


