Robust Homography-Based Control for Camera Positioning in Piecewise Planar Environments

D. Santosh Kumar and C.V. Jawahar

Center for Visual Information Technology
International Institute of Information Technology
Hyderabad 500032, India
{santosh@students., jawahar@}iiit.ac.in

Abstract. This paper presents a vision-based control for positioning a camera with respect to an unknown piecewise planar object. We introduce a novel homography-based approach that integrates information from multiple homographies to reliably estimate the relative displacement of the camera. This approach is robust to image measurement errors and provides a stable estimate of the camera motion that is free from degeneracies in the task space. We also develop a new control formulation that meets the contradictory requirements of producing a decoupled camera trajectory and ensuring object visibility by only utilizing the homography relating the two views. Experimental results validate the efficiency and robustness of our approach and demonstrate its applicability.

1   Robotic Vision

The use of computer vision techniques to control robotic systems has received great popularity in recent times [1]. Images captured by cameras attached to a robot provide ample information about its surroundings that assists it in efficiently navigating the environment. This field, known as *Visual Servoing* [2], has gained recent prominence due to the widespread availability of high quality cameras and low cost microprocessors. In addition to robotics, visual servoing algorithms also find interesting applications for interactive vision systems such as video conferencing, tracking, active vision, augmented reality etc. The visual feedback increases the accuracy of the overall vision system and relaxes the requirement of high precision accessories.

Many servoing techniques have been proposed and extensively studied in literature. In [3], optical flow is used to control the pose of the camera in conjunction with a Jacobian-based adaptive controller. In [4], $3D$ object pose is estimated and utilized to regulate the camera pose error. The class of algorithms similar to the former method constitute the popular *Image-based* Visual Servoing techniques while the latter pertain to *Position-based* approaches. For the relative merits and demerits of the above techniques, the reader may refer to [2]. Recently, a new group of algorithms have been proposed [5,6,7] that exploit a combination of the above methods to estimate the camera displacement between the desired and the current pose. They combine the traditional Jacobian-based control with

other techniques to form the class of *Hybrid Visual Servoing* algorithms. These methods yield a decoupled, straight-line camera trajectory and possess a large singularity-free task space.

Hybrid algorithms can essentially be classified into two primary categories. Algorithms in the first category are generally based on the computation of the essential matrix relating the two camera views [7,8]. Although the relative camera displacement can be obtained even for unknown (non-planar) scenes, a problem with epipolar geometry is that, it degenerates in certain critical cases (for example, when the target is planar or when the relative displacement is a pure rotation) and hence is not suitable for servoing. Note that a positioning task is accomplished only when the current and the desired images of the scene are similar, which corresponds to the degenerate case. The second class of algorithms determine the relative camera displacement by computing the homography induced by a scene plane relating the two views. However, a major drawback of these methods is the implicit assumption of the planarity of the scene, which prevents their application to real world scenarios as the world is often made up of non-planar regions. It must be emphasized that in either cases, the degeneracies critically affect the convergence and predictability of the system. Thus dealing with such degeneracies is of vital importance in the design of a stable system.

In summary, the desirable characteristics of a *hybrid* visual-control algorithm are

- – Absence of degeneracies in its task space
- – Applicability to both planar and non-planar environments
- – Robustness to image measurement errors
- – Continuity in velocity instruction and smooth convergence behavior
- – Independence from prior knowledge of the object model and initialization of parameters

In this paper, we propose a new homography-based servoing algorithm that achieves the above features. Our method integrates homographies induced by multiple scene planes using geometric and subspace constraints to efficiently estimate the motion and structure parameters (Fig. 1). Another contribution of this paper is the development of a modified control law that provides the
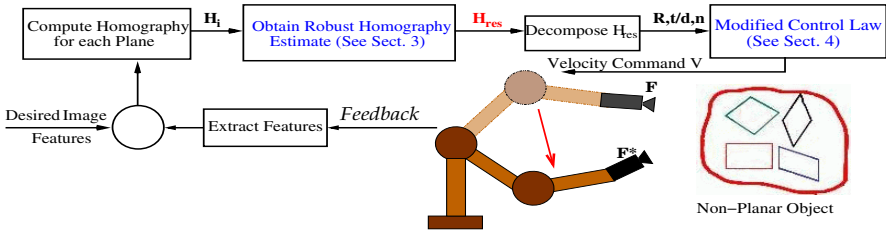

Fig. 1. Visual-feedback control: Multiple homographies are integrated to obtain a robust homography, which is used in the modified control law to gain superior performance

complementary characteristics of producing a decoupled camera trajectory and ensuring object visibility by only using the homography transformation relating the two camera poses.

2 Homography-Based Visual Control

A visual servo control compares the current image of a target with the desired image and the difference (or 'error') is used to drive the camera towards the goal position. Often the task is not just to regulate the image error but also to ensure a realizable camera trajectory. In such scenarios, homography-based control acts as a convenient option as it regulates the error in camera pose by estimating the $3D$ motion parameters only using image information.

If all the object points lie on a $3D$ plane, their coordinates in the current image I and the goal image I^* are related by a 'collineation' [9]. Assume that a point P lies on a plane whose normal vector is n as shown in Fig. 2. The point
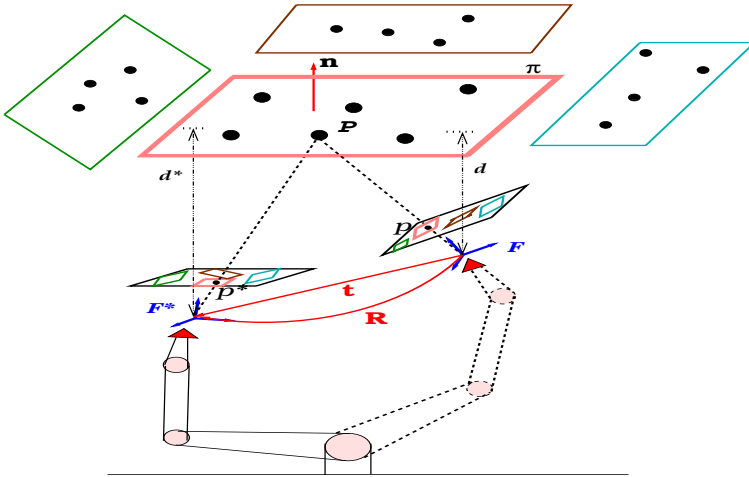

Fig. 2. Homography-based Visual Servoing

expressed in current camera frame $\mathcal{F}$ is related to goal camera frame $\mathcal{F}^*$ by a rotation matrix R and translation vector t as

$$P^* = R\,P + t = (R + t\frac{n^T}{d})P, \tag{1}$$

where $d = n^T P$ is the distance of the plane π from the current camera center. Assuming the camera intrinsic parameters are known, the image coordinates of the $3D$ points are given by $p = \frac{P}{Z}$ and $p^* = \frac{P^*}{Z^*}$ respectively. This transforms (1) to

$$\frac{Z^*}{Z}p^* = (R + t\frac{n^T}{d})p, \tag{2}$$

which can be rewritten as $\alpha p^* = Hp$ where $H_{3\times3} = R + t\frac{n^T}{d}$ is called the 'homography' matrix up to a scale factor α [9].

The recovered homography can be decomposed to obtain the rotation matrix R, the scaled translation vector $\frac{t}{d}$ and the plane normal n using the procedure described in [10]. Unfortunately, in the most general case the decomposition of H yields four different solutions (two of them being the 'opposites' of the other). They can be reduced to two solutions by applying the visibility constraint (*i.e.,* all the features must lie within the camera field of view). Further ambiguity can be resolved by decomposing an additional homography induced by another scene plane. Two pairs of solutions (S_1, S_2) and (S'_1, S'_2) are obtained respectively and a compatible pair (S_i, S'_j) among them is found, *i.e.,* a pair with common motion $(R, \frac{t}{d})$. In general, there is only one compatible pair, and hence the unique solution can be obtained. Thus using information from multiple planes, H can be decomposed unambiguously to obtain the motion and structure parameters. These parameters are used in the control law to generate the optimal velocity instruction.

2.1 Degenerate Configurations and the Use of Multiple Planes

Some of the limitations of the existing *hybrid* techniques to estimate the relative camera displacement were reviewed in Sect. 1. Recently, another method was proposed by Malis *et al.* [6] to compute the relative orientation between the two camera views for a non-planar object using the concept of 'virtual parallax' [11]. By defining a plane using three arbitrary points on the object, they estimate the homography using this virtual plane and perform the positioning task.

A single homography estimate is not sufficient when a camera has to undergo large displacements in visual servoing as the control can be affected by degenerate configurations. Degeneracies in the task space can result either due to occlusion of the feature points, the camera center approaching the world (virtual) plane, the camera centers and the feature points arriving in a singular configuration [9] or due to singular homographies. In either of the cases, when a degeneracy is reached, the plane in consideration is switched *i.e.,* the points used to define the virtual plane are changed and a new plane using three different points is defined. This switching causes a discontinuity in the velocity command and leads to the instability of the control system. In Fig. 3, the effect of switching is demonstrated, where a positioning task with respect to a piecewise planar object was studied.

The other drawbacks in defining a non-planar object using arbitrary planes include

- **Unfavorable for planar scenes.** The methods using virtual parallax are theoretically inefficient to deal with planar objects as the epipolar geometry degenerates in this case [6].
- **Initialization of plane parameters.** In order to resolve the ambiguity in homography decomposition, *a priori* information about the normal vector of the virtual plane is required.
- **Assumption of point features.** Point correspondences are not available in many practical situations or could be noisy. Since the virtual plane is defined
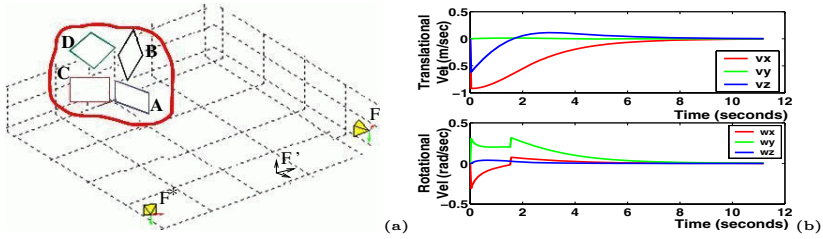
Fig. 3. Velocity Screw using virtual parallax algorithm: (a) Servoing begun using plane A reaches a degeneracy at F' whose origin intersects the plane (b) Discontinuity in the velocity screw is due to the switching of planes (A to C) at F'

  explicitly using the non-coplanar points on the object, these methods may not be applicable when such features are not available.

– **Effect of measurement errors.** Homography estimation is affected due to measurement errors (*'drift'*) in the correspondences. By choosing a different set of points (that are error-free) to define the virtual plane, one can obtain better results.

  It must be emphasized that the above limitations are caused by the fact that only information from a single plane is being utilized to perform the positioning task. The bottleneck has been the fact that there exists no single homography relating the two camera views that can be absolutely relied upon. Nevertheless, by selectively exploiting the information available from multiple planes, one can avoid the above drawbacks and achieve superior performance.

3 Homography Estimation Using Multiple Planes

The objective of the servoing task is to drive the disparity between the current and the desired camera configurations to zero. The homographies relating the two camera poses induced by different planar regions are used to guide the positioning task.

  Our approach proceeds initially by partially tessellating the non-planar scene into piecewise planar patches. This is done by a simple partitioning of the image features into homogeneous planar regions (See Fig. 6(a)). Interest regions are detected and the regions subject to planarity constraint form a set of matching regions [12]. The seed regions act as a 'driver' to guide the evolution of planar patches in the image. Any interest region detector with the ability to detect robust and stable regions can be employed here. For each pair of matching regions, a plane-induced homography is calculated.

  Even though a single homography is sufficient to determine the motion parameters (*rigidity constraint*), information from multiple homographies can be combined to obtain a reliable estimate of the camera displacement. However, to avoid the estimation of multiple homographies at each instant, the constraints on homographies can be exploited to reduce the computations. Recall from (2)

that any H induced by a $3D$ scene plane is described by $H_{3\times3} \approx R + tn^T$. Given a homography matrix H_π induced by some $3D$ plane π, all other homographies H_i can be described as $\lambda_i H_\pi + tn_i^T$ for a fixed pair of cameras [9]. This observation results from the fact that all the homographies differ only in their scale λ_i and plane n_i parameters. Consider k homography matrices $H_1, H_2, \ldots, H_k$, each expressed as a column vector in a $9 \times k$ matrix. The rank of this matrix is known to be utmost four [13]. Hence the space of all homographies between two fixed camera views is embedded in a 4-dimensional linear subspace of $\Re^9$. This observation follows the fundamental fact that multiple planar patches in the scene share the common global camera geometry (*i.e., R, t*).

3.1   Computation of the Reliable Homography

Given the rank-4 constraint, any new homography can be computed as a weighted combination of four linearly independent homographies. The four homographies are in general selected such that they are induced by planes that possess largest area and best visibility (if the centroid of the features in a planar region is within a threshold distance from the nearest image boundary, then it satisfies the visibility constraint) since they are the most reliable.

The resultant homography H_{res} is defined as

$$H_{res} = \lambda_1 H_1 + \lambda_2 H_2 + \lambda_3 H_3 + \lambda_4 H_4, \qquad (3)$$

where the weights λ_i are assigned such that good homographies receive higher weights while the degenerate or errored estimates are given low priority. By appropriately choosing the λ_i's, a reliable homography can be deduced. Recall that, in general, any homography in the subspace can be expressed as a linear combination of four base homographies. In our case, H_{res} is one such 'valid' homography possessing certain desired characteristics.

The principle behind the weight assignment is to prefer valid homographies and reject singular ones in order to prevent abrupt switching of planes during a degeneracy. It must be emphasized that most of the degeneracies are not arbitrary changes and in general, can be predicted in advance. For instance, distance between a camera and a (virtual) plane gradually regresses to zero. Likewise, occlusion of planes can be anticipated by the persistent decrease in area of the planar region (or the number of features). Other degenerate cases can also be predicted in a similar manner and thus homographies that are likely to confront a degenerate configuration can be rejected.

Assignment of weights. Let us define the constraints to assign the weights and hence the parameter λ_i that is used in the computation of H_{res}.

– **Re-projection Error.** This constraint measures the accuracy of the estimated homography. A high error in re-projection indicates a poor estimate and such H should receive less weight as parameters obtained from it will be unreliable. Thus the weights are set inversely proportional to the re-projection error. This ensures that planar regions that are affected

by the cumulative tracking errors (*'drift error'*) are avoided and thereby guaranteeing the robustness of H_{res} to image measurement errors. The exact weight λ_i^e is defined by first calculating the re-projection error *i.e.,* $e = \sum_k d(p_k^*, Hp_k) = \sum_k \|\frac{p_k^*}{\|p_k^*\|} - \frac{H*p_k}{\|H*p_k\|}\|$ and then assigning it using a one-sided Normal distribution $N(e_{thres}, \sigma_e)$ where e_{thres} is the tolerable re-projection error and σ_e is the variance.

– **Homography Determinant.** This quantity signifies the 'goodness' of a homography estimate. If the determinant is tending toward zero, it suggests the arrival of a degeneracy and hence such a homography should acquire low weight. Therefore the weights are set directly proportional to the value of the determinant D. This constraint ascertains the resultant homography to be free of singularities. Here again, the weights λ_i^D are set using a one-sided Normal density function $N(D_{thres}, \sigma_D)$ where D_{thres} is the minimum acceptable determinant.

– **Area of the Plane.** Occlusion of a plane can be detected by measuring the gradient of the plane area dA. If the area of the planar region decreases drastically, then it indicates a possible occlusion of this plane in the near future. Thus the λ_i's are to be set inversely proportional to the value of dA. More precisely, the weight λ_i^{dA} is set using a one-sided Normal distribution $N(dA_{thres}, \sigma_{dA})$ where dA_{thres} is the minimum acceptable gradient.

These weights are normalized and summed together to obtain the resultant weight λ_i. The final expression for H_{res} is calculated as

$$H_{res} = \sum_{i=1}^4 \lambda_i H_i, \quad \text{where} \quad \sum_i \lambda_i = 1.$$

Hence a judicious assignment of weights using the above constraints helps in deducing a 'virtual' homography with the desirable characteristics. A change of bases might be required in case one of the H_i degenerate. However, the degenerate homography would automatically procure a low λ value and its replacement does not affect the stability of the system. This approach is applicable even if the scene consists of less than four planar regions. In such a case, the unavailable homographies in (3) acquire zero weight. It must be emphasized that the method utilizes additional homographies to obtain a reliable homography estimate rather than computing the optimal estimate. The parameters obtained from decomposition of H_{res} are used in the modified control law to compute the camera trajectory.

4   Modified Control Design

Given the stable estimate of the motion and structure parameters, our focus is to design a robust control that not only produces a decoupled camera trajectory but also guarantees feature visibility. Classical approaches such as the $3D$ control algorithms compute an optimal camera trajectory but very often violate

the visibility criteria. $2D$ controls ensure the features to remain in the camera field of view, although they suffer from non-optimal trajectory, computational complexity of calculating the Jacobian pseudo-inverse and the demand for $3D$ depth estimates. Note that providing the contradictory requirements of either controls poses a daunting challenge in the design of an optimum control scheme. Though a few attempts in this direction have been made [5,6,14], the devised controls do not satisfy all the above requirements.

Much of the information that is required for performing the positioning task is readily available from the homography transformation. The presence of multiple planes in the scene further compliments this fact. We exploit this result to fulfill the requirements of the desired optimal control.

Proposed Control. We first introduce the Cartesian ($3D$) control law and then proceed to derive the robust control. Given the parameters obtained from homography decomposition, the translational velocity to go directly to the goal is determined as $-\lambda_v(\frac{t}{d})\,d$, where λ_v is a gain factor and d is the distance to the plane (See Fig. 2). The rotational velocity is computed as $-\lambda_\omega u\theta$, where λ_ω is again a gain factor and u, θ denote the rotation axis and angle that are obtained using the Rodriguez formula for the rotation matrix R as $\theta = \arccos(\frac{1}{2}(tr(R)-1))$ and $[\mathbf{u}]_\times = \frac{R - R^T}{2\,\sin c(\theta)}$ [4].

However, a direct control in the Cartesian space might result in the features leaving the camera field of view. To enforce the visibility constraint, we use a single image point to control two axes of rotation (around x and y) and the final axis of rotation is controlled directly using the rotation matrix. This is done as follows: We know from the image-based visual servoing control [2]

$$\begin{bmatrix} u - u^* \\ v - v^* \end{bmatrix}_{2\times 1} = \begin{bmatrix} \underbrace{-\frac{1}{Z} & 0 & \frac{u}{Z}}_{L_\nu} & \underbrace{uv & -(1+u^2)}_{L_{\omega_{xy}}} & \underbrace{v}_{L_{\omega_z}} \\ 0 & -\frac{1}{Z} & \frac{v}{Z} & 1+v^2 & -uv & -u \end{bmatrix}_{2\times 6} \begin{bmatrix} \nu_{3\times 1} \\ \omega_{3\times 1} \end{bmatrix}_{6\times 1}, \quad (4)$$

where $p = [u\ v\ 1]^T = [x\ 1]^T$, $p^* = [u^*\ v^*\ 1]^T = [x^*\ 1]^T$, $Z = \mathbf{Z}(P)$ (See Fig. 2) and $[\nu\ \omega]^T$ denotes the camera velocity. Equation (4) relates the motion of image features *i.e.*, $x - x^*$ to the camera motion using the 2×6 Jacobian matrix L. It can be rewritten as $x - x^* = [L_\nu\ L_{\omega_{xy}}\ L_{\omega_z}][\nu\ \omega_{xy}\ \omega_z]^T$. Observe that a simple rearrangement of terms yields

$$\omega_{xy} = L_{\omega_{xy}}^{-1}[(x - x^*) - L_\nu \nu - L_{\omega_z} \omega_z], \quad (5)$$

where $\nu = (\frac{t}{d})\hat{d}$ and $\omega_z = \mathbf{u_z}\theta$. In (5), the rotational motion ω_{xy} is controlled not only to minimize the differences between the current and the goal image features but also to compensate the effects caused by translation on the image. This ensures a straight-line feature trajectory in the image and thereby guarantees object visibility. Estimates of the values Z and d are required in (4) that can be obtained as follows: Firstly, observe that

$$det(H) = det(R + \frac{tn^T}{d}) = det(R + \frac{t(n^{*T}R)}{d}) \quad (6)$$

$$= det(I + \frac{tn^{*T}}{d})det(R) = \frac{d + n^{*T}t}{d} \tag{7}$$

where (6) uses the fact that $n^* = Rn$ (See Fig. 2). Equation (7) can be further simplified using the result $d^* - d = n^{*T}P^* - n^T P = n^{*T}(P^* - RP) = n^{*T}t$. Hence we have $\hat{d} = \frac{\hat{d}^*}{det(H)}$. Using (7), Z can be calculated as

$$\frac{Z}{d^*} = \frac{Z}{d^*}\frac{d}{n^T P} = \frac{1}{n^T p}\frac{1}{det(H)}. \tag{8}$$

Thus we have $\hat{Z} = \frac{\hat{d}^*}{n^T p}\frac{1}{det(H)}$, where $\hat{d}^*$ is an estimate of the constant distance to the plane in the desired camera frame. In general, this quantity is considered as a gain ratio [6] and a coarse estimate obtained from a simple stereo technique is adequate. Consequently, all the parameters required for the control are now available directly from the homography decomposition.

In summary, the resultant expression for the velocity $\mathbf{v}$ is given as

$$\mathbf{v} = \begin{bmatrix} -\lambda_v I_{3\times3} & 0_{3\times2} & 0_{3\times1} \\ 0_{2\times3} & -\lambda_{\omega_{xy}} I_{2\times2} & 0_{2\times1} \\ 0_{1\times3} & 0_{1\times2} & -\lambda_{\omega_z} I_{1\times1} \end{bmatrix} \begin{bmatrix} \nu \\ \omega_{xy} \\ \omega_z \end{bmatrix} \left(= \begin{bmatrix} (\frac{t}{d})\,\hat{d} \\ L_{\omega_{xy}}^{-1}[(x - x^*) - L_\nu\nu - L_{\omega_z}\omega_z] \\ \mathbf{u_z}\theta \end{bmatrix} \right) \tag{9}$$

Equation (9) has only one singularity that occurs at $Z = 0$ (See expression for L_ν). However, as discussed in the earlier section, this degenerate configuration is avoided by the reliable homography computation algorithm. Thus by incorporating image features into the $3D$ control, an efficient control offering the complimentary features of object visibility and decoupled trajectory has been developed.

5   Experimental Results

In our experiments, we constructed an arbitrary configuration of planes as shown in Fig. 4(a). The projection of points belonging to these planar regions onto the image were considered as features. A perspective camera projection model was assumed. The basic implementation of the proposed algorithm given below was used to perform the positioning task.

1. Extract features from the current image and partition them into piecewise planar regions
2. Compute homography H_i induced by each region
3. Select four independent homographies induced by the regions that have the largest areas and best visibility (Only the selected regions need to be tracked in the successive iterations)
4. Determine the weights using the geometric constraints and compute the normalized weight λ_i for the selected homographies (Sect. 3.1)
5. Determine the robust homography H_{res} using (3)
6. Decompose H_{res} to obtain the motion and structure parameters (Resolve ambiguity using an additional homography)

7. Use the control law to obtain the velocity instruction **v** (See (9))
8. Repeat above steps until convergence

We analyzed the performance of our algorithm by generating several random initial camera configurations and then moving the camera to a fixed desired pose in a multi-plane scenario as shown in Fig. 4(a). Observe that a camera can frequently encounter degenerate cases during the positioning task in such a scene. However, in almost all the cases, the proposed algorithm was uninfluenced by degeneracies. In Fig. 4(c), the velocity command generated by the proposed approach for the particular scenario as tested in Fig. 3(a) is shown. Fig. 4(b) shows the variation in weights corresponding to the homographies. Observe that the weight corresponding to degenerate H tends towards the minimum value as the camera approaches the degeneracy. The smooth velocity screw in Fig. 4(c) demonstrates the stable behavior of the algorithm unlike in Fig. 3(b). Fig. 4(d) displays the camera trajectory. Note that the expression for Z in (8) requires at
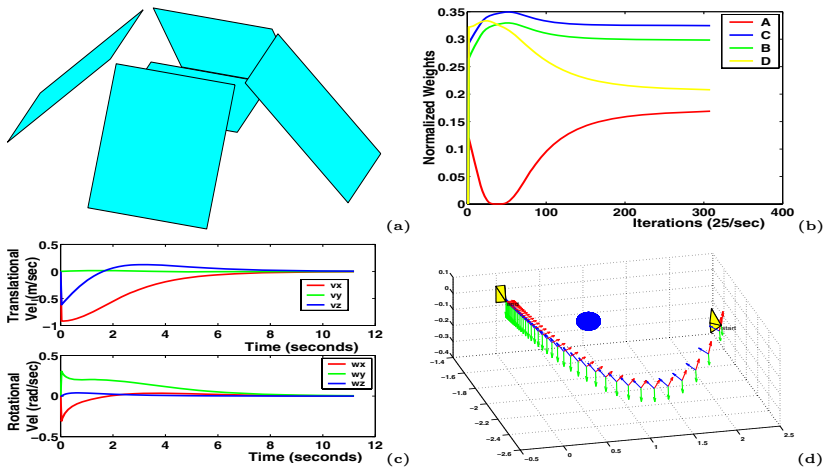

Fig. 4. (a) Non-Planar scene considered in the experiments. (b) Normalized weight values (c) Velocity Screw and (d) Camera Trajectory obtained for the scenario described in Fig. 3(a) . Smooth convergence even in presence of degeneracies confirms the stable behavior of the proposed approach.

least one feature p belonging to the planar region. However as a virtual homography is being used in our case, it might not correspond to any physical plane in the scene. In our method, we obtained this feature by finding the intersection of the plane inducing the virtual homography H_{res} with other scene planes as described in [15].

Analysis of the Control Law. The performance of the control law was analyzed in simulation. Fig. 5 shows the velocity screw and the image feature trajectory obtained during a positioning task using the proposed, $3D$ and the
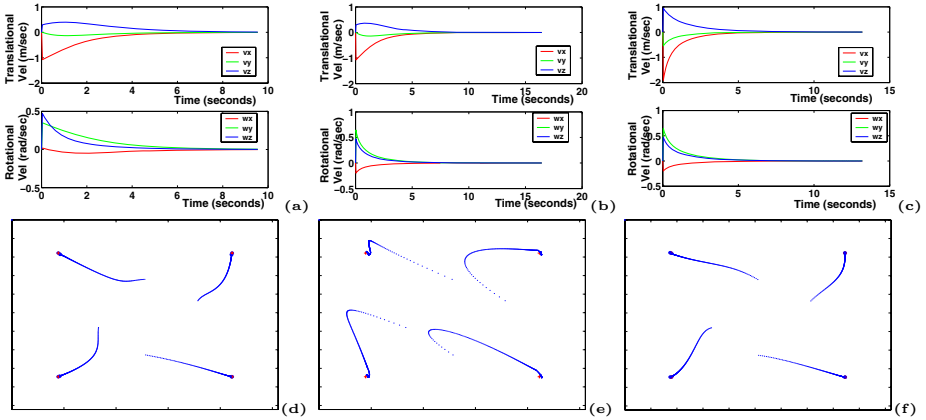
Fig. 5. Analysis of proposed control: Fig.(a),(b),(c) show the velocity screw obtained in case of proposed, $3D$ and 21/2D controls respectively while (d),(e),(f) display the feature trajectory. Similarity of velocity screws in (a) and (b) confirms the optimal trajectory behavior of the proposed control while near straight-line image feature trajectory in (d) ascertains the feature visibility.
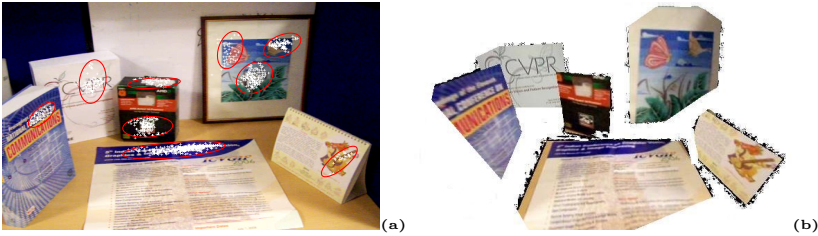

Fig. 6. Planar scene reconstruction using inter-image homographies: (a) A sample frame along with the detected interest regions on the scene planes (b) Reconstruction result

21/2D [6] controls respectively. The velocity screw obtained using the proposed control is very similar to the one obtained using the $3D$ control. Further, the feature trajectory almost follows a straight line. These two observations ascertain our claims of decoupled (straight-line) camera trajectory and object visibility using the proposed control. Inter-image homographies are an interesting tool for reconstruction of planar surfaces. The decomposition of homographies provide the $3D$ plane parameters required to reconstruct the scene. By considering a common feature belonging to two planes n_i and n, a relationship could be derived between their distances using (8) as

$$Z = \frac{d_i}{n_i^T p} \frac{1}{det(H)} = \frac{d}{n^T p} \frac{1}{det(H)} \quad i.e., \quad d_i = \frac{n_i^T p}{n^T p} d, \tag{10}$$

where p denotes the common image feature. Thus given the plane normals n_i, the $3D$ scene could be reconstructed up to a scale factor d (See Fig. 6(b)). Given an estimate of d, the exact scene can be reconstructed.

6 Conclusion

A novel homography-based control capable of positioning a camera even in presence of non-planar objects has been developed for the first time in this paper. A robust homography estimate was efficiently computed using multiple homographies by employing geometric and subspace constraints. This homography estimate was used in a modified control law to compute the optimal camera trajectory. The method performed better in comparison to existing servoing algorithms and avoided their critical drawbacks. In future, we plan to investigate further the utility of multi-plane homography-based formulations for efficiently solving other classical computer vision problems.

References

1. DeSouza, G., Kak, A.: Vision for mobile robot navigation: A survey. IEEE Transactions on Pattern Analysis and Machine Intelligence **24** (2002) 237–267
2. Hutchinson, S.A., Hager, G.D., Corke, P.I.: A tutorial on visual servo control. IEEE Transactions on Robotics and Automation **12** (1996) 651–670
3. Chaumette, F., Espiau, B.: A new approach to visual servoing in robotics. IEEE Transactions on Robotics and Automation **8** (1992) 313–327
4. Wilson, W.J., Hulls, C.C.W., Bell, G.S.: Relative end effector control using cartesian position based visual sovoing. IEEE Transactions on Robotics and Automation **12** (1996) 684–696
5. Taylor, C.J., Ostrowski, J.P., Jung, S.H.: Robust visual servoing based on relative orientation. IEEE Computer Society Conference on Computer Vision and Pattern Recognition **2** (1999) 574–580
6. Malis, E., Chaumette, F.: 2 1/2D visual servoing with respect to unknown objects through a new estimation scheme of camera displacement. International Journal of Computer Vision **37** (2000) 79–97
7. Rives, P.: Visual servoing based on epipolar geometry. IEEE/RSJ International Conference on Intelligent Robots and Systems **1** (2000) 602–607
8. Basri, R., Rivlin, E., Shimshoni, I.: Visual homing: surfing on the epipoles. IEEE International Conference on Computer Vision (1998) 863–869
9. Hartley, R., Zisserman, A.: Multiple view geometry in computer vision. Cambridge University Press (2003)
10. Faugeras, O., Lustman, F.: Motion and strucutre from motion in a piecewise planar environment. International Journal of Pattern Recognition and Artificial Intelligence **2** (1988) 485–508
11. Boufama, B., Mohr, R.: Epipole and fundamental matrix estimation using the virtual parallax property. IEEE International Conference on Computer Vision (1995) 1030–1036
12. Fraundorfer, F., Bischof, H.: Detecting distinguished regions by saliency. Scandinavian Conference on Image Analysis (2003) 208–215

13. Shashua, A., Avidan, S.: The rank-4 constraint in multiple view geometry. European Conference on Computer Vision **2** (1996) 196–206
14. Deguchi, K.: Optimal motion control for image-based visual servoing by decoupling translation and rotation. IEEE/RSJ International Conference on Intelligent Robots and Systems **2** (1998) 705–711
15. Johansson, B.: View synthesis and 3D reconstruction of piecewise planar scenes using intersection lines between the planes. IEEE International Conference on Computer Vision **1** (1999) 54–59