Three-D Motion and Dense Structure Estimation Using Convex Projections

A. Aydın Alatan, A. Tanju Erdem, and Levent Onural
Department of Electrical and Electronics Engineering
Bilkent University, TR-06533, Ankara, Turkey
e-mail: alatan@ee.bilkent.edu.tr

ABSTRACT

We propose a novel method for estimating the 3-D motion and dense structure (depth) of an object from its two 2-D images. The proposed method is an iterative algorithm based on the theory of projections onto convex sets (POCS) that involves successive projections onto closed convex constraint sets. We seek a solution for the 3-D motion and structure information that satisfies the following constraints: (i) Rigid motion—the 3-D motion (rotation and translation) parameters are the same for each point on the object. (ii) Smoothness of the structure—the depth values of the neighboring points on the object vary smoothly. (iii) Temporal correspondence—the intensities in the given 2-D images match under the 3-D motion and structure parameters. We mathematically derive the projection operators onto these sets and discuss the convergence properties of successive projections. Experimental results show that the proposed method significantly improves the initial motion and structure estimates.

Keywords : 3-D motion estimation, structure estimation, projections onto convex sets, video coding.

1 INTRODUCTION

3-D motion estimation refers to finding the actual motion of an object in a 3-D scene, which is observed through consecutive 2-D video frames. The 2-D projection of the actual motion of an object depends not only on the 3-D motion parameters, but also on the object depth information (structure) which is simply defined as the distance of the object surface points from the camera. Hence, in video compression applications where 2-D projections are utilized for motion compensated prediction, both depth and 3-D motion information should be estimated.

There are different ways to approach the 3-D motion and structure estimation problem. The methods which estimate 3-D rigid motion from consecutive monocular frames can actually be divided into two major classes, as *direct* and *correspondence based* methods. Direct methods use spatio-temporal gradients in the image to find a solution to the 3-D motion estimation problem.[1] Currently, there is no general solution to direct methods and it is only possible to obtain a solution by making some

SPIE Vol. 3024 • 0277-786X/97/$10.00

simplifying assumptions about motion and/or structure.[1] On the other hand, any 3-D motion estimation method which requires some dense or sparse set of 2-D motion vectors for finding the 3-D motion parameters is said to be correspondence based.[2] These methods require some point matches between frames, which can be obtained by one of the feature matching[3] or 2-D motion estimation algorithms.[4] However, incorrect matches may lead to unstable solutions and the performance of any correspondence based method mainly depends on this initial matching step.

In this paper, we introduce a novel solution to the problem of 3-D motion and structure estimation. The proposed method can be used in video coding applications to obtain the 3-D motion and the structure of a moving object between consecutive frames. Given two input frames, we define certain constraints such that the 3-D motion and structure information should obey these constraints. The intersection of all these constraints (if it exists) contains a solution to the motion and structure estimation problem. The theory of convex projections not only shows how to approach the intersection of these convex sets from an initial point, it also guarantees the convergence of the iterations to a point (solution) in the intersection set.[5] Although, a similar approach is used in an earlier work[6] to estimate 3-D motion and structure, the proposed constraint sets in that work are not convex and hence, the convergence is not guaranteed.

In Section 2, we give a brief overview of modeling 3-D motion and the method of projections onto convex sets (POCS). In Section 3, we introduce the proposed method, the constraint sets, and the corresponding projection operators. Experimental results are provided in Section 4. Finally, concluding remarks are given in Section 5.

2 BACKGROUND

2.1 Modeling of 3-D Motion

Let $\mathbf{P}$ define an object in the 3-D object space and let $\mathbf{p} \in \mathbf{P}$ be an object point whose 3-D coordinates at time t is given by $\mathbf{X_p}(t) = [X_\mathbf{p}(t)\ Y_\mathbf{p}(t)\ Z_\mathbf{p}(t)]^T$. The perspective projection of $\mathbf{X_p}(t)$ onto the image plane is written as $\mathbf{x_p}(t) = [x_\mathbf{p}(t)\ y_\mathbf{p}(t)]^T$, which is shown in Figure 1. For any rigid motion from time $t-1$ to t, the 3-D coordinates of object point $\mathbf{p}$ at time $t-1$ can be written in terms of $\mathbf{X_p}(t)$ as

$$\mathbf{X_p}(t-1) = \mathbf{R}\,\mathbf{X_p}(t) + \mathbf{T} \tag{1}$$

where $\mathbf{R}$ is a 3x3 rotation matrix, $\mathbf{T}$ is a 3x1 translation vector. It should be noted that $\mathbf{R}$ and $\mathbf{T}$ do not reflect the "real" motion from time $t-1$ to t, but rather an "inverse" motion from time t to $t-1$ for video coding purposes. Among different rotation matrix representations, $\mathbf{R}_{x,y,z}$ models the overall motion as three consecutive rotations around coordinates axes.[9] For the rotations around (x, y, z) coordinate axes, the corresponding rotation matrix can be written as

$$\mathbf{R}_{x,y,z} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & cos(w_x) & sin(w_x) \\ 0 & -sin(w_x) & cos(w_x) \end{bmatrix} \cdot \begin{bmatrix} cos(w_y) & 0 & -sin(w_y) \\ 0 & 1 & 0 \\ sin(w_y) & 0 & cos(w_y) \end{bmatrix} \cdot \begin{bmatrix} cos(w_z) & sin(w_z) & 0 \\ -sin(w_z) & cos(w_z) & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{2}$$

where w_x, w_y and w_z are the rotation angles for x, y and z axes, respectively.

After perspective projection of the 3-D object points into 2-D image plane, the equations below are obtained[2] (the focal length in the perspective mapping is chosen to be the unit length with no loss of generality).

$$
\begin{aligned}
x_{\mathbf{p}}(t-1) &= \frac{r_{11}\cdot x_{\mathbf{p}}(t) + r_{12}\cdot y_{\mathbf{p}}(t) + r_{13} + \frac{T_x}{Z_{\mathbf{p}}(\mathbf{x_p},t)}}{r_{31}\cdot x_{\mathbf{p}}(t) + r_{32}\cdot y_{\mathbf{p}}(t) + r_{33} + \frac{T_z}{Z_{\mathbf{p}}(\mathbf{x_p},t)}} \\
y_{\mathbf{p}}(t-1) &= \frac{r_{21}\cdot x_{\mathbf{p}}(t) + r_{22}\cdot y_{\mathbf{p}}(t) + r_{23} + \frac{T_y}{Z_{\mathbf{p}}(\mathbf{x_p},t)}}{r_{31}\cdot x_{\mathbf{p}}(t) + r_{32}\cdot y_{\mathbf{p}}(t) + r_{33} + \frac{T_z}{Z_{\mathbf{p}}(\mathbf{x_p},t)}}
\end{aligned}
\tag{3}
$$

$\mathbf{x_p}(t-1) = [x_{\mathbf{p}}(t-1)\ \ y_{\mathbf{p}}(t-1)]^T$ is the projected 2-D coordinates of the object point $\mathbf{p}$ at time $t-1$. Notice that $Z_{\mathbf{p}}(\mathbf{x_p}, t)$ is the third component of the vector $\mathbf{X_p}(t)$ whose perspective projection gives $\mathbf{x_p}(t)$ and simply called as "depth value". However, it should be noted that "depth field" term is being used as the set of depth values defined on the 2-D lattice, Λ. Hence the depth field reflects only the Z values of the projected 3-D object points. In Equation (3), the parameters $r_{i,j}$, which are a function of rotation angles,w_x, w_y and w_z, denote the elements of the rotation matrix $\mathbf{R}_{x,y,z}$. For notational simplicity, $\mathbf{p}$ subscript will not be used to label the object point coordinates in the rest of the paper.

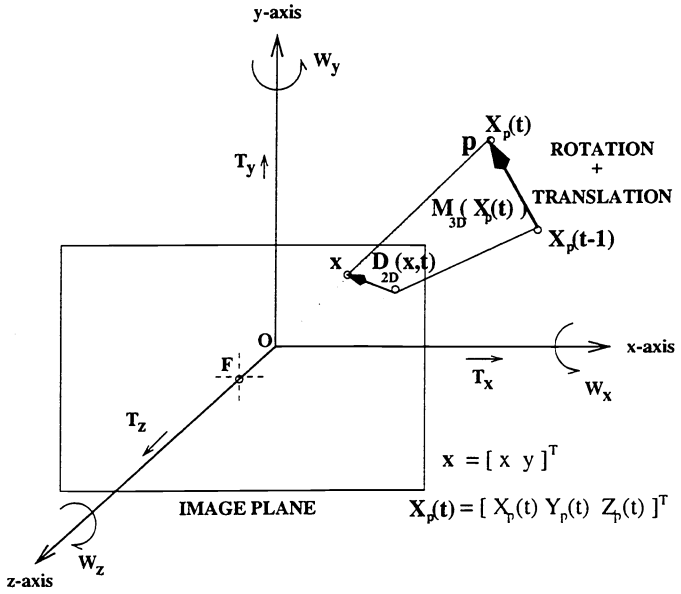

Figure 1: 3-D coordinate system

2.2 Overview of the POCS method

In the method of projections onto convex sets (POCS), the unknown quantity $\mathbf{q}$ (in our case the 3-D motion and the depth field) is assumed to be an element of an appropriate Hilbert space $\mathcal{H}$.[7] Each *a priori* information or constraint restricts the solution to a closed convex set in $\mathcal{H}$. For each piece of information. there is a corresponding constraint set, C_i. The unknown quantity $\mathbf{q}$ is an element of the intersection of all these constraint sets, if this intersection is nonempty.

Given the constraint sets C_i, $i = 1, \cdots, m$, and their respective projection operators $\mathbf{P}_i$, $i = 1, \cdots, m$, the sequence generated by

$$\mathbf{q}_{k+1} = \mathbf{P}_m \cdots \mathbf{P}_1 \mathbf{q}_k, \quad k = 0, 1, \cdots,$$

converges to a solution in the intersection of $C_1, \cdots, C_m$. Figure 2 demonstrates the convergence of $\mathbf{q}$ for $m = 2$. The projection operators for each constraint set minimizes the distance between the initial location and the corresponding set, as it is shown in Figure 2. The initial location can be arbitrarily chosen from $\mathcal{H}$. More discussion on the fundamentals of POCS method can be found in.[8]

Hence, in order to apply the POCS method for the estimation 3-D motion and structure, certain convex constraint sets and their corresponding projection operators are derived in the next section.

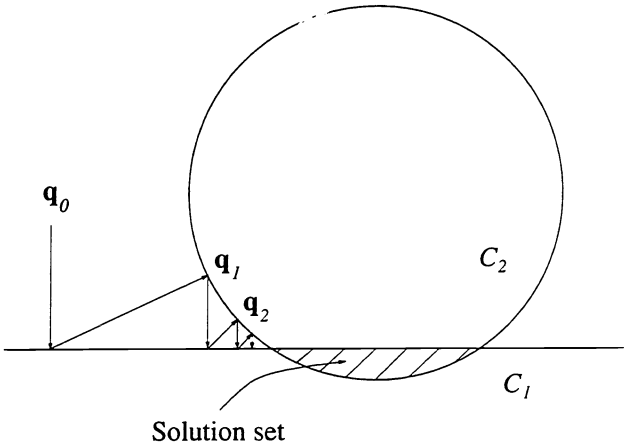

Figure 2: The method of projection onto convex sets.

3 PROPOSED METHOD

We let $\mathbf{x}$ and $\mathbf{x}'$ denote the corresponding 2-D image coordinates of an object point $\mathbf{p}$ in Equation (3) at times t and $t-1$, respectively. Let $\mathbf{m}(\mathbf{x})$ denote the 6×1 motion vector which is made of the rotation angles (w_x, w_y, w_z) around and the translations (T_x, T_y, T_z) along the three coordinate axes at point $\mathbf{x}$ (Figure 1). We define $\mathbf{q}(\mathbf{x}) = [\mathbf{m}(\mathbf{x}) \;\; Z(\mathbf{x})]^T$ as the 7×1 motion-structure vector at $\mathbf{x}$, where $Z(\mathbf{x})$ denotes the depth value at $\mathbf{x}$. Finally, let $\mathcal{S}$ denote the 2-D support of the object at time t; $\mathbf{x}_1, \cdots, \mathbf{x}_N$ denote the points in $\mathcal{S}$, and $\mathbf{q} = [\mathbf{q}(\mathbf{x}_1) \; \cdots \; \mathbf{q}(\mathbf{x}_N)]^T$ denote the $7N \times 1$ vector of 3-D motion and depth parameters that is to be found.

3.1 Definition of Convex Sets

There are three pieces of *a priori* information that non-deformable 3-D motion and structure should obey :

- Rigid motion: The 3-D motion (rotation and translation) parameters are the same for each point on the object.
- Smoothness of the structure: The depth values of the neighboring points on the object vary smoothly.
- Temporal correspondence: The intensities in the given 2-D images match under the 3-D motion and structure parameters.

The following closed convex set is defined to represent the rigid motion constraint

$$C_r \doteq \{\mathbf{q}: \ \mathbf{m}(\mathbf{x}_i) = \mathbf{m}(\mathbf{x}_j), \quad \forall \mathbf{x}_i, \mathbf{x}_j \in \mathcal{S}\}. \tag{4}$$

The smoothness constraint, on the other hand, is represented by the collection of the following closed convex sets

$$C_{s,k} \doteq \{\mathbf{q}: \ |Z(\mathbf{x}_i) - Z(\mathbf{x}_j)| \leq \delta_s, \quad \mathbf{x}_i, \mathbf{x}_j \in \mathcal{N}_k\}, \quad k = 1, \cdots, K, \tag{5}$$

where $\{\mathcal{N}_1, \cdots, \mathcal{N}_K\}$ is the collection of all pairs of neighboring points in $\mathcal{S}$; we say that $\mathbf{x}_i$ and $\mathbf{x}_j$ are neighbors if $\mathbf{x}_i - \mathbf{x}_j = [1\ 0]^T$, $[-1\ 0]^T$, $[0\ 1]^T$, or $[0\ -1]^T$. The quantity δ_s in Equation (5) is an *a priori* bound reflecting the statistical confidence with which the actual motion-structure vector is a member of the set $C_{s,k}$.

Finally, the constraint set representing temporal correspondence can be expressed as a collection of the sets

$$\{\mathbf{q}: \ |I_t(\mathbf{x}) - I_{t-1}(\mathbf{x}')| \leq \delta_t\}, \quad \mathbf{x} \in \mathcal{S}, \tag{6}$$

where I_t denotes the intensity distribution at time t, and the quantity δ_t is the *a priori* confidence bound. We note that the set given in Equation (6) is not convex. We perform a 2-step linearization to obtain a convex constraint set that approximates the set given in Equation (6). First, $I_{t-1}(\mathbf{x}')$ is linearized using spatial gradients of the intensity distribution around a given initial estimate of the motion-structure vector $\tilde{\mathbf{q}}$ (which can be found using one of the techniques in,[7] such as the E-matrix algorithm),

$$I_{t-1}(\mathbf{x}') \simeq I_{t-1}(\tilde{\mathbf{x}}') + \left(\frac{\partial I_{t-1}(\mathbf{x})}{\partial \mathbf{x}}\bigg|_{\mathbf{x}=\tilde{\mathbf{x}}'} \right)^T \left(\mathbf{x}' - \tilde{\mathbf{x}}'\right), \tag{7}$$

where $\tilde{\mathbf{x}}' = \mathbf{f}(\mathbf{x}, \tilde{\mathbf{q}}(\mathbf{x}))$, and $\mathbf{f}$ represents the vector function that evaluates the coordinates of a point at time $t-1$ given the coordinates of the point at time t and the motion-structure vector between frames t and $t-1$. The explicit form of $\mathbf{f}$ is given in Equation (3). In the second step, $\mathbf{x}'$ is linearized with respect to the motion-structure parameters as

$$\mathbf{x}' \simeq \tilde{\mathbf{x}}' + \left(\frac{\partial \mathbf{f}(\mathbf{x}, \mathbf{q}(\mathbf{x}))}{\partial \mathbf{q}(\mathbf{x})}\bigg|_{\mathbf{q}(\mathbf{x})=\tilde{\mathbf{q}}(\mathbf{x})} \right)^T (\mathbf{q}(\mathbf{x}) - \tilde{\mathbf{q}}(\mathbf{x})). \tag{8}$$

Combining Equations (6), (7) and (8), we obtain a closed convex set approximating the temporal correspondence constraint

$$C_{t,n} = \{\mathbf{q}: \ |\mathrm{DID}(\mathbf{x}_n) - \mathbf{k}^T(\mathbf{x}_n)\,(\mathbf{q}(\mathbf{x}_n) - \tilde{\mathbf{q}}(\mathbf{x}_n))\,| \leq \delta_t\}, \quad n = 1, \cdots, N, \tag{9}$$

where $\mathrm{DID}(\mathbf{x}) \doteq I_t(\mathbf{x}) - I_{t-1}(\tilde{\mathbf{x}}')$ and $\mathbf{k}(\mathbf{x})$ is a 7×1 vector defined in terms of the products of partial derivatives of $\mathbf{f}$ and spatial gradients of I_{t-1} as

$$\mathbf{k}(\mathbf{x}) = \frac{\partial \mathbf{f}(\mathbf{x}, \mathbf{q}(\mathbf{x}))}{\partial \mathbf{q}(\mathbf{x})}\bigg|_{\mathbf{q}(\mathbf{x})=\tilde{\mathbf{q}}(\mathbf{x})} \frac{\partial I_{t-1}(\mathbf{x})}{\partial \mathbf{x}}\bigg|_{\mathbf{x}=\tilde{\mathbf{x}}'}. \tag{10}$$

3.2 Projection Operators

The projection operator, which finds an element in the corresponding constraint set with minimum distance from the initial value, is found as a result of a constrained minimization. While the existence of the projected point in the convex set is written as a constraint, the distance between the initial given and final projected points are tried to be minimized.[8]

The constrained minimization

$$\min_{\mathbf{m}(\mathbf{x}),\ \mathbf{x}\in\mathcal{S}} \sum_{\mathbf{x}\in\mathcal{S}} ||\mathbf{m}(\mathbf{x}) - \mathbf{m}_0(\mathbf{x})||^2 \tag{11}$$

$$\textit{subject to} \quad \mathbf{m}(\mathbf{x}_i) = \mathbf{m}(\mathbf{x}_j), \quad \forall \mathbf{x}_i, \mathbf{x}_j \in \mathcal{S}$$

where the subscript 0 denotes the initial value of the quantity before the projection, the projected motion-structure vector $\mathbf{q}$ inside C_r can be obtained using the projection

$$P_r\mathbf{q} = [\mathbf{q}^\star(\mathbf{x}_1) \ \cdots \ \mathbf{q}^\star(\mathbf{x}_N)]^T \quad \text{where} \quad \mathbf{q}^\star(\mathbf{x}_i) = [\mathbf{m}^\star \ \ Z(\mathbf{x}_i)]^T, \quad \mathbf{m}^\star = \frac{1}{N}\sum_{\mathbf{x}\in\mathcal{S}} \mathbf{m}(\mathbf{x}). \tag{12}$$

Similarly, the result of the constrained minimization[10]

$$\min_{Z(\mathbf{x}_i),Z(\mathbf{x}_j)} \left((Z(\mathbf{x}_i) - Z_0(\mathbf{x}_i))^2 + (Z(\mathbf{x}_j) - Z_0(\mathbf{x}_j))^2 \right) \tag{13}$$

$$\textit{subject to} \quad |Z(\mathbf{x}_i) - Z(\mathbf{x}_j)| \le \delta_s, \quad \mathbf{x}_i, \mathbf{x}_j \in \mathcal{N}_k, \quad k = 1, \cdots, K,$$

gives the projection of $\mathbf{q}$ onto $C_{s,k}$ as

$$P_{s,k}\mathbf{q} = [\mathbf{q}^\star(\mathbf{x}_1) \ \cdots \ \mathbf{q}^\star(\mathbf{x}_N)]^T \quad \text{where} \quad \mathbf{q}^\star(\mathbf{x}_i) = [\mathbf{m}(\mathbf{x}_i) \ \ Z^\star(\mathbf{x}_i)]^T, \tag{14}$$

$$Z^\star(\mathbf{x}_i) = \begin{cases} Z(\mathbf{x}_i) & \text{if } \mathbf{x}_i \notin \mathcal{N}_k \\ Z(\mathbf{x}_i) & \text{if } |Z(\mathbf{x}_i) - Z(\mathbf{x}_j)| \le \delta_s, \ \mathbf{x}_i, \mathbf{x}_j \in \mathcal{N}_k \\ \frac{1}{2}[Z(\mathbf{x}_i) + Z(\mathbf{x}_j) + \delta_s] & \text{if } Z(\mathbf{x}_i) - Z(\mathbf{x}_j) > \delta_s, \ \mathbf{x}_i, \mathbf{x}_j \in \mathcal{N}_k \\ \frac{1}{2}[Z(\mathbf{x}_i) + Z(\mathbf{x}_j) - \delta_s] & \text{if } Z(\mathbf{x}_i) - Z(\mathbf{x}_j) < -\delta_s, \ \mathbf{x}_i, \mathbf{x}_j \in \mathcal{N}_k \end{cases} \tag{15}$$

Finally, the projection of $\mathbf{q}$ onto $C_{t,n}$ can be found as below at the end of a similar constrained minimization:

$$P_{t.n}\mathbf{q} = [\mathbf{q}^\star(\mathbf{x}_1) \ \cdots \ \mathbf{q}^\star(\mathbf{x}_N)]^T, \quad \mathbf{q}^\star(\mathbf{x}_i) = \begin{cases} \mathbf{q}(\mathbf{x}_i) & \text{if } i \ne n \\ \mathbf{q}(\mathbf{x}_n) & \text{if } |\mathrm{DID}(\mathbf{x}_n)| \le \delta_t, \quad i = n \\ \mathbf{q}(\mathbf{x}_n) + \dfrac{\mathrm{DID}(\mathbf{x}_n) - \delta_t}{\mathbf{k}^T(\mathbf{x}_n)\mathbf{k}(\mathbf{x}_n)}\mathbf{k}(\mathbf{x}_n) & \text{if } \mathrm{DID}(\mathbf{x}_n) > \delta_t, \quad i = n \\ \mathbf{q}(\mathbf{x}_n) + \dfrac{\mathrm{DID}(\mathbf{x}_n) + \delta_t}{\mathbf{k}^T(\mathbf{x}_n)\mathbf{k}(\mathbf{x}_n)}\mathbf{k}(\mathbf{x}_n) & \text{if } \mathrm{DID}(\mathbf{x}_n) < -\delta_t, \quad i = n \end{cases} . \tag{16}$$

3.3 Algorithm

In the proposed algorithm, we employ the following iterations to improve the initial estimate of the motion-structure vector:

$$\mathbf{q}_{\ell+1} = P_r P_{s,K} \cdots P_{s,1} P_{t,N} \cdots P_{t,1} \mathbf{q}_\ell, \quad \ell = 0, 1, 2, \cdots, \tag{17}$$

where $\mathbf{q}_0 = \tilde{\mathbf{q}}$, i.e., the given initial estimate of the motion-structure vector. At the beginning of the ℓth iteration, $\tilde{\mathbf{q}}$ in Equation (10) is updated to $\mathbf{q}_\ell$ in order to improve the linear approximations in Equations (7) and (8). Because $C_{t,1}, \cdots, C_{t,N}$ in Equation (9) depend on $\tilde{\mathbf{q}}$, the projection operators $P_{t,1}, \cdots, P_{t,N}$ in Equation (17) vary at each iteration. Therefore, the theory of POCS cannot be directly applied here to guarantee the convergence of Equation (17). However, experimental results show that, the estimate of the motion-structure vector improves with the increased values of ℓ in Equation (17).

4 RESULTS

The experiments are carried out on an artificial sequence called *Salecube*. The two frames of the *Salecube* sequence used for motion and structure estimation are given in Figure 5(a) and (b), respectively. As observed from Figure 5, the "cube" moves in front of a stationary background. The texture on faces of the cube is obtained as a result of warping certain regions of still images such as those in the *Salesman* sequence. We inject the true 3-D motion and structure parameters of the moving cube with noise and use the resulting noisy motion and structure data to initialize the iterations in the algorithm proposed in Section 3.3. The threshold δ_t is determined according to the variance of the noise on the intensities and during the experiments this parameter is selected as 1. Similarly, δ_s is selected using the variance of the initial depth estimate and its value is chosen as 50. The intensity gradients, which are used in the intensity matching projection operator, are obtained after fitting a surface to the blurred intensities of the original input frames.

The $PSNR$ of the reconstructed Frame 2 is shown in Figure 3 for 50 iterations. As seen from Figure 3, a convergence of projections is achieved after about 10 iterations.

As shown in Figure 4, the errors between the true and estimated values decreases for all 3-D motion parameters, except the error on w_z slightly increases during iterations. Since the projection operators $P_{t,1}, \cdots, P_{t,N}$ vary after each iteration and in such a case POCS does not guarantee convergence, the increase in the error for w_z is possible. However, the reduction in the errors for rotation parameters is significant at the end of the iterations.

Furthermore, when this method is used in video coding algorithms with 3-D motion models, the most important performance parameter will be the $PSNR$ of the reconstructed current frame, which is obtained using the available previous frame and the estimated 3-D motion and structure parameters. The $PSNR$ plot in Figure 3.(b) increases monotonically during the iterations and about 10 dB gain is obtained in quality as a result of the proposed algorithm.

In Figure 5, the reconstructed and difference images are shown before and after the proposed algorithms. The difference images clearly show the improvements on the reconstructed image for Frame 2.

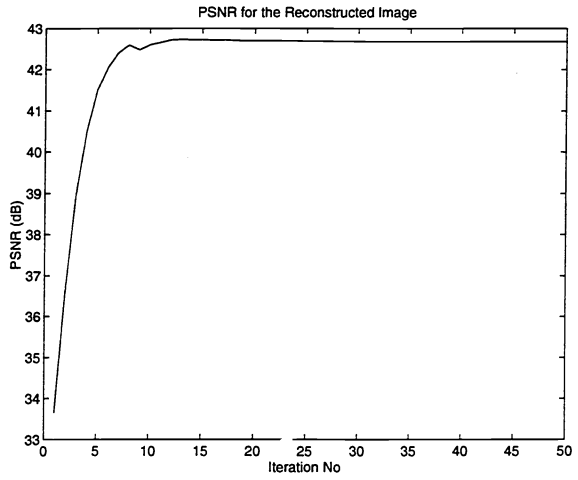

Figure 3: $PSNR$ of the reconstructed Frame 2 for 50 iterations.

5 CONCLUSIONS

A novel 3-D motion and structure estimation algorithm is proposed. The constraints sets defined in this paper not only leads to a solution for both depth and 3-D motion, but they also impose advantageous properties on the obtained motion and structure parameters for video coding application. While the rigid body and the smooth surface constraints give efficient description of the motion and structure parameters for video coding purposes, the intensity matching constraint improves the visual quality of the reconstructed frame. Hence, the proposed algorithm is indeed suitable for video compression applications.

Although, the defined constraint sets are convex, the method of POCS does not guarantee convergence to a solution, since one of the constraint sets (intensity matching) changes after each iteration. Nevertheless, the experimental results suggest that if an initial estimate of the motion and structure information is available, the proposed algorithm still converges to an acceptable solution.

6 REFERENCES

[1] B. K. P. Horn. *Robot Vision*, pp. 401–417. MIT Press, Cambridge, 1986.

[2] T. S. Huang and A. N. Netravali "Motion and Structure from Feature Correspondences: A Review," *IEEE Proceedings*, vol. 82, pp. 252–268, February 1994.

[3] J. Weng, N. Ahuja and T. S. Huang "Matching Two Perspective Views," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 14, pp. 806–825, August 1992.

[4] J. L. Barron, D. J. Fleet and S. S. Beauchemin "Performance of Optical Flow Techniques," *International Journal of Computer Vision*, vol. 12, pp. 43–77, January 1994.

[5] D.C. Youla and H. Webb "Image Restoration by the Method of Convex Projections: Part 1-Theory," *IEEE Trans. on Medical Imaging*, vol. 1, pp. 81–94, October 1982.

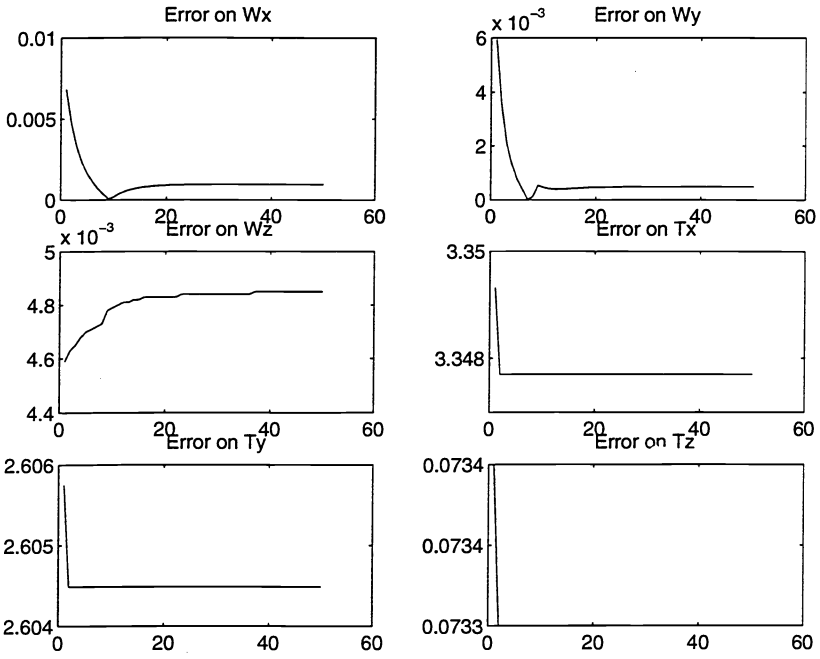

Figure 4: The error plots between the true and the estimated motion parameters (w_x,w_y,w_z.t_x,t_y,t_z) during iterations.

[6] A. Kara, D. M. Wilkes and K. Kawamura "3-D Structure Reconstruction from Point Correspondences between Two Perspective Projections," *CVGIP-Image Understanding*, vol. 60, pp. 392–397, November 1994.

[7] A.M. Tekalp. *Digital Video Processing.* Prentice Hall, 1995.

[8] H. Stark, ed. *Image Recovery : Theory and Application.* Academic Press, 1987.

[9] K. Shoemake "Animating Rotation with Quanternion Curves," in *Proceedings of SIGGRAPH'85*, pp. 245–254, San Francisco, July 1985.

[10] M. I. Sezan and H. Stark "Incorporation of Priori Moment Information into Signal Recovery and Synthesis Problems," *Journal of Mathematical Analysis and Applications*, vol. 122, pp. 172–186, 1987.

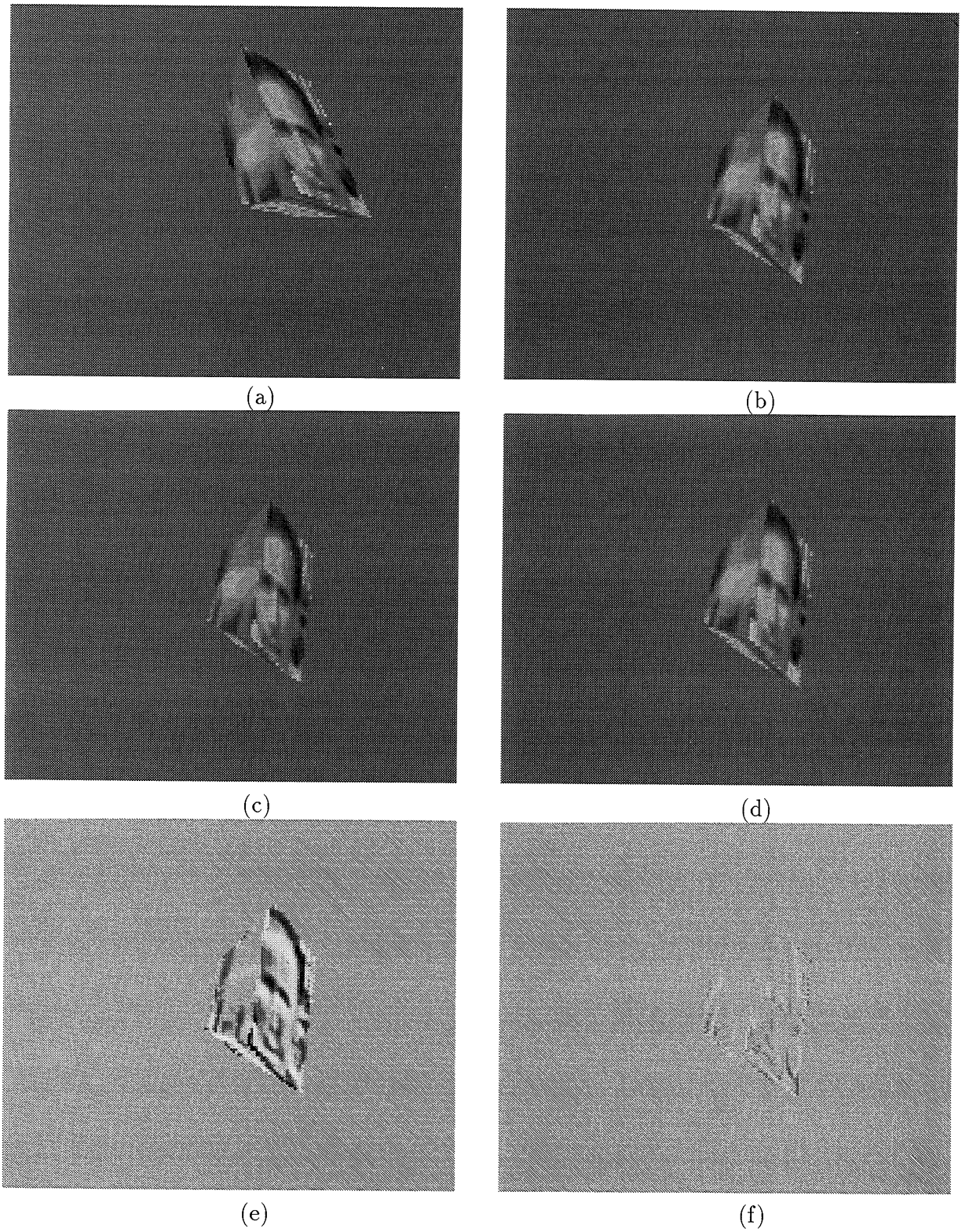

Figure 5: Results on the *Salecube* sequence. (a) Original Frame 1 and (b) original Frame 2; reconstructed Frame 2 (c) before and (d) after convex projections; frame difference between the reconstructed and original Frame 2 (e) before and (f) after convex projections.