

Conformal Geometric Algebra for 3D Object Recognition and Visual Tracking Using Stereo and Omnidirectional Robot Vision

Eduardo Bayro-Corrochano, Julio Zamora-Esquivel, and Carlos López-Franco

Computer Science Department, GEOVIS Laboratory,
Centro de Investigación y de Estudios Avanzados,
CINVESTAV, Guadalajara, Jalisco 44550, Mexico
edb@gdl.cinvestav.mx
<http://www.gdl.cinvestav.mx/~edb>

Abstract. In this paper the authors use the framework of conformal geometric algebra for the treatment of robot vision tasks. In this mathematical system we calculated projective invariants using omnidirectional vision for object recognition. We show the power of the mathematical system for handling differential kinematics in visual guided tracking.

1 Introduction

This paper shows the power of conformal geometric algebra for different tasks of robot vision. In this framework we calculate projective invariants using omnidirectional vision. These invariants are utilized for object recognition. We also treat the problem of the control of a robot binocular system which is used for 3D visual tracking. For the control strategy we utilize a novel geometric formulation of the involved Jacobian for the differential kinematics.

The rest of this paper is organized as follows: We give a brief description of the geometric algebra and also of the conformal geometric algebra in section II. In section III we explain the projective invariants. In section IV we explain the projective invariants using omnidirectional vision. Section V is devoted to the differential kinematics and control of a pan-tilt unit. The experimental analysis is given in section VI and the conclusions are in section VI.

2 Geometric Algebra

In general, a geometric algebra G_n is a n -dimensional vector space V^n over the reals. We also denote with $G_{p,q,r}$ a geometric algebra over $V^{p,q,r}$ where p, q, r denote the signature p, q, r of the algebra. If $p \neq 0$ and $q = r = 0$ the metric is Euclidean G_n , if just $r = 0$ the metric is pseudoeuclidean $G_{p,q}$ and if non of them are zero the metric is degenerate. See [3,2] for a more detailed introduction to conformal geometric algebra.

We will use the letter e to denote the vector basis e_i . In a geometric algebra $G_{p,q,r}$, the geometric product of two basis vectors is defined as

$$e_i e_j = \begin{cases} 1 & \text{for } i = j \in 1, \dots, p \\ -1 & \text{for } i = j \in p+1, \dots, p+q \\ 0 & \text{for } i = j \in p+q+1, \dots, p+q+r \\ e_i \wedge e_j & \text{for } i \neq j \end{cases} \quad (1)$$

2.1 Conformal Geometric Algebra

In the Euclidean space the composite of displacements is complicated because rotations are multiplicative but translations are additive. In order to make translations multiplicative too, we use the Conformal Geometric Algebra [3,2].

In the generalized homogeneous coordinates for points in the Euclidean space, we need that they be null vectors and also lie on the intersection of the null cone N^{n+1} (the set of all null vectors) with the hyperplane

$$P^{n+1}(e, e_0) = \{X \in R^{n+1,1} \mid e(X - e_0) = 0\}, \quad (2)$$

that is

$$N_e^n = \mathcal{N}^{n+1} \cap \mathcal{P}^{n+1}(e, e_0) = \{x \in \mathcal{R}^{n+1,1} \mid X^2 = 0, X \cdot e = -1\} \quad (3)$$

which is called the homogeneous model of \mathcal{E}^n , also called the horosphere (see Fig. 1) in hyperbolic geometry.

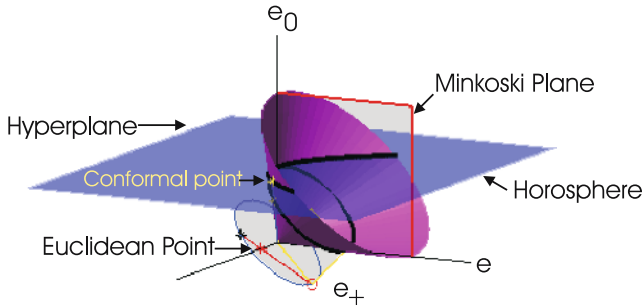


Fig. 1. Simplex at a_0 with tangent $a_1 \wedge a_2$

The points that satisfy the restrictions $X^2 = 0$ and $X \cdot e = -1$ are

$$X = \mathbf{x} + \frac{1}{2}\mathbf{x}^2 e + e_0 \quad (4)$$

where $\mathbf{x} \in \mathcal{R}^n$ and $X \in \mathcal{N}^n$. The origin is $e_0 = \frac{1}{2}(e_{n+1} - e_{n+2})$ and the point at infinity $e = e_{n+1} + e_{n+2}$.

Table 1. Entities in conformal geometric algebra

Entity	IPNS Representation	OPNS (Dual) Representation
Sphere	$S = \mathbf{p} + \frac{1}{2}(\mathbf{p}^2 - \rho^2)e + e_0$	$S^* = A \wedge B \wedge C \wedge D$
Point	$X = \mathbf{x} + \frac{1}{2}\mathbf{x}^2e + e_0$	$X^* = S_1 \wedge S_2 \wedge S_3 \wedge S_4$
Plane		$\Pi^* = A \wedge B \wedge C \wedge e$
Line		$L^* = A \wedge B \wedge e$
Circle	$Z = S_1 \wedge S_2$	$Z^* = A \wedge B \wedge C$
Point Pair	$PP = S_1 \wedge S_2 \wedge S_3$	

Note that this is a bijective mapping. From now and in the rest of the paper the conformal points will be denoted by an italic uppercase letter (X), and the Euclidean points will be denoted by boldpoint at lowercase letters \mathbf{x} .

In table 1 we show the geometric entities of the conformal geometric algebra. Note that in the IPNS representation the point is a sphere with radius zero. In the dual representation the sphere is calculated using 4 points that lie on it.

Simplexes and Conformal Points. Evaluating the outer product of r linearly independent conformal points a_0, a_1, \dots, a_r , where $r \leq n$ and n is the maximum grade of the algebra. The outer product of r conformal points is

$$a_0 \wedge a_1 \wedge \dots \wedge a_r = \mathbf{A}_r + e_0 \mathbf{A}_r^+ + \frac{1}{2} e \mathbf{A}_r^- - \frac{1}{2} E \mathbf{A}_r^\pm, \quad (5)$$

where

$$\begin{aligned}
 \mathbf{A}_r &= \mathbf{a}_0 \wedge \mathbf{a}_1 \wedge \dots \wedge \mathbf{a}_r, \\
 \mathbf{A}_r^+ &= \sum_{i=0}^r (-1)^i \mathbf{a}_0 \wedge \dots \wedge \check{\mathbf{a}}_i \wedge \dots \wedge \mathbf{a}_r = (\mathbf{a}_1 - \mathbf{a}_0) \wedge \dots \wedge (\mathbf{a}_r - \mathbf{a}_0), \\
 \mathbf{A}_r^- &= \sum_{i=0}^r (-1)^i \mathbf{a}_i^2 \mathbf{a}_0 \wedge \dots \wedge \check{\mathbf{a}}_i \wedge \dots \wedge \mathbf{a}_r, \\
 \mathbf{A}_r^\pm &= \sum_{i=0}^r \sum_{j=i+1}^r (-1)^{i+j} (\mathbf{a}_i^2 - \mathbf{a}_j^2) \mathbf{a}_0 \wedge \dots \wedge \check{\mathbf{a}}_i \wedge \dots \wedge \check{\mathbf{a}}_j \wedge \dots \wedge \mathbf{a}_r.
 \end{aligned} \quad (6)$$

Note that A_r is the moment of the simplex with tangent (boundary) A_r^+ . The outer product $a_0 \wedge a_1 \wedge \dots \wedge a_r$ represents a sphere when $\mathbf{A}_r = 0$

$$a_0 \wedge a_1 \wedge \dots \wedge a_r = -[e_0 - \frac{1}{2} e \mathbf{A}_r^- (\mathbf{A}_r^+)^{-1} + \frac{1}{2} \mathbf{A}_r^\pm (\mathbf{A}_r^+)^{-1}] E \mathbf{A}_r^+ \quad (7)$$

where the center and radius of the sphere

$$c = \frac{1}{2} \mathbf{A}_r^\pm (\mathbf{A}^+)^{-1}, \quad \rho^2 = c^2 + \mathbf{A}_r^- (\mathbf{A}^+)^{-1}. \quad (8)$$

3D Rigid Motion. In conformal geometric algebra we can perform rotations by means of an entity called rotor which is defined by

$$R = \exp\left(\frac{\theta}{2}\mathbf{l}\right), \quad (9)$$

where \mathbf{l} is the bivector representing the dual of the rotation axis. To rotate an entity, we simply multiply it by the rotor R from the left and the reverse of the rotor \tilde{R} from the right,

$$Y = RX\tilde{R}. \quad (10)$$

If we want to translate an entity we use a translator which is defined as

$$T = \left(1 + \frac{et}{2}\right) = \exp\left(\frac{\mathbf{et}}{2}\right). \quad (11)$$

With this representation the translator can be applied multiplicatively to an entity similarly to the rotor, by multiplying the entity from the left by the translator and from the right with the reverse of the translator,

$$Y = TX\tilde{T}. \quad (12)$$

Finally, the rigid motion can be expressed using a *motor* which is the combination of a rotor and a translator

$$M = TR,, \quad (13)$$

thus the rigid body motion of an entity is described with

$$Y = MX\tilde{M}. \quad (14)$$

Also a motor can be defined using the exponential representation with a line representing its axis

$$M = \exp\left(\frac{-\theta}{2}I_C L^*\right), \quad (15)$$

note that the line must be normalized to one.

3 Invariants

An invariant is a property that remains unchanged under certain class of transformation. Within the context of vision, we are interested in determining the invariants of an object under perspective projection. The cross-ratio of four collinear points is a well known 1D-invariant under projective transformations but it can be extended to 2D, so we can use it for image invariants. In the 2D case we need five points in the 3D case we need six points. In the 3D space these invariants can be interpreted as the cross-ratio of tetrahedral volumes.

Now, for the 2D case we need five points, an example of a 2D invariant is

$$Inv_2 = \frac{(\mathbf{X}_5 \wedge \mathbf{X}_4 \wedge \mathbf{X}_3) I_{p2}^{-1} (\mathbf{X}_5 \wedge \mathbf{X}_2 \wedge \mathbf{X}_1) I_{p2}^{-1}}{(\mathbf{X}_5 \wedge \mathbf{X}_1 \wedge \mathbf{X}_3) I_{p2}^{-1} (\mathbf{X}_5 \wedge \mathbf{X}_2 \wedge \mathbf{X}_4) I_{p2}^{-1}}, \quad (16)$$

where $I_{p2} = e_1 \wedge e_2 \wedge e_-$ denotes the pseudoscalar of the 2D projective space.

If we use conformal points the outer product of three points leads to a circle, so with four circles we can compute the 2D invariants. Also note that we use the A_r (6) part of the circle (the moment of the simplex) to calculate the invariant.

$$C_1 = X_5 \wedge X_4 \wedge X_3, C_2 = X_5 \wedge X_2 \wedge X_1, \quad (17)$$

$$C_3 = X_5 \wedge X_1 \wedge X_3, C_4 = X_5 \wedge X_2 \wedge X_4. \quad (18)$$

Let $A_{r,k}$ denote the A_r part of the k -circle C_k where $k = 1 \dots 4$. Then the invariant using the moment A_r of the simplex is

$$Inv_2 = \frac{A_{r,1} I_E^{-1} A_{r,2} I_E^{-1}}{A_{r,3}^+ I_E^{-1} A_{r,4}^+ I_E^{-1}}. \quad (19)$$

4 Invariants and Omnidirectional Vision

The projective invariants do not hold in the catadioptric image, but they do in the image sphere. Therefore we must take some points in the catadioptric image and project them to the sphere. Once we do this we can proceed to calculate the invariants using four circles.

First we will show briefly that projective invariants in the plane are equivalent to projective invariants in the S^2 sphere (image sphere), see Fig. 2. According our previous work [1] we define the point F (in this case it will be equal to e_0), then the unit sphere is

$$S = e_0 - \frac{1}{2}e. \quad (20)$$

Now, let x_1, x_2, \dots, x_5 be points in the Euclidean space with conformal representation

$$X_i = \mathbf{x}_i + \frac{1}{2}\mathbf{x}_i^2 e + e_0, \text{ for } i = 1 \dots 5. \quad (21)$$

Then we project the points in the space to the sphere and that give us the projected points say U_1, U_2, \dots, U_5 .

In the other hand, the image plane Π_I (in order to compare the invariants) is defined as

$$\Pi_I = e_2 + e. \quad (22)$$

We project first the points to the plane and then we intersect the plane with each line

$$Q_i = L_i^* \cdot \Pi_I \text{ for } i = 1 \dots 5. \quad (23)$$

The point Q_i is a *flatpoint* which is the outer product of a conformal point with the null vector e (the point at infinity). To obtain the conformal point from the *flatpoint* we can use

$$V_i = \frac{Q_i \wedge e_0}{(-Q_i \cdot E)E} + \frac{1}{2} \left(\frac{Q_i \wedge e_0}{(-Q_i \cdot E)E} \right)^2 e + e_0 . \quad (24)$$

Using (18) we calculate the two sets of four circles, one for the points U_i and one for V_i . With each set of circles we calculate the two invariants using (19), after comparing this two invariants we will see that them are the same. Therefore, we now know that if we project the points in the catadioptric image to the sphere we have again the projective invariants.

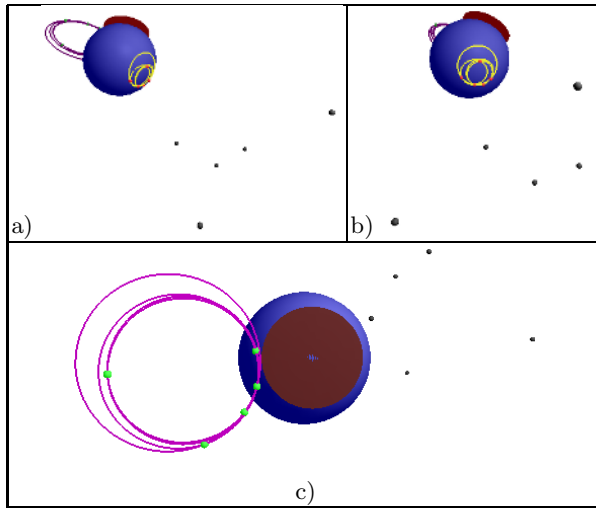


Fig. 2. Different views of points in the space projected to the (image) sphere and to the (image) plane used to compare the calculated invariants. a) Global view of points projected to the sphere and to the plane, b) Points projected in the sphere with the circles formed to calculate the invariants and c) Points projected in the plane with its circles formed to calculate the invariants.

We have seen a brief introduction to several topics necessary to understand the experimental results. In the next section we will see an application of the given theory.

5 Differential Kinematic Control for a Pan-Tilt Unit

We will show an example using our formulation of the Jacobian. This is the control of a pan-tilt unit.

5.1 System

We can implement velocity control for a pan-tilt unit (PTU Fig. 3.a) easily assuming three degree of freedom (we call it virtual component), the PTU has similar kinematic behavior as a robot of three D.O.F.

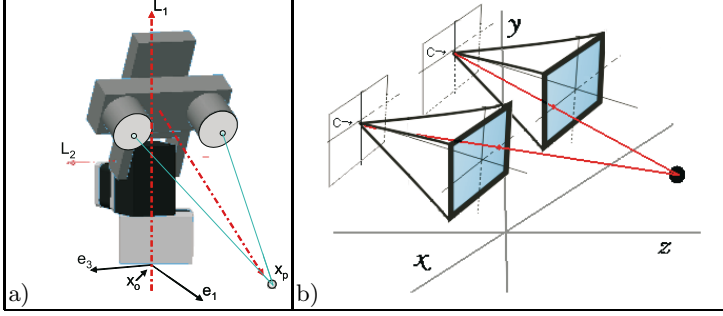


Fig. 3. a) Binocular stereo system fastened on a pan tilt unit. b) Abstraction of the stereo system.

In order to carry out a velocity control, we need first to compute the direct kinematics, this is very easy to do, because we know the axis lines:

$$L_1 = -e_{31}, \quad L_2 = e_{12} + d_1 e_1 e_\infty, \quad L_3 = e_1 e_\infty. \quad (25)$$

Since $M_i = e^{-\frac{1}{2}q_i L_i}$ and $\widetilde{M}_i = e^{\frac{1}{2}q_i L_i}$, we can compute the position of end effector as:

$$x_p(q) = x'_p = M_1 M_2 M_3 x_p \widetilde{M}_3 \widetilde{M}_2 \widetilde{M}_1, \quad (26)$$

The estate variable representation of the system is as follows

$$\begin{cases} \dot{x}'_p = x' \cdot (L'_1 \ L'_2 \ L'_3) \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} \\ y = x'_p \end{cases} \quad (27)$$

where the position of end effector at home position x_p is the conformal mapping of $x_{p_e} = d_3 e_1 + (d_1 + d_2) e_2$, the line L'_i is the current position of L_i and u_i is the velocity of the i -junction of the system. As L_3 is an axis at infinity M_3 is a translator, that is, the virtual component is a prismatic junction.

5.2 Linearization Via Feedback

Now the following state feedback control law is chosen in order to get a new linear an controllable system.

$$\begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = (x'_p \cdot L'_1 \ x'_p \cdot L'_2 \ x'_p \cdot L'_3)^{-1} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \quad (28)$$

Where $V = (v_1, v_2, v_3)^T$ is the new input to the linear system, then we rewrite the equations of the system

$$\begin{cases} \dot{x}'_p = V \\ y = x'_p \end{cases} \quad (29)$$

5.3 Asymptotic Output Tracking

The problem of follow a constant reference x_t is solved computing the error between end effector position x'_p and the target position x_t as $e_r = (x'_p \wedge x_t) \cdot e_\infty$, the control law is then given by.

$$V = -ke \quad (30)$$

This error is small if the control system is doing it's job, it is mapped to an error in the joint space using the inverse Jacobian.

$$U = J^{-1}V \quad (31)$$

Computing the Jacobian $J = x'_p \cdot (L'_1 \ L'_2 \ L'_3)$

$$j_1 = x'_p \cdot (L_1), \ j_2 = x'_p \cdot (M_1 L_2 \widetilde{M}_1), \ j_3 = x'_p \cdot (M_1 M_2 L_3 \widetilde{M}_2 \widetilde{M}_1) \quad (32)$$

Once that we have the Jacobian is easy to compute the dq_i using Crammer's rule.

$$\begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = (j_1 \wedge j_2 \wedge j_3)^{-1} \cdot \begin{pmatrix} V \wedge j_2 \wedge j_3 \\ j_1 \wedge V \wedge j_3 \\ j_1 \wedge j_2 \wedge V \end{pmatrix} \quad (33)$$

This is possible because $j_1 \wedge j_2 \wedge j_3 = \det(J)I_e$. Finally we have dq_i which will tend to reduce these errors. Due to the fact that the Jacobian has singularities then we should use the pseudo inverse of Jacobian.

5.4 Pseudo-Inverse of Jacobian

To avoid singularities we compute the pseudo inverse of Jacobian matrix $J = [j_1 \ j_2]$. Using the pseudo-inverse of Moore-Penrose

$$J^+ = (J^T J)^{-1} J^T \quad (34)$$

Now evaluating J in (34)

$$J^+ = \frac{1}{\det(J^T J)} \begin{pmatrix} (j_2 \cdot j_2)j_1 - (j_2 \cdot j_1)j_2 \\ (j_1 \cdot j_1)j_2 - (j_2 \cdot j_1)j_1 \end{pmatrix} \quad (35)$$

And Using Clifford algebra we could simplify further this equation

$$\det(J^T J) = (j_1 \cdot j_1)(j_2 \cdot j_2) - (j_1 \cdot j_2)^2 = (|j_1||j_2|)^2 - (|j_1||j_2|)^2 \cos^2(\theta), \quad (36)$$

$$= (|j_1||j_2|)^2 \sin^2(\theta) = |j_1 \wedge j_2|^2 \quad (37)$$

calling θ the angle between vectors. By the way each row of J^+ could be simplify as follows: $(j_2 \cdot j_2)j_1 - (j_2 \cdot j_1)j_2 = j_2 \cdot (j_2 \wedge j_1)$ and $(j_1 \cdot j_1)j_2 - (j_2 \cdot j_1)j_1 = j_1 \cdot (j_1 \wedge j_2)$.

Now the equation (34) can be rewritten as

$$J^+ = \frac{1}{|j_1 \wedge j_2|^2} \begin{pmatrix} j_2 \cdot (j_2 \wedge j_1) \\ j_1 \cdot (j_1 \wedge j_2) \end{pmatrix} = \begin{pmatrix} j_2 \cdot (j_2 \wedge j_1)^{-1} \\ j_1 \cdot (j_1 \wedge j_2)^{-1} \end{pmatrix} \quad (38)$$

Using this equation we can compute the input as $U = J^+ V$ that is equal to

$$U = (j_1 \wedge j_2)^{-1} \cdot \begin{pmatrix} V \wedge j_2 \\ j_1 \wedge V \end{pmatrix} \quad (39)$$

5.5 Visual Tracking

The target point is calculate using two calibrated cameras (see Figure 3.b), on each camera we estimate the center of mass of the object in movement in order to do a retroprojection and estimate the 3D point. to compute the mass center first we subtract the current image I_c to an image in memory I_a , the image in memory is the average of the last N images, this help us to eliminate the background.

$$I_k(t) = I_c(t) - I_a(t-1) * N, \quad I_a(t) = (I_a(t-1) * N + I_c)/(N+1) \quad (40)$$

After that the moment of x and y is computed and they are divided by the mass (pixels in movement) that is, the intensity difference between the current image and the image on memory give us the mass center.

$$x_o = \frac{\int_0^n \int_0^m I_k y dx dy}{\int_0^n \int_0^m I_k dx dy}, \quad y_o = \frac{\int_0^n \int_0^m I_k x dx dy}{\int_0^n \int_0^m I_k dx dy} \quad (41)$$

When the camera moves the background changes and its necessary to reset N to 0 to restart the process of track.

6 Experimental Results

In this section we present two experiments: the first illustrates the use of the theory of invariants and omnidirectional vision for object recognition and the second the control of a binocular head for tracking.

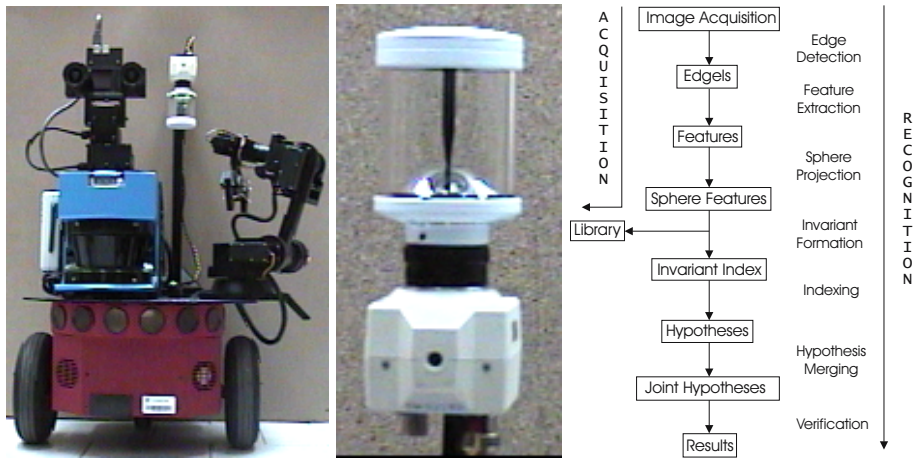


Fig. 4. a) Mobile robot. b) Omnidirectional vision system. c) Recognition procedure.

6.1 Object Recognition

The omnidirectional image has the advantage of a bigger field of view, see Fig. 4.a-b. This capability allows to see all the objects around the robot without moving it. In contrast to the stereo system, which does not see all the objects or in some cases none of them (see Fig. 5).

Before we use the omnidirectional system we must calibrate it with this we mean find the mirror center, focal length, skew and aspect ratio. The objective of the experiment is that the robot should recognize an object from different objects lying on three tables located around the robot. The recognition process consists of various steps that are show in Fig. 4.c .

To recognize an object we first take features from the catadioptric image, then these features are projected onto the unit sphere. With this features in the sphere we calculate the circles formed with them (see Eq. 18). Finally, the invariants are calculated with Eq. 19 which are equivalent to the projective invariants. These invariants are compared with the previously acquired invariants in the library to identify the object. The key points of an object are selected by hand. If they are accurate enough, our procedure can recognize the objects correctly. In general this kind of invariants are a bit sensitive to noise, due to the illumination changes and computations. In order to diminish the effect of noise in the data, we can compute several invariants related with the object, so that the accuracy of the recognition is increased. Utilizing an automatic corner detector the procedure of object recognition using our method can be carried out in real time.

Once that the object is recognized we rotate the robot until the object is in front of the stereo system. Since the object is now visible to the stereo camera, we can use an inverse kinematic approach to grasp the object. In our case we chose for the approach of [4] which is very interesting. Such approach models the

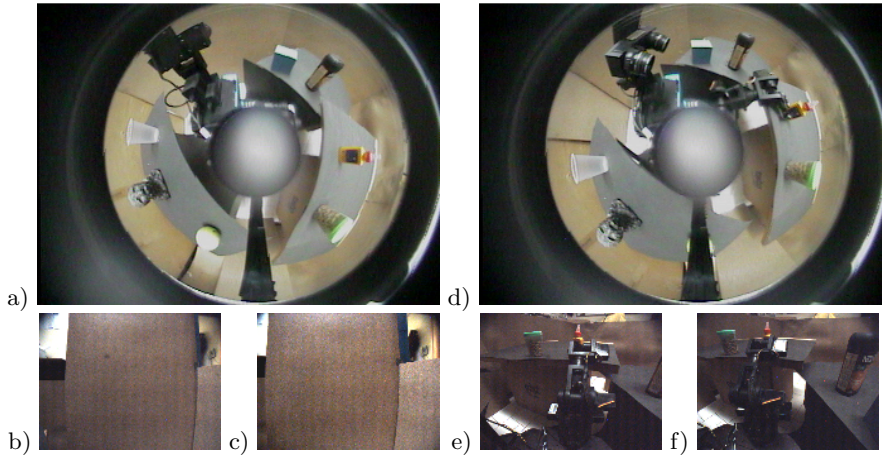


Fig. 5. Initial state of the experiment: a) Omnidirectional view, b-c) Left and right images of the stereo system (out of target). Robot grasps an object: d) Omnidirectional view, e-f) Left and right images of the stereo system (looking at the target).

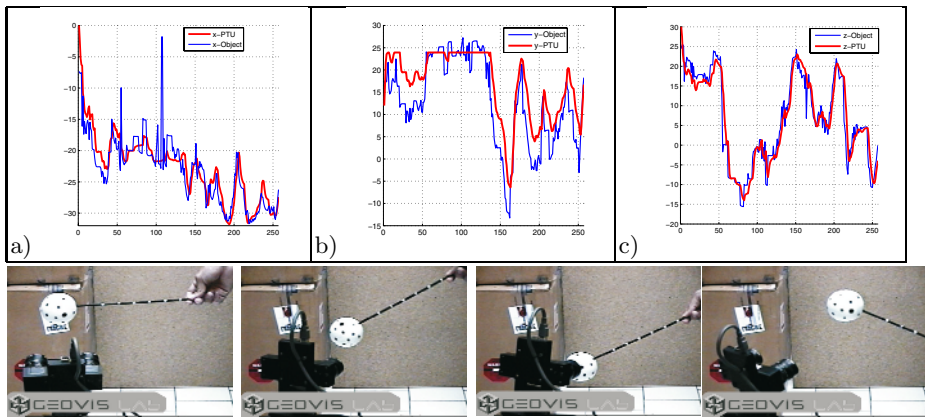


Fig. 6. (Upper row) Velocity components: a) x; b) y; c) z, (the rough curves are of the 3D object motion). (Lower row) Some views of a tracking sequence.

joints of the robot arm using spheres, circles, lines and planes which are entities very easy to handle in conformal geometric algebra. In Figures 5.d-f we show the robot grasping an object.

6.2 Visually Controlled Tracking

In Figure 6 we can appreciate the smooth trajectory of the tracking. The rough behavior of the 3D object motion is compensated by a PD controller using our

geometric Jacobian approach. Note that 3D motion of the pan-tilt unit is not disturbed by the big peaks of the 3D object motion.

7 Conclusions

In this article we have chosen the coordinate-free system of conformal geometric algebra for the design of algorithms useful for robot perception and action. In this framework we calculate the invariants of circles in the sphere and used them to recognize objects with the advantage of the bigger field of view offered by the omnidirectional vision system. We also showed an interesting application of 3D tracking using a new formulation of a geometric Jacobian for the differential kinematics.

Acknowledgment

We are thankful to CONACYT Proyecto 49 Fondos de Salud for supporting this work.

References

1. Bayro-Corrochano E. and Lpez-Franco C. [2004]. Omnidirectional vision: unified model using conformal geometry. In Proc. of the 8th European Conf. on Computer Vision, ECCV2004, Prage, Czech Republic, Part I, pp. 536-548.
2. Bayro-Corrochano E. [2005]. Robot perception and action using conformal geometric algebra. In the *Handbook of Geometric Computing. Applications in Patter Recognition, Computer Vision, Neuralcomputing and Robotics*, Eduardo Bayro-Corrochano (ed.), Chap. 14, Springer Verlag, Heidelberg, pp 405-457.
3. H. Li, D. Hestenes and A. Rockwood. [2001] Generalized homogeneous coordinates for computational geometry. In *Geometric Computing with Clifford Algebra*, G. Sommer (Ed.), Springer-Verlag, pp. 27-59.
4. Zamora-Esquivel J. and Bayro-Corrochano E. [2005] Static and differential geometry of robot devices using conformal computational geometry. Submitted elsewhere.