A new Gesture Representation for Sign language analysis

Boris Lenseigne*, **Frédérick Gianni***, **Patrice Dalle***

*IRIT
Université Paul Sabatier 118 route de Narbonne 31062 Toulouse cedex 4
{lenseign,gianni,dalle}@irit.fr

Abstract

Computer aided human gesture analysis requires a model of gestures, and an acquisition system which builds the representation of the gesture according to this model. In the specific case of computer Vision, those representations are mostly based on primitives described from a perceptual point of view, but some recent issues in Sign language studies propose to use a proprioceptive description of gestures for signs analysis. As it helps to deal with ambiguities in monocular posture reconstruction too, we propose a new representation of the gestures based on angular values of the arm joints based on a single-camera computer vision algorithm.

1. Gesture representation

1.1. Previous work

Most of the descriptions of Sign language vocabulary relies on linguistic studies and are those used in notation or transcription systems, SignWriting[1], Hamnosys[2] (Prillwitz and al., 1989). In the case of computer aided Sign language analysis, we distinguish systems using a specific hardware such as data gloves (Braffort, 1996)(Starner T., 1998)(Vogler C., 1999) and those using cameras. Data gloves based applications process directly on values provided by the sensors. in he case of a computer vision system, gesture model cans be bidimentionnal or tridimentionnal. When several cameras are used (Wren C., 1999)(Vogler C., 1998)(Somers M.G., 2003) $3D$ reconstruction is possible and gestures can be analyzed directly in $3D$ space. In the case of a single camera, gesture analysis can be performed directly in $2D$ images (Starner T., 1998)(Tanibata N., 2002) or some additional image processing has to be performed for a partial $3D$ estimation. In this case, visual aspect of gestures is deduced from $3D$ models (Horain P., 2002)(Athitsos V., 2004), or $3D$ model is used to constrain the reconstruction (Lenseigne B., 2004). Both solutions leading to ambiguities.

Thus notation systems and vision-based gesture analysis systems use a representation of signs derived from a tridimentionnal perceptive description. Gestures are located in a speaker-centered frame (ScF) (fig. 5) but described from external point of view. Those descriptions are based on the definition of a set of elementary motions and for each elementary motion a set of parameters. Widely used primitives are, straight, curve and complex motions. Such a classification is only suitable for standard vocabulary description, leads to a classification of gestures in terms of geometrical primitives and to a description of gestures from the observer's point of view.

1.2. A new gesture representation

A different way to represent gesture is to use a proprioceptive point of view. In such a case, motion analysis and classification rely on the way gesture is performed. This approach is presented in recent linguistic research (Boutet, 2001) which suggests that an articulation-based representation may have appropriate properties to allow the representation of the function of the gesture. So that, using joint values to represent gesture is an interesting choice. This assumption leads us to propose a method, based on a single camera to compute a gesture representation based on joint angle evolution in time.

2. Computing articulation values from a single image

Articulations values calculation is performed in two stages : a geometrical reconstruction of the $3D$ posture of the arm and the computation of corresponding articulations values. As we use a single camera, a direct $3D$ reconstruction of the arm is not possible, and the geometrical method provides us with a set of four possible configuration of the arm in a given image. A configuration is represented by the $3D$ Cartesian coordinates of each joint (shoulder, elbow and wrist). Those coordinates are grouped together to form a set of four possible motions for the arm and joint values can be computed for each trajectory to build articulation-based motion representation.

2.1. Geometric resolution

In this section we describe how to reconstruct a set of possible $3D$ pose of a human arm from a single picture using a single calibrated camera. This configuration is defined by the position of each segment's limits (shoulder, elbow and wrist) in Cartesian coordinates. Given an image, we are able to reduce the space of possible poses for the arm to four configurations only using a simple model of the scene, the camera and some assumption about it.

2.1.1. Requirements

Our technique is based on several assumptions, which may be crippling under an uncontroled environment. However they could be raised if reconstruction can be performed with a scale factor which does not affect joint values computation.

[1] `http://www.signwriting.org/`

[2] `http://www.sign-lang.uni-hamburg.de/Projects/HamNoSys.html`

Acquisition device : The acquisition device is made up of a single camera, which has been calibrated in order to be able to calculate the equation of the projective ray across a given pixel, which suppose that the perspective transformation matrix $\mathbf{C}$ is known. Many techniques of calibration were previously proposed, for instance in (Gurdjos P., 2002) or (Heikkilä, 2000).

Tracking the articulations : We also make the assumption that we are able to identify the $2D$ positions of the three articulations of the arm in the image. Tracking techniques abound and depend on the problem to solve (degree of homogeneity of the background, the use of markers, motion models, etc...). The precision needed for tracking depends on the precision needed for reconstruction. A study of the influence of tracking errors on reconstruction can be found in (Lenseigne B., 2004).

Arm pose : We only speak here about rebuilding the posture of an arm, without considering the hand. Within the framework of a geometric resolution, we define the posture by the position in space of the articulations (shoulder, elbow, wrist), i.e. if coordinates are expressed in the camera frame :

- for the shoulder : $P_1 = (X_1, Y_1, Z_1)^T$
- for the elbow : $P_2 = (X_2, Y_2, Z_2)^T$
- for the wrist : $P_3 = (X_3, Y_3, Z_3)^T$

Using this representation, the estimating of the posture of the arm is reduced to the calculation of three points in space.

2.1.2. Geometrical model of the arm

The arm is modeled by the articular system connecting the shoulder to the wrist. This system consists of articulations (a ball-and-socket joint for the shoulder and a revolving joint for the elbow) connecting rigid segments (arm and forearm) noted l_i, the segment l_i connecting the articulations P_i and P_{i+1}. The position of the final body corresponds to the wrist position, i.e. with P_3, end of the segment l_2. Since those articulations allows only pure rotations of the segment l_i around the articulation P_i, we can define the set of the reachable positions by the articulation $P_j (j = 2, 3)$ as a sphere, centered on the preceding articulation P_{j-1} and whose ray is $\|l_i\|$[3] (cf. figure 1). Using

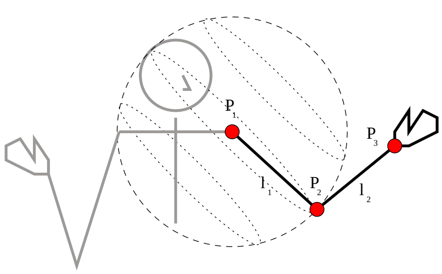

Figure 1: Model of the articular system of the arm. The sphere represents the set of the possible positions for the elbow.

this model, the reachable space for each articulation position becomes a sphere whose parameters are known if we determine the position of the preceding articulation and the length of each segment of the arm, which means that we have to know, from the beginning, the $3D$ position of the shoulder. This can be problematic in an uncontrolled environment. However, when the problem is to obtain qualitative or relational values, or for angular values calculation, a reconstruction with a scale factor can be sufficient. The position of the shoulder could then be fixed as a preliminary. Identically, dimensions of each segment can be fixed arbitrarily as long as the ratio of their respective lengths is respected.

2.1.3. Algorithm

The method we present exploits a simple geometrical model of the scene and especially of the structure of the arm. We suppose that the coordinates of the points corresponding to the articulations are known, in the image. They can be writen in homogeneous coordinates as :

- for the shoulder : $\widetilde{\mathrm{p}_1} = (u_1, v_1, 1)^T$
- for the elbow : $\widetilde{\mathrm{p}_2} = (u_2, v_2, 1)^T$
- for the wrist : $\widetilde{\mathrm{p}_3} = (u_3, v_3, 1)^T$

After the calibration of the camera, we can compute for each point in the image the associated projection ray, which is the line (passing by the optical center and the point image considered) containing the $3D$ counterpart of this point.

Set of possible configurations for the elbow : knowing P_1 the (possibly arbitrary) position of the shoulder in space, the set of possible positions for the elbow can be defined as the sphere S_1 centered on the shoulder and whose ray is the length $\|l_1\|$ of the arm. The Cartesian equation of such a sphere is :

$$(X_1 - x)^2 + (Y_1 - y)^2 + (Z_1 - z)^2 - \|l_1\|^2 = 0 \quad (1)$$

Equation of the projection ray : $\widetilde{\mathrm{p}}_1$ is the position of the shoulder in the image, expressed in homogeneous coordinates. The calibration of the camera gives us the perspective transformation C matrix defining the transformation from a $3D$ frame associated to the camera[4], to the $2D$ image frame [5]. The matrix defining the perspective transformation which forms the image is traditionnaly written as follows :

$$\mathbf{C} = \begin{bmatrix} fk_u & 0 & u_0 \\ 0 & fk_v & v_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (2)$$

Where:

- f is the focal length ;
- k_u, k_v are scale factor, horizontal and vertical, in $pixels/mm$
- (u_0, v_0) is the position of the principal point in the image frame (the projection of the optical center of the camera).

[3] $\|l_i\|$ is the norm of the segment l_i

[4] the origin of this frame is in the optical center

[5] the origin of the image frame is in the left higher corner of the image

This matrix let us deduct the position in the image frame of a point $\widetilde{\mathrm{p}}_{\mathrm{i}} = (u_i, v_i, 1)^T$ projection of a point whose coordinates are expressed in the camera frame $P_i = (X_i, Y_i, Z_i)^T$:

$$\begin{bmatrix} u_i \\ v_i \\ 1 \end{bmatrix} = \begin{bmatrix} fk_u & 0 & u_0 \\ 0 & fk_v & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_i \\ Y_i \\ Z_i \end{bmatrix} \quad (3)$$

The inverse of this matrix is used to calculate, for each point p_i in the image, the equation of the associated projection ray in space. The projection ray is the line passing through the focal point of the camera and the considered point in the image plane. The original $3D$ point is necessarily located on this line. Here is a parametric equation of the projection ray, where λ is a simple multiplying coefficient :

$$R_i(\lambda) = \lambda \widetilde{\mathbf{p}}_{\mathbf{i}} \quad (4)$$

$\widetilde{\mathbf{p}}_{\mathbf{i}}$ represents the coordinates of the image point in the camera frame :

$$\widetilde{\mathbf{p}}_{\mathbf{i}} = C^{-1}\widetilde{\mathrm{p}}_{\mathrm{i}} \text{ with } C^{-1} = \begin{bmatrix} \frac{1}{fk_u} & 0 & \frac{u_0}{fk_u} \\ 0 & \frac{1}{fk_v} & \frac{v_0}{fk_v} \\ 0 & 0 & 1 \end{bmatrix} \quad (5)$$

$$\text{So that : } \widetilde{\mathbf{p}}_{\mathbf{i}} = \begin{bmatrix} \frac{(u_i-u_0)}{fk_u} \\ \frac{(v_i-v_0)}{fk_v} \\ 1 \end{bmatrix} \quad (6)$$

Therefore the $3D$ position we search is the intersection of the surface of the sphere S_1 defining the set of the possible configurations for the elbow, and the projection ray $r_i(\lambda)$. Calculation of those intersections in the camera frame consists in determining values for λ such as :

$$\begin{aligned} &(X_1 - \lambda\tfrac{(u_i-u_0)}{fk_u})^2 + (Y_1 - \lambda\tfrac{(v_i-v_0)}{fk_v})^2 \\ &+(Z_1-\lambda)^2 - \|l_1\|^2 = 0 \end{aligned} \quad (7)$$

This is a second degree polynomial $a\lambda^2 + b\lambda + c = 0$ whose coefficients are :

$$\begin{aligned} a &= (\tfrac{(u_i-u_0)}{fk_u})^2 + (\tfrac{(v_i-v_0)}{fk_v})^2 + 1; \\ b &= 2[(\tfrac{(u_i-u_0)}{fk_u})(-X_1) + (\tfrac{(v_i-v_0)}{fk_v})(-Y_1) - Z_1]; \\ c &= X_1^2 + Y_1^2 + Z_1^2 - l_1^2 \end{aligned} \quad (8)$$

Solving this polynomial gives two possible values for λ, possibly a single double one, the positions $\widehat{p}_{2,j} (j = 1, 2)$ possible for the elbow comes now directly since $r(\lambda) = \lambda\widetilde{\mathbf{p}}_{\mathbf{i}}$.

Using the same technique, we are able to calculate the possible positions $\widehat{p}_{3,j} (j = 1..4)$ of the wrist, considering the two spheres whose centers are given by the estimated positions of the elbow and rays by the length of the forearm. We can calculate for each value of the position of the elbow two possible positions for the wrist and thus four possible configurations for the arm.

This algorithm allows us to reduce the set of possible configurations for an arm to four possibilities for a single image. Elbow's positions are symmetric in regard of a plane parallel to the image and containing the shoulder. Calculation of the wrist's position is performed from each possible elbow position so that we obtain four possible positions for the wrist. In the same way as for the elbow, each couple of solutions is symmetric in regard of a plane parallel to the image and containing the corresponding elbow position.

2.2. Extension to image sequences analysis

In the case of image sequences, we calculate a set of $3D$ points candidates for each image. During the sequence those points have to be merged to build trajectories. For each branch of the solution tree (except particular configuration) there are two points to assign to a pair of trajectories. Since it is not possible to know directly which point must be attached to a given trajectory, we introduce a linearity criterion : we calculate the angle α between vectors $\overrightarrow{\mathcal{V}}_{i,j,k}$ and $\overrightarrow{\mathcal{V}}_{i,j,k+1}$, where $\overrightarrow{\mathcal{V}}_{i,j,k}$ is defined by points $\widehat{P}_{i,j,k-1}$ and $\widehat{P}_{i,j,k}$, and $\overrightarrow{\mathcal{V}}_{i,j,k+1}$ by points $\widehat{P}_{i,j,k}$ and $\widehat{P}_{i,j,k+1}$. $\widehat{P}_{i,j,k}$ is the $j^{th}(j = 1, 2)$ estimated space coordinates of the articulation i in the k^{th} image of the sequence. We must therefore calculate the norm of the cross product $\|\overrightarrow{\mathcal{V}}_{i,j,k} \wedge \overrightarrow{\mathcal{V}}_{i,j,k+1}\| = \|\overrightarrow{\mathcal{V}}_{i,j,k}\|\|\overrightarrow{\mathcal{V}}_{i,j,k+1}\| sin(\alpha)$. (figure 2).

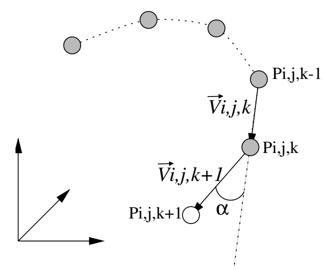

Figure 2: Building trajectories : we first compute the cross product between the last guiding vector of the current trajectory and the new one build by using the (white) candidate point. Linearity criterion consists in merging to the current trajectory the point which minimises the norm of this cross product.

The candidate for which the norm is weakest is affected to the corresponding trajectory. The second point of the branch is then affected to the other one.

Particular configurations : The construction of the trajectories described above can be done correctly in the general case where the algorithm gives two intersections between the projection ray and the sphere. However there are configurations where this assumption is false. Those configurations must be taken into account in the algorithm ; they can also be used as detector for particular movements. There are two categories of particular configurations:

1. The polynomial (8) has only a single solution. It happens when the considered segment (arm or forearm) is included in a plane parallel to the image plane. In this case, the projection ray is tangent to the sphere and there will only be a single "intersection" with the sphere. This point is then added to the both trajectories : it indeed corresponds to a case where the two possible trajectories will cross.

2. The polynomial (8) does not have any solution. In the absence of noise, this case can occur only for the wrist : after having calculated the two possible positions for the elbow, we define the pair of spheres which forms the set of the possible positions of the wrist. There are cases where, based on the "wrong" position of the elbow, the sphere does not have any intersection with the

i	θ_i	d_i	α_i	a_i
1	θ_1	0	$-\pi/2$	0
2	θ_2	0	$\pi/2$	0
3	θ_3	l_1	$-\pi/2$	0
4	θ_4	0	$\pi/2$	0
5	0	l_2	0	0

Table 1: DH parameters describing the human arm system

projection ray. Those configurations directly allows us to cut a complete branch from the solution tree.

2.2.1. Angular values calculation

The parametric model of the human arm is based on the modified Denavit-Hartenberg (DH) parameters description (Denavit J., 1955). This representation provides a systematic method for describing relationships between adjacent links. As long as the frames attached to each articulation are positionned using DH algorithm (Cf. 2.2.1.. The model consists in a $4x4$ homogeneous transformation matrix corresponding to the transformation from link 1 to link 3, which describes, in fact, the arm system. This matrix is parametrized with angular values of each joint and link lengths. This matrix constitute the direct geometrical model. Whereas the inverse geometrical model provides the joint angular values in function of the joint Cartesian coordinates.

Modified parameters of Denavit-Hartenberg : The DH method is systematic as long as the axis system R_i attached to each joint (figure 3) is defined using the following rules :

1. O_{i-1} is the perpendicular common to link L_{i-1} and L_i axes located on link L_{i-1} ;

2. axis x_{i-1} is the unit vector of the common perpendicular oriented from link L_{i-1} to link L_{i-1} ;

3. z_i is the unit vector of link L_i ;

4. axis y_i is set so that : $y_i = z_i \wedge x_i$

5. relationships between frame R_i and R_{i-1} are defined by the following parameters :

   - α_i is the offset angle from axis z_{i-1} to z_i around x_{i-1} ;
   - d_i : the distance from the origin of the $(i-1)^{th}$ coordinate frame to the intersection of the z_{i-1} axis with the x_i ;
   - θ_i : the joint angle from x_{i-1} to x_i turning around z_i ;
   - a_i : the offset distance from the intersection of the z_{i-1} axis with the x_i axis.

With the joint frame O and 1 jointed, the arm model is given by the D-H parameters is shown in table 1.

DH parameters are used to write an homogeneous transformation matrix for each joint. The generic form of the matrix for a revolving joint is :

$$ {}^{i-1}T_i = \begin{bmatrix} \cos\theta_i & -\sin\theta_i & 0 & a_i \\ \cos\alpha_i \sin\theta_i & \cos\alpha_i\cos\theta_i & -\sin\alpha_i & -d_i\sin\alpha_i \\ \sin\alpha_i\sin\theta_i & \sin\alpha_i & \cos\theta_i & d_i\cos\alpha_i \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (9) $$

Where $\theta_i, \alpha_i, d_i, a_i$ are the DH parameters.

Direct geometrical model : The direct geometrical model gives the transformation from Cartesian coordinate space to angular values of the each joint. The $4x4$ matrix ${}^0T_5$ specifies the homogeneous transformation from frame 0 (the shoulder) to frame 5 (the wrist) (figure 3). This matrix is built by multiplying the successive homogeneous transformation matrices ${}^{i-1}T_i, i = 0, .., 5$.

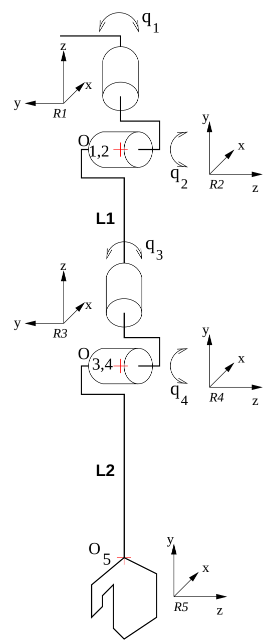

Figure 3: Arm model showing the frames used in direct geometric model calculation

This model is parametrized by θ_i, the joints angular values, and allows cartesian coordinates calculation. For angular coordinates calculation we need to inverse this model.

Inverse geometrical model : The inverse geometrical model is parametrized by the Cartesian coordinates of the wrist and returns the angular value θ_i for each joint. The first way to calculate this model would be to calculate the inverse of ${}^0T_5$, but in regard of the complexity of the calculation, splitting up the kinematic chain will be a far better solution. We calculate angular values for each joint separately by defining the inverse transformation ${}^3T_0$ that gives us the shoulder's joints angular values from elbows's Cartesian coordinates (expressed in frame R_0) and ${}^5T_3$ which gives us elbow's angular values from wrist's Cartesian coordinates expressed in R'_2 frame. R'_2 is a virtual frame oriented as R_2 and centered on the elbow.

Considering only the shoulder, we can write DH parameters (table 2) for the shoulder-elbow system, and define the homogeneous transformation matrix ${}^0T_3$ transformation matrix by multiplying the elementary transformation matrices (9)which specifies the transformation from frame

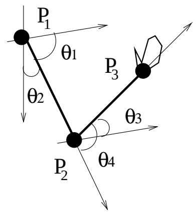

Figure 4: The inverse geometric model of the arm gives the joint angular values θ_1, θ_2 (shoulder joint) , θ_3, θ_4 (elbow joint) knowing shoulder, elbow and wrist Cartesian coordinates P_1, P_2, P_3

i	θ_i	d_i	α_i	a_i
1	θ_1	0	$-\pi/2$	0
2	θ_2	0	$\pi/2$	0
3	0	$-l_1$	0	0

Table 2: DH parameters for shoulder-elbow system

R_{i-1} to frame $R_i (i = 1, 2, 3)$:

$$^0T_3 = \begin{bmatrix} \cos\theta_{12} & -\sin\theta_1 & \cos\theta_{12} & -l_1\cos\theta_1\sin\theta_2 \\ \sin\theta_1\cos\theta_2 & \cos\theta_1 & \sin\theta_{12} & -l_1\sin\theta_{12} \\ -\sin\theta_2 & 0 & \cos\theta_2 & l_1\cos\theta_2 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Where $\cos\theta_{12}$ stands for $\cos\theta_1 * \cos\theta_2$.
The fourth column of $^0T_3$ represents the direct geometric model, so the inverse geometric model is :

$$\begin{cases} \theta_1 = \arctan(y_2/x_2) \\ \theta_2 = \arccos(z_2/l_1) \end{cases} \quad (10)$$

Doing the same calculation for the wrist brings (wrist's cartesian coordinates have to be expressed in elbow-centered frame R'_2) :

$$\begin{cases} \theta_3 = \arccos(z_3/l_2) \\ \theta_4 = \arctan(y_3/x_3) \end{cases} \quad (11)$$

As the arm model is redundant, direct inversion using the analytical solution will lead to unexpected reversal in angular values. To avoid it, we use a numerical resolution method to compute the first two joint values (θ_1, θ_2). This method can be initialized with previously computed values so that the new ones stay as close as possible to them which leads to smooth trajectories. Solution is computed iteratively using the pseudo-inverse of the arm system Jacobian (Klein C.A., 1983). Only the last two values are analytically calculated. This approach allows us to obtain a set of angular values corresponding to the given $3D$ joints position, even when the arm has a singular configuration.

3. Articulation-based motion representation

Articulation-based motion representation could be used to distinguish, among geometrical solutions, the good one, so that the first point to study is the variation of joint values for each solution. The second one concerns the possibility to use those representations to differentiate gestures based on the way they are made. Preliminary experiences have been made using a video corpora of elementary gestures.

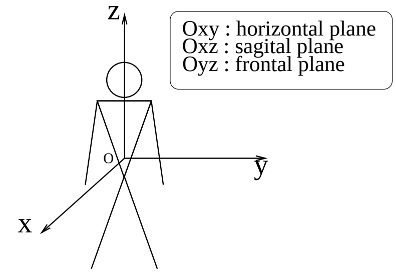

Figure 5: Representation of speaker-centered frame (ScF) showing the planes where most of the elementary gestures are realized.

The results (fig. 6) concern two circular motions of the left hand done in a plane parallel to Oyz plane of ScF, the first one with the arm in extension (gesture **A**) and the other one with the elbow bended (gesture **B**). The third gesture presented is a circular one made with the elbow bended in a plane parallel to Oxz plane of ScF (fig. 5) (gesture **C**). So that gestures **A** and **B** have quite similar aspect from the viewer's point of view and that gesture **B** and **C** are performed by moving articulations in a similar manner. Joints values are computed on each solution provided by the geometric reconstruction algorithm.

Figure (6) presents angular values evolution for each joint of the arm model and for three different gestures. Those values are presented in polar coordinates and ρ parameter stands for time (which means that gesture duration has been normalized). Different curves correspond to angular values computation for each geometrical solution. If we except noise on angular values implied by geometrical reconstruction, different angular trajectories for a same angle can be either confused (fig.6, θ_1 and θ_4 variations for gesture **A**) or symmetric (fig.6, θ_1 and θ_2 variations for gesture **B** and **C**). So that for each solution, changes in angular value variation occur at the same time.

One can remark too, that gesture **B** and **C** have closer signatures than gesture **A** and **B** in the sense that θ_1,θ_2 and θ_3 variations have the same kind of symmetry for those gestures : θ_1 and θ_2 are symmetric in regard of on a horizontal axis and θ_3 values present symmetries in regard of a vertical one. And that angular values for each articulation take values in the same part of the angular space.

4. Conclusion

Articulation-based motion representations are used to improve results computed by a single-camera geometrical algorithm which estimates possible poses of a human arm, being given a single image. This algorithm, provides us with a set of four possible motions for the arm in space. We made the assumption that using such a representation of gesture could allow us to use any of those solutions for gesture analysis. Primary experimentations on simple gestures brought out relationships such as symmetries or confusion between angular values for the different solutions, which is due to symmetries between the different solutions. On the other hand, recent linguistic issues made the assumption that using a proprioceptive representation of gesture is more suitable for Sign language analysis than a description based on elementary gestures described from an ob-

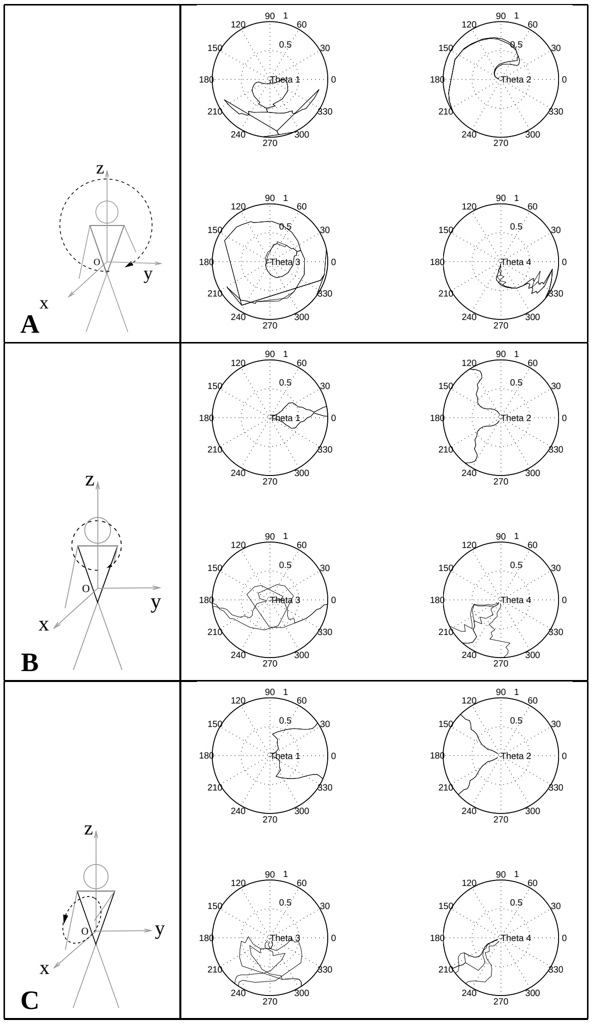

Figure 6: *On the left :* gesture **A,B,C** representation in ScF. Gesture **A** and **B** have similar visual aspect from the viewer's point of view, while gesture **B** and **C** are performed with similar articulation motion. *On the right :* joint values computed on each gesture and for each solution provided by geometrical reconstruction. Each solution is displayed as different curve. Each graph presents the evolution of the angular value for a given angle, from left to right, from top to bottom : $\theta_1, \theta_2, \theta_3, \theta_4$. Angle values are displayed in polar coordinates and ρ parameter stands for the time so that a constant angle value for an angle would be displayed as straight line starting at the center of polar frame.

servator point of view. Our algorithm make it possible to build such a articulation-based motion representation from single-camera data. Considering gestures performed in a similar manner with different orientations and comparing the results to gestures performed in a different manner but similar form observers point of view, we could observe that using our method will lead to a different gesture classification than the ones based on visual aspect in image or tridimentionnal representations. Further researchs have to be perform to bring out useful criterions to analyze real Sign language gestures from this point of view, but primary results are encouraging.

5. References

Athitsos V., Sclaroff S., 2004. Database indexing methods for 3d hand pose estimation. In *Gesture Workshop*, volume 2915 of *Lecture Notes in Computer Science*. Genova, Italy: Springer.

Boutet, D., 2001. Une approche morphogénétique du sens dans la gestuelle conversationnelle. In C. Calvé et al. (ed.), *Oralité et Gestualité*. Paris: l'Harmattan.

Braffort, A., 1996. Argo : An architecture for sign language recognition and interpretation. In P. Harling and al. (eds.), *Progress in Gestual Interaction*. Springer.

Denavit J., Hartenberg R.S., 1955. A kinematic notation for lower-pair mechanisms based on matrices. *Journal of Applied Mechanics*:215–221.

Gurdjos P., Payrissat R., 2002. Calibrage plan d'une caméra en mouvement à focale variable. In *RFIA 2002*. AFRIF-AFIA.

Heikkilä, J., 2000. Geometric camera calibration using circular control points. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(10):1066–1077.

Horain P., Bomb M., 2002. 3d model based gesture acquisition using a single camera. In *Proceedings of the IEEE Workshop on Applications of Computer Vision (WACV 2002)*. Orlando.

Klein C.A., Huang C.H., 1983. Review of pseudoinverse control for use with kinematically redundant manipulators. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-13(3):245–250.

Lenseigne B., Gianni F., Dalle P., 2004. Estimation monovue de la posture du bras, méthode et évaluation. In *RFIA 2004*. Toulouse: AFRIF-AFIA.

Prillwitz, S. and al., 1989. *HamNoSys. Version 2.0; Hamburg Notation System for Sign Languages. An introductory guide*. Hamburg : Signum.

Somers M.G., Whyte R.N., 2003. Hand posture matching for irish sign language interpretation. In *ACM Proceedings of the 1st international symposium on Information and communication technologies*. Dublin Ireland: Trinity College Dublin.

Starner T., Pentland A., Weaver J., 1998. Real-time american sign language recognition using desk and wearable computer based video. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(12):1371–1375.

Tanibata N., Shirai Y., Shimada N., 2002. Extraction of hand features for recognition of sign language words. In *15th International Conference on Vision Interface*. Calgary, Canada.

Vogler C., Metaxas D., 1998. Asl recognition based on a coupling between hmms and 3d motion analysis. In *Proceedings of the International Conference on Computer Vision*. Mumbai, India.

Vogler C., Metaxas D., 1999. Parallel hidden models for american sign language recognition. In *Proceedings of the International Conference on Computer Vision*.

Wren C., Pentland A., 1999. Understanding purposeful human motion. In *Proc. IEEE International Workshop on Modelling People (MPEOPLE)*.