SIGN LANGUAGE FORMAL DESCRIPTION AND SYNTHESIS

O Losson, J-M Vannobel
University of Science and Technology of Lille, France

ABSTRACT

Special needs of deaf people appeal henceforward to sign language synthesis. The system presented here is based on a hierarchical description of sign, trying to take the different grammatical processes into account. Stress is laid on hand configurations specification thanks to finger shapes primitives and hand global properties, and on location and orientation computation issues. We then expose the results achieved from the corresponding written form of signs, leading to their computer virtual animation.

1. Introduction

Sign language appears to be the main means of communication within the deaf community. In spite of a repression cycle it has suffered from in several countries until recent past, it remains a living language, in the full acceptance of the term, with communication capacities similar to those of oral languages.

Our project of synthesizing sign language is based on the established facts that half of deaf people encounter difficulties in reading and, as a result, suffer from subeducation. Being able to synthesize their mother tongue would undoubtedly be of major interest for those persons. In education as well as in daily life – such as in an emergency – the generation process of synthetic signs offers many advantages in comparison with video: facility and swiftness of generation and transmission, availability of the virtual signer and smoothness of transitions between lexical items.

The first three-dimensional synthesis of sign has been achieved in the early eighties (Shantz and Poizner, 1982). But it suffered from the lack of power of computers at the time, which entailed low-level (joint angles) specification of body postures. Surprisingly enough, very few attempts have been made since then to carry out the same goal.

This study is in keeping with the global aim of translating written French into signs. But the starting point considered lies in a textual representation of the signed sentence which is to be animated; as a consequence, we will not deal with sign language syntax here. The main core of the system is a sign formal description, grounded on a set of primitives isolated beforehand, stemming from both linguistic works and more synthesis-oriented research.

2. Formal Description

Figure 1: General hierarchic structure of sign specification.
2.1. GRAMMATICAL BASES AND MODULATION PROCESSES

Following Stokoe’s footsteps, linguists have proved that sign languages, like oral ones, were doubly articulated into phonemic and morphemic (or monemic) levels. This decomposition has definitely given evidence for deaf people's signs system to be counted among all languages in the world. Second-level articulation units have been identified on the minimal pairs criterion (Nève, 1997), i.e. two signs differ from one another in only one of these cheremes. Belonging to one of four spatial types - hand configuration, manual location and orientation, and movement -, they combine to form morphemes, lexemes (signs) and whole statements; their psycholinguistic actuality has been besides fully validated.

Sign language has its own grammar, based on spatial, physiological and temporal dimensions. For instance, strongly iconic classifiers and size-and-shape specifiers, but also pronouns and index references, appeal to mechanisms involving the signer’s space. Non-manual expressions both play paralinguistic (kinetic stress, face expression) and grammatical roles (marking syntactic clauses). At last, it has been shown that special modulation processes (Klima and Bellugi, 1979; Namir and Schlesinger, 1978) affecting repetition, shape, amplitude and kinematics of movement (Loomis et al, 1983), were used to express subtle variations in meaning.

Liddell and Johnson (1989) have proposed a highly detailed description of sign based on the partition between segmental (holds and movements) tier and articulatory bundle tier (containing features of the hand). One of the major interest of this approach is to tackle with particular phonological (hold deletion, assimilation) but also morphological processes of sign language, including the one – of fundamental importance – of subject and object agreement.

2.2. A SIGN LANGUAGE FORMAL DESCRIPTION

The proposed sign language formal description system tries to take the widest range of such processes into account. To achieve that purpose, it has been split into two levels. Sentence, on the one hand, has its own parameters (localization and indexic references, grammatical clause descriptor, …). On the other hand, signs description may inherit from some of those discourse parameters. We will focus here on sign specification.

In order to provide the most general possible synthesis, with abilities to describe not only French- but every Sign Language, one of the basic underlying principles was to identify primitives at all levels, aimed at gradual combination into more and more complex structures (see figure 1). In that way, a hand configuration is described in terms of digital primitives and global hand properties (see details in section 3). Together with orientation, location and an optional manual point, it makes up what we call hand specification. Specifying a contact point is convenient in those many cases where the latter, rather than the wrist, must be located at the given position. Movement, as for it, is composed of one main move, with specific path and dynamics. Optional preceding and following holds, as well as some superimposed secondary movement (waving of fingers for instance), may complete movement specification.

A shift primitive is built as a bundle of hand specifications at the beginning and end of the sign, and of the movement itself. A transitional hand specification may be added too if necessary. Depending on the number and activity of articulators implied in a sign, a shift can be described - either as a single shift primitive (if the sole strong hand is present), - or a (dominant hand) shift primitive plus a hand specification for the weak hand, - or two shift primitives if both hands are active. The description is completed with a spatial relationship and a descriptor for synchronization between the two movements in this case (see table 1).

We introduce the notion of macro-shift in order to be able to describe quite complex signs (such as compounds). This is made up of one or more shifts, optionally followed

| Sign | Meaning | Description |
|------|---------|-------------|
| French Sign Language (FSL) : | **Waterfall** | Hand specification for the static weak hand : Hand configuration [Flat], close to [torso], facing [down] |
| Shift primitive for the strong hand : | Initial hand specification : [5] at [weak hand back] facing [down] | Final hand specification : [5] at [lower torso] facing [back] |
| [Arc-shaped] movement while [wiggling] of fingers | |
| American Sign Language (ASL) and FSL : | **Judge** | Shift primitive for the strong hand (implicit for the weak hand) : Hand spec. : [clip] at [ipsilateral upper torso] (initial) or [ipsilateral lower torso] (final) facing [contralateral] |
| [Straight] movement | Spatial relationship : [symmetrical] hands, [alternate] move repeated shift according to [simple repetition] |

Table 1. Two examples of signs with main features
by a repetition. Isolation of the different kinds of repetitions, temporal and spatial relations, has been made by a systematic census on a French signs corpus (Moody, 1986).

At last, we have endowed our sign description with grammatical and non-manual information. Facial expression is of particularly salient importance when considering grammatical parameters inherited from the sentence level.

3. Hand Configurations

Hand configurations in sign languages have given rise to several coding systems (Kurokawa, 1992; Lee, 1994). Among them, HamNoSys (Prillwitz and Zienert, 1990) seems to be able to encode the widest diversity of hand shapes. We have set up a new descriptive method, based on the examination of different inventories, and split into three levels: finger shape primitives, hand global properties and constraints on joints.

3.1. FINGER CONFIGURATION PRIMITIVES

All fingers except thumb have the same behaviour in terms of movement: they may flex at their interphalangeal joints, and both flexion and abduction (spreading of fingers) are permitted at the metacarpophalangeal (MCP) joint. Thumb has a larger reachable workspace due to greater articulatory complexity. Especially, it may contact other fingers in various ways. In those cases, its shape will be considered here as constrained in a hand global property (primitives specifying only free thumb configurations).

Figure 2: Example of #Flat finger configuration primitive.

We have isolated seven finger shape primitives, and five ones for the thumb, as being used in sign languages. For instance, the configuration called ‘#Flat’ (see figure 2) has flexed MCP joint and extended proximal and distal interphalangeal joints. The very peculiar ‘#E’ configuration has been taken into consideration at the present level, although fingers interaction is involved. As a matter of fact, it almost merely appears in the corresponding manual alphabet entry.

3.2. HAND PROPERTIES AND CONSTRAINTS ON JOINTS

Hand properties imply more than one digit. Thumb and finger contact (or opposition), and abduction or crossing of fingers are here taken into account. Many kinds of such relationships were found in sign languages:

- Contact, or opposition without contact, between thumb and another finger:
  - Ventral contact: thumb tip is in touch with the middle phalanx of the finger in question, that may have any configuration (#Hook in most cases).
  - Flat contact between pads of thumb and of another finger, which should have #Flat configuration.
  - Contact between tips of thumb and of another finger, which should have #O configuration.
- Crossing of index and middle fingers. Other kinds of crossing are far more constraining for joints and absent from sign language.
- Intercalation of thumb, mainly between index and middle fingers. Thumb may also functionally come between middle and ring fingers, or between ring and middle fingers; nevertheless, those situations are seldom used in sign language.
- Dorsal contact: thumb pad covers a finger on the back of its middle phalanx.
- Abduction / adduction is also considered here, insofar as it generally involves several fingers, if not the whole hand.

Figure 2 above shows an example of a flat contact between the thumb and index finger.

Final step for hand configurations specification is the application of constraints on joints. The metacarpophalangeal structure of ligaments implies indeed more or less strong interdependence between flexions of neighbouring fingers at that joint. Inequalities proposed in (Lee and Kunii, 1993) have been applied on MCP joints as flexion limits, in order to achieve realistic hand shapes synthesis. Moreover, in a clenched hand, fingers converge towards the scaphoid point. That phenomenon has been included in the synthesis process by setting artificial abduction, proportional to flexion, when the latter exceeds two thirds of its static maximal value.

3.3. HAND CONFIGURATION SYNTHESIS

The kinematic skeleton of the thumb, as for the whole body, is mathematically modeled by ideal joints and flat segments, each one being defined in the local coordinate system attached to the proximal joint. Axes are chosen such that x is the main axis of each segment, oriented from the proximal to the distal joint, and that (sz) is the plane in which segment points are given (palm for instance in figure 3), y being defined by the right-hand rule. Rotations about those axes will be considered in the following order: abduction-adduction (yaw) $\theta$ about the y-axis, then flexion-extension (pitch) $\phi$ about the x-axis, and finally axial rotation (roll) $\psi$ about the z-axis.
The major issue in hand configurations synthesis is thumb posture computation. As called to mind above, this articulator has more degrees of freedom than other fingers. Several methods have been described to position such articulated structures, even when multiple goals are to be achieved (Badler et al., 1987). We have rather rationally tried to simplify our initial model so as to obtain a non-redundant system and solve it. Axial rotation takes place at the carpometacarpal (CMC) joint, especially during opposition movement. This one will here be undertaken by a fixed initial rotation, about the axis of the first metacarpal (see figure 3). Neglecting the low adduction-abduction angle, only flexion will be considered at the MCP joint, which is moreover assumed not to vary in a wide range of values around ten degrees. Thumb tip may therefore be expressed in the CMC local coordinate system as

\[
x_T = \begin{pmatrix}
c_\theta_1 c_\phi_1 & -c_\theta_1 s_\phi_1 & s_\phi_1 & s_\theta_1 c_\phi_1 & p_1 c_\phi_1 \\
c_\phi_1 & c_\phi_1 & -s_\phi_1 & -s_\theta_1 c_\phi_1 & p_1 s_\phi_1 \\
-s_\theta_1 c_\phi_1 & s_\theta_1 s_\phi_1 & -c_\phi_1 & -s_\theta_1 s_\phi_1 & -s_\theta_1 c_\phi_1 & -p_3 s_\phi_1 \\
0 & 0 & 0 & 0 & 1 \\
\end{pmatrix}
\]

where \( c \) means cosine, \( s \) means sine, and \( p_1, p_2, p_3 \) are segment lengths, from the proximal to the distal one.

Solving of those inverse kinematics equations is achieved thanks to a method presented in (Kobrinski and Kobrinski, 1989), assuming that the distance from \( x_T \) to the target position \( x \) expresses as a function of only one of the unknowns, others remaining fixed. Let \( \alpha_i \) be the \( i \)-th rotation angle and let us determine its optimal value \( \alpha_i^* \). Rewriting \( x_T \) as

\[
x_T = a_{\alpha} \sin \alpha_i + b_{\alpha} \cos \alpha_i + d_{\alpha},
\]

where vectors \( a_{\alpha}, b_{\alpha} \) and \( d_{\alpha} \) do not depend upon \( \alpha \), the minimum of distance \( |x - x_T|^2 \) is reached when

\[
tg \alpha_i^* = \frac{i a_{\alpha}(d_{\alpha} - x)}{i b_{\alpha}(d_{\alpha} - x)}
\]

From the two possible values of \( \alpha_i^* \), we must keep the one leading to a positive value for the second-order derivative, or for the quantity \( i (a_{\alpha} \sin \alpha_i + b_{\alpha} \cos \alpha_i)(d_{\alpha} - x) \). This method provides realistic shapes for the thumb, as shown in figure 4.

4. Location and Orientation

4.1 LOCATION

Ranges of locations in sign language are limited by the reachable workspace of the hand (Lenarcic and Umek, 1994): almost always above the waist, ahead of the front plane, roughly within half spheres placed in front of the torso, in front and both sides of the head. In our system, hand may be assigned a location in two ways:

- in space, with main points on a grid defined as intersections of three planes (horizontal, parallel to the frontal plane, parallel to the sagittal plane);
- on the signer’s body (including the weak hand), at various locations inspired by linguistic studies (Liddel and Johnson, 1989).

Points on the body can be expressed in terms of spatial coordinates in a 3D synthesis. Therefore, what we have to do here is, given a target location for the wrist, compute the four angles of the arm (three at the shoulder - \( \theta_s, \phi_s, \psi_s \) – and one at the elbow - \( \phi_e \)).

| Hand Prop. Parametor(s) | #Pad | #Ventral | #Covered | #Covered | #Tip | #ThumbBetween |
|-------------------------|------|----------|----------|----------|------|---------------|
| Config. name            | [Beak] | [Key]    | [B] [H]  | [R] [K]  | [R]  | [T]           |

Figure 4: Wireframe 3D synthesis of hand configurations using different hand properties.
Redundancy of that positioning problem has been solved as explained below.

- First, one natural configuration of the arm is selected for each main point in space, among a collection of possible quadruplets obtained by successive iterations on $\theta_i$ in the following equations:

$$s\phi_i = \frac{BC \pm A\sqrt{A^2 + B^2 - C^2}}{A^2 + B^2}$$

with

$$A = 2l_1xc\theta_s - 2l_1zs\theta_s$$
$$B = 2l_1y$$
$$C = l_2^2 - l_2^2 + x^2 + y^2 + z^2$$

$$l_zc\phi_s = xc\theta_s c\phi_s + ysc\phi_s - zsc\phi_s - l_1$$
$$l_zs\psi s\phi_e = xs\theta_s + zc\theta_s$$

- In the second time, we can get the four arm angles for any point in space from those configurations of the neighbouring points $N$. The starting point is computing the average angles weighted by the inverse distance to each of them:

$$\bar{q} = \left\{ \bar{\theta}_s, \bar{\phi}_s, \bar{\psi}_s, \bar{\psi}_e \right\}$$

with

$$\bar{q} = \sum_{i \in N} \frac{1}{d(P,i)}\alpha, \quad \alpha \in \left\{ \theta_s, \phi_s, \psi_s, \psi_e \right\}$$

We then keep the solution, among the set $Poss$ of possible ones, which is the nearest from that mean value:

$$q^* = \min_{q, q \in Poss} \left| q - \bar{q} \right|$$

The above strategy keeps us out of generating unnatural configurations for the arm, while avoiding greedy algorithms of inverse kinematics.

4.2 ORIENTATION

According to linguistic studies, palm orientation seems to be sufficient to encode hand orientation. Six main directions only (up, down, front, back, ipsilateral and contralateral) are generally isolated as a result of phonological substitution. We cannot release from specifying hand orientation more precisely in our synthesis perspective. This is done by defining both a forward vector ($\vec{n}$) normal to the palm, and a vector ($\vec{i}$) oriented according to the (virtually) pointing index finger, or to the axis of the third metacarpal, from CMC3 to MCP3.

The problem of hand global orientation consists in computing:

- the radius-ulnar pronation-supination angle $\psi_e$.
- two angles at the wrist joint: adduction-abduction $\theta_w$ and flexion-extension $\phi_w$.

In mathematical terms, let $r_0 A_{k\bar{e}} = (s \ n \ a)$ be the current orientation of the local coordinate system at the elbow, after shoulder rotations and elbow flexion have been applied. The required global orientation of the hand thus expresses as:

$$r_0 A_{k\bar{e}} = (i \ n \ i \wedge n)$$
$$= (\sigma \ n \ a)$$
$$= r_0 A_{k\bar{e}} A(x_2, \psi_e)A(y_3, \theta_w)A(z_3, \phi_w)$$

Identification leads to the desired angles:

$$\psi_e = -\arctan \left( \frac{\dot{n}a}{\dot{n}c} \right)$$
$$\theta_w = \arctan \left( \frac{\dot{s}c \dot{c} \psi}{\dot{a}a} \right)$$
$$\phi_w = -\arctan \left( \frac{\dot{s}v}{\dot{s}c} \right)$$

It often appears easier and convenient to specify a particular point $M$ on the hand that is to be settled at the given place, rather than setting location of the wrist $W$. In sign I-give-you for example (see section 5), the index finger tip should be located on the upper torso at the beginning of the movement. Since coordinates of $M$ are known in the
In the last place, specifying relative orientation has been contemplated by using a set of predefined values for each angle $\psi_e$, $\theta_w$, and $\phi_w$. In a number of cases besides, only pronation-supination of the forearm is required to describe the orientation of the hand. For the moment, this method is solely used when the wrist joint – no hand point – is concerned by the target location.

5. Computer Synthesis of Sign

5.1 SYNOPSIS

To specify any sign according to the structure presented in section 2, we have preferred a textual extensive description of the sign, rather than a symbol-based coding system (on grounds of legibility and easiness of data exchange). The sign compiler implements a lexical analyser, a parser, and the evaluator itself. It builds the hierarchic sign specification structure in memory from predefined entities like hand configurations, while inheriting from grammatical information from the sentence level through sign parameters (see figure 6).

The graphical synthesis module generates the signs after each one has been previously evaluated in this way. The underlying model of the human body is a hierarchic tree of upper body segments and joints, from the torso to the eyebrows and distal phalanges. Each segment is described by a collection of 2D-points and has one or more children segments together with the proximal joint attached (including the local and global coordinate systems). Moreover, segments are regrouped within macro-segments (fingers, hands, head, ...) accepting high-level messages.

Input sentences look as streams of sign-words and escape codes handling grammatical features such as indexic references and role play, time setting, clause type (condition, wh-question, yes/no question, imperative), grammatical repetition, etc. Signs are described in a formal language with special syntax close to the Smalltalk object-oriented programming language.

5.2 RESULTS

In the current progress of our work, the system computes arm postures as well as hand configuration and orientation from the given sign textual description. Symbols used to specify hand shapes, contact points and spatial locations and orientations, are evaluated so as to transmit suitable objects to the human body structure. The graphical synthesis has been achieved by means of connected deformable polygons, with distance-dependent lighting and removal of hidden surfaces. Primitives of facial expressions have been added too, as part of lexical items, but conveying also crucial grammatical information.

Two parameters have been inserted in the sign above to take the subject and object agreement into account: thanks to local variables, we are able to determine the sign features with the different possible values (I/me, you and he/him) for the agent and patient parameters. The syntax used leads to a very readable description of the sign, in which the formal hierarchic sign specification clearly appears. It is built and referenced to by sending messages to objects. A graphical interface is provided with the editor in order to make sign edition easier. So far, it allows the user to add hand specifications with direct tridimensional visualization of the selected elements.

For the moment, movement remains rather simple: ‘natural’ shifts (through joint angles extrapolation from one hand specification to another) and the main types (straight...
specified by the user. More complex movements may be then considered as series of goals (Lebourque and Gibet, 1997), and synchronization between the two articulators be handled by event-driven processes with Grafcet or Petri-net representations. Beside, we have recently developed a specific application intended for the analysis of sign duration and dynamics from video sequences. This should enable us to specify tense of signs better, just as pauses between signs at the sentence level.

6. Conclusion

One of the guidelines of the system described here was to be able to generate any kind of sign, as far as possible. The proposed formal description is generic enough to take up such a challenge. Not only does it provide primitives and more complex structures found in every sign, but it also tries to take the widest variety of sign language grammatical processes into consideration. Such an open system lets the user entirely free to adjust the encoding sharpness of signs.

Movement must still be specified precisely in terms of path dimension and dynamics, together with hands arrangement and synchronization. But the virtual signer already shows interesting results for simple signs, especially a high-level specification of hand configuration, location and orientation. In collaboration with native deaf signs, we now intend to test the generated synthetic signs in order to optimize their expressive potential.

REFERENCES

N I Badler, K H Manoochehri and G Walters (1987), Articulated figure positioning by multiple constraints, IEEE Computer Graphics and Application, 7, 6, pp. 28-38.

E Klima and U Bellugi (1979), The signs of language, Harvard University Press, Cambridge, London.

A A Kobrinski and A E Kobrinski (1989), Bras manipulateurs des robots - Architecture et théorie, Mir editions, Moscow [in French].

T Kurokawa (1992), Gesture coding and a gesture dictionary for a nonverbal interface, IEICE Trans. Fundamentals, E75-A, 2, pp. 112-121.

T Lebourque and S Gibet (1997), Synthesis of hand-arm gestures, In Progress in gestural interaction - Proc. Gesture Workshop '96 (P A Harling & A D N Edwards, Eds), Springer-Verlag, Berlin, pp. 217-225.

J Lee, T L Kunii (1993), Constraint-based hand animation, In Models and Techniques in Computer Animation World (N Magnenat-Thalmann & D Thalmann, Eds.), Springer-Verlag, Berlin, pp. 110-127.

J Lee (1994), Notational representation of sign language: a structural description of hand configurations, In Computers for handicapped persons – Proc. 4th Int. Conf., ICCHP '94 Vienna, Austria (W L Zagler, G Busby, B R Wagner, Eds.), Springer-Verlag, Berlin, pp. 38-45.

J Lenarcic and A Ume (1994), Simple model of human arm reachable workspace, IEEE Trans. on Systems, Man and Cybernetics, 24, 8, pp. 1239-1246.

S K Liddell and R E Johnson (1989), American Sign Language: the phonological base, Sign Language Studies, 64, pp. 195-277.

J Loomis, H Poizner and U Bellugi, Computer graphics modeling of American Sign Language, Computer Graphics, 17, 3, pp. 105-114.

B Moody (1986), La langue des signes – Tome 2 : dictionnaire bilingue élémentaire, International Visual Theatre (I.V.T.), Editions Ellipses, Paris [in French].

L Namir and I M Schlesinger (1978), The grammar of Sign Language, In Sign Language of the Deaf : Psychological, Linguistic and Sociological Perspectives (I M Schlesinger & L NAMIR, Eds), Academic Press, New York, pp. 97-140.

F-X Nève (1997), Essai de grammaire de la Langue des Signes Française, Bibliothèque de la Faculté de Philosophie et Lettres de l'Université de Liège, Liège (Belgium) [in French].

S Prillwitz and H Zienert (1990), Hamburg Notation System for Sign Language: Development of a sign writing with computer application, In Current Trends in European Sign Language Research - Proc. 3rd European Congress on Sign Language Research (S Prillwitz & T Vollhaber, Eds), Signum, Hamburg, pp. 355-380.

M Shantz and H Poizner (1982), A computer program to synthesize American Sign Language, Behavior Research Methods and Instrumentation, 14, 5, pp. 467-474.
BIOGRAPHIES

Olivier Losson (photo on CDROM) is a PhD student at the I3D (Interaction, Image & Decision Engineering) Laboratory in Lille, France. He received his engineering degree from the Ecole Centrale de Lille in 1994, while orienting himself towards research through Petri-nets and event-driven processes. He then started a PhD thesis in the team of Pr. Toulotte, specialized in enriched palliative communication. His main study concerns sign language of the deaf, but his interests include every research applied to handicap. He also works for the Regional Institute for Research about Handicap.

Contact information:
Oliver Losson
Laboratoire I3D (CEP) - IRRH
Université des Sciences et Technologies de Lille
Bâtiment P2, Cité Scientifique
59655 Villeneuve d'Ascq
FRANCE
Email: lo@cal.univ-lille1.fr

Jean-Marc Vannobel (photo on CDROM) received his Ph.D. in 1995 in industrial automation from the Science and Technology University of Lille. He is currently an associate professor in the I3D laboratory at the same university. His research interests focus on gesture communication, especially for deaf people at present, including both recognition and synthesis of French Sign Language and lipreading based on hands and lips video tracking.

Contact information:
Jean-Marc Vannobel
Laboratoire I3D (CEP) - IRRH
Université des Sciences et Technologies de Lille
Bâtiment P2, Cité Scientifique
59655 Villeneuve d'Ascq
FRANCE
Email: jmv@cal.univ-lille1.fr