Constructing new coordinate system suitable for sign animation

T Kuroda[1], K Okamoto[2], T Takemura[3], K Nagase[4] and H Yoshihara[5]

[1,3,4,5]Department of Medical Informatics, Kyoto University Hospital,
54, Kawahara-cho, Shogo-in, Sakyo-ku, Kyoto, JAPAN

[1,2,3,4,5]Graduate School of Informatics, Kyoto University,
36-1, Yoshida-honmachi, Sakyo-ku, Kyoto, JAPAN

[1,2,3,4,5]*{Tomohiro.Kuroda, kazuya, takemura, nagase, lob}@kuhp.kyoto-u.ac.jp*

www.kuhp.kyoto-u.ac.jp/~mi/

ABSTRACT

This paper proposes new coordinate system suitable for denoting sign language motion. As the proposed coordinate system consists of polar coordinate systems whose origins are certain points of human body, postures shown on the system can be proportional for avatars with any possible shape and fit with existing subjective sign notation systems. This paper extracted coordinate origins from Japanese-Japanese Sign Language Dictionary via morphological analysis. Selected 85 points are successfully mapped on H-ANIM standard humanoid avatar.

1. INTRODUCTION

Sign linguistic engineering is a group of research to develop communication aid (called sign information system) for the Deaf and the Hearing Impaired, who have communication barrier in social lives, using information technology (Nagashima and Kanda 2001). Figure 1 shows the general view of sign information systems. Any sign information systems, such as virtual reality based sign language telephones, automatic translation system between sign languages and phonetic languages, and even sign language dictionary, share same basic structure. Most of sign information systems use three-dimensional computer graphic model of human beings called avatar to show signs. To obtain signs, though several systems using image recognition technologies, most of them are utilizing motion capture systems consists of data-gloves and position and orientation sensors either magnetic-field-based, ultrasound-based, and image-based. The authors also developing sign language telephone named S-TEL (Kuroda et al 1998) and consequent systems based on motion capture system consists of two data-gloves and three position and orientation sensors.

Most of sign animations handles motion as sets of rotation angles of each joint and generate animation applying the data to an avatar model, just as same as MPEG-4 H-ANIM standard (ISO/IEC, 2001). However, a certain body motion parameter to show a certain sign on a certain avatar cannot produce the same sign when it applied on another avatar with different body shape. Thus, the sign linguistic system needs to prepare an avatar, which holds the same body shape as a certain person to produce proper signs. Although authors developed a system to morph a standard avatar to fit a certain person through image recognition (Kuroda et al, 2001), it cannot give users full selection of avatar's outlook to have favorable view. Thus, to give flexibility and usability on sign information systems, a new motion coding system, which is independent from body shape and easily transferable from or to standard motion data such as a captured motion, data via a certain motion capture system or H-ANIM.

This paper proposes new coordinate system suitable for sign animation. Notations of textbooks or dictionaries of sign language present a position and orientation of hands in a certain posture in relation to a certain point of human body.

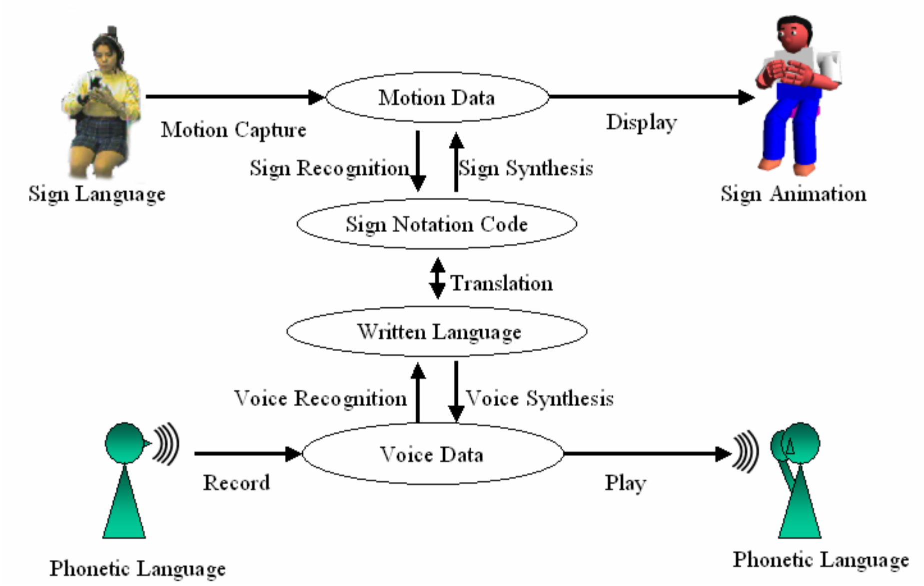

Figure 1. *General overview of sign information systems.*

2. FOREGOING CODING SYSTEMS

Foregoing researches on sign language developed several types of sign coding systems. Most of the sign notation system follows framework of Stokoe (1965), which claims signs are consists of "Cheremes", that is, hand posture (Dez), position (Tab) and motion (Sig). Thus, most of the notation system denotes signs as a series of combination of hand posture, orientation and position, although some of them denote motions as shape of trajectories. Sometimes several non-manual signals, such as facial expressions, are added on them.

HamNoSys (Hanke, 2004), one of the most well-known sign notation system, also denote signs as a combination of hand posture, orientation and position. HamNoSys denote position of hands as combination of upper/middle/bottom of the shoulder and left and right of body, although it adds several more codes on update of version 4.0, the resolution is still quite rough.

SignWriting (Satton, 1998), another one of the most well-known sign notation system, denotes hand posture and position, orientation and speed. As SignWriting is designed for easy denoting and reading of sign language and is derived from dance notation system, the notation is quite pictorial and subjective. Thus, automatic conversion from or to motion capture data seems quite difficult.

Kurokawa (1992) developed a motion coding system for nonverbal interface and applied the coding system to denote sign language. However, resolution of the coding system, again, quite rough to synthesis signs animation from given code as the system designed mainly for motion recognition.

As discussed above, foregoing sign notation systems are based on subjective coding and its resolution are not detailed enough, new coding system, which mediates such subjective coding system and motion data, such as FACS (Ekman and Freisen, 1978) for facial expression, is essential for sign language information systems. As hand posture can be mapped via simple vector quantization as authors presented on previous conference and implemented on StrinGlove® (Kuroda, 2004), the coding system to denote hand position and orientation can be sufficient.

3. DESIGNING NEW COORDINATE SYSTEM

3.1 Basic approach

As discussed above, the target of this research is to develop hand posture/orientation coding system to mediate raw captured or kinematics data and existing subjective sign notation codes.

Careful observation of foregoing sign notation systems, especially new location symbols added on HamNoSys in version 4.0 (Schmaling and Hanke, 2004), gives interesting insight of sign posture coordination. Figure 2 shows several new added location symbols. As shown in the figure, new version of

Proc. 6th Intl Conf. Disability, Virtual Reality & Assoc. Tech., Esbjerg, Denmark, 2006
©2006 ICDVRAT/University of Reading, UK; ISBN 07 049 98 65 3

HamNoSys gives locations of nominal hand by several body parts such as lip, ear lobe, nose and arms. On the other hand, quick review over several textbooks of sign languages also gives positions of hands or fingers by referring certain body parts. Therefore, human being seems to position hands relative to a certain body parts, such as lip, nose, chest, elbow, etc. Thus, coding system that let us denote relation between hands and other body parts can fit subjective insight of human beings as well as existing sign notation code.

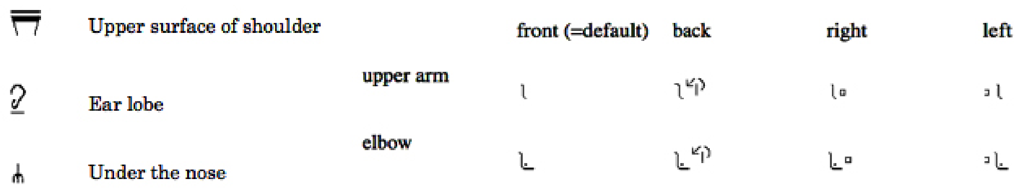

Figure 2. *New location symbols on HamNoSys 4.0 and its usage.*

On the other hand, to enable automatic conversion from or to “raw” motion data, such as captured motion data or generated kinematics data, the coding system must not be sparse or rough. The coding system should have enough resolution to show its precise position in a certain coordinate.

The authors propose to develop new coordinate system suitable for sign motion rather than coding system. The coordinate system consists of sets of polar coordinate system whose origins are certain points (the reference points) on human body, such as “inner corner of right eyelid” or “eye”. As the reference points are proportional to shape of body, the proposed system can proportion position and orientation of hands to various avatar models including cartoon-like model with big head as shown in Fig. 3.

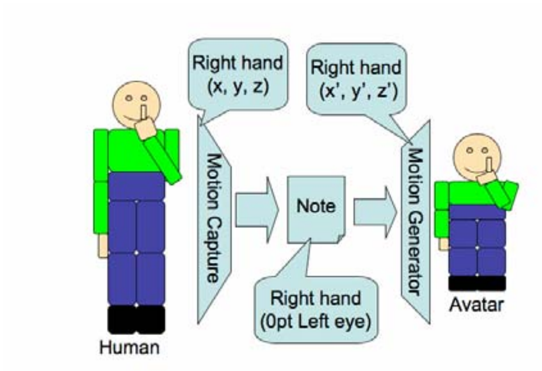

Figure 3. *Basic concept of the proposed coordinate system.*

This coordinate system allows automatic conversion from “raw” motion data by calculating out distance of hand from several reference points and to “raw” motion data by mapping them on certain avatar. Additionally, a certain motion denoted on subjective notation code can be mapped onto the coordinate by giving distance on a certain rule.

3.2 Picking up the reference points

To obtain reference points, this paper analyzes notations of Japanese – Japanese sign language dictionary issued by the Deaf association of Japan (Yonekawa, 1997). The dictionary has 8320 headwords (Japanese) and one headword has a pair of gestures (Japanese signs) and notation of each sign gestures in Japanese as shown in Fig. 4. In total, the dictionary has 16640 sign notations.

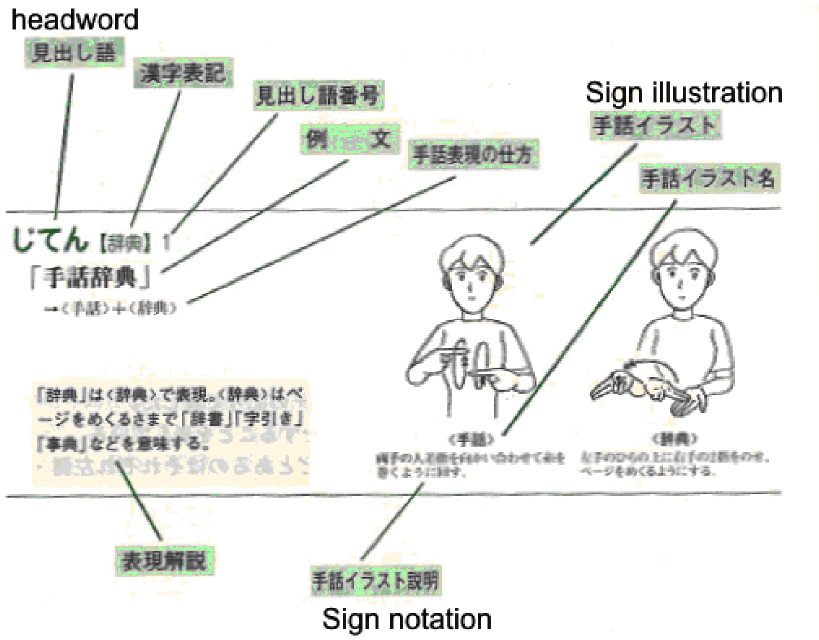

Figure 4. *Explanatory notes of Japanese-Japanese Sign Dictionary [Yonekawa, 1997]*

The authors scanned all pages of the dictionary and obtained notations via optical character reader. The obtained notations are analyzed through ChaSen, a morphological parser of the Japanese Language developed by NAIST Computer Linguistic Laboratory (Matsumoto et al, 2001). In morphological analysis, the authors applied medical corpus developed for natural language processing in medicine by the authors (Takemura and Ashida, 2001) to obtain proper anatomical nouns. The dictionary had 3249 noun or compound noun candidates including compound noun connected by the possessive particle “NO”. To extract possible reference point, the authors selected 691 candidates including part of body manually. Analysis of the 691 candidates shows that the extracted compound nouns are consists of single prepositional modifier such as “HIDARI (left)” or “MIGI (right)”, single or multiple nouns of body parts concatenated by possessive particle, and single postpositional modifier following a possessive particle such as “SHITA (under)” or “UCHIGAWA (inside of)” as shown in Fig 5, although possessive particles sometimes drop. Along this analysis, the authors divided selected compound noun candidates by the possessive particle into 142 compound nouns including part of body and 46 postpositional modifiers, and divide the 142 compound nouns into 4 prepositional modifiers and 85 nouns indicates part of body.

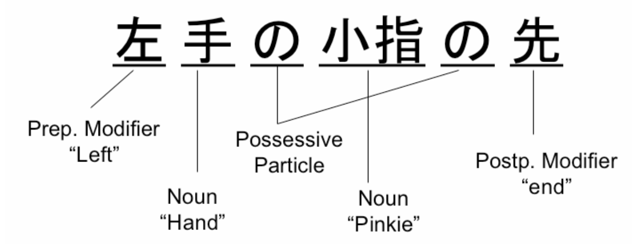

Figure 5. *Typical expression of reference points in sign notation.*

3.3 Applying the obtained reference points

The extracted 85 nouns obtained through semi-automatic morphological analysis may express same body part due to nature of Japanese Language, such as hiragana and Chinese character explanation for exactly same words. Manual investigation founds the nouns consists of 43 different body parts. The extracted 43 parts are successfully mapped on either a certain point of kinematical model or reference model of MPEG4 H-ANIM standard as Fig 6. Therefore, the proposed coordination system can be proportional to various avatars.

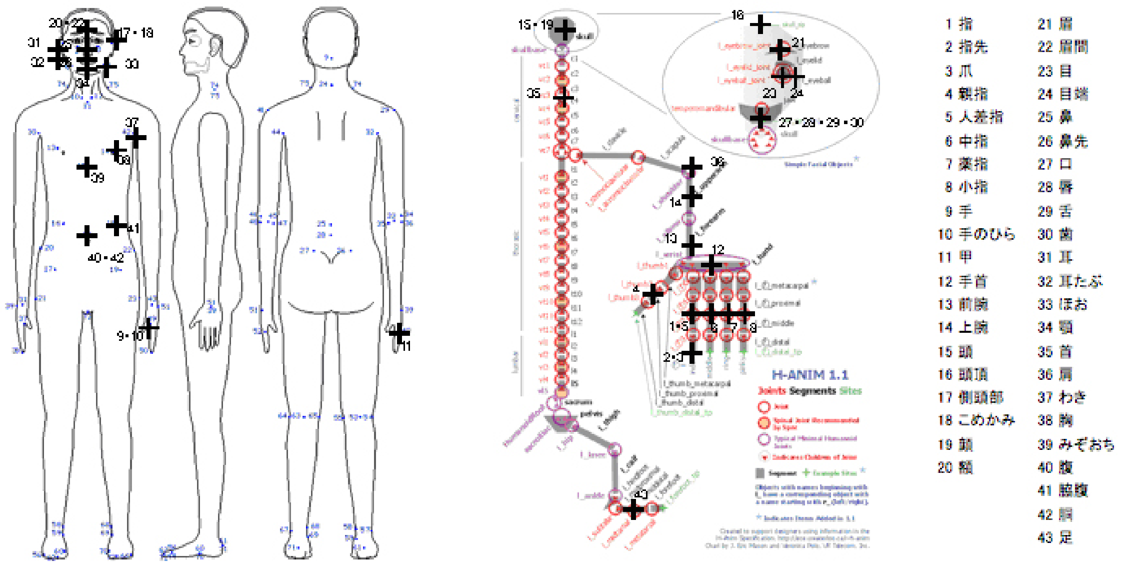

Figure 6. *Applying the reference points on H-ANIM avatar.*

4. DISCUSSION

The analysis of Japanese-Japanese sign language gives 85 body positions, which can be mapped on H-ANIM standard avatar. Thus the coordinate system may denote given sign motion in a certain manner.

On the other hand, the extracted modifiers and noun may produce 19975 compound nouns to show position or orientations of hands in disregard of multiple usages of nouns and drop of possessive particle. 321 out of the 19975 candidates appear 30919 times among 16640 notations in the dictionary in total. Regarding that the obtained notation including many errors of optical character reader, the proposed coordinate system can denote most of sign expressions appear in Japanese Sign Language Dictionary.

However, the further investigation found that the 321 compound nouns are also used to denote posture of hands. Thus, further morphological analysis regarding with verbs related to the 321 candidates may results in reducing number of nouns and modifiers to compose expression of position or orientation of hands. Additionally, further investigation of sign notations appears that some signs are expressed figuratively without using the compound words. To examine whether the obtained nouns are sufficient to show most of signs, to denote such signs using obtained reference points are required.

Although this paper proposed and mapped new coordinate system logically suitable for signs, sign language or human posture can be quite complicated. For example, to show “eye”, a signer can indicate his/her eye from any direction of eye if the fingertip of index is 0pt (or small amount) from targeting eye. Thus, to define motion of sign posture, to define dominant part of hand is indispensable. However, as the dominant part can be defined as nearest point to the reference point, or dependent on hand posture itself, automatic conversion from “raw” data into the proposed coordinate may be realized to a certain extent.

5. CONCLUSION

This paper proposed new coordinate system suitable for sign animation, and defined possible origin of coordinates through morphological analysis of Japanese Sign Language Dictionary. The investigation clears that the proposed coordinate can proportional to avatar and be applicable for most of sign expressions. The authors are performing further morphological analysis to refines possible reference points and to clear sufficiency of obtained reference points. Additionally, the authors are developing coordinate conversion to evaluate availability of the proposed coordinate system.

Acknowledgements: This research is partly supported by Grant-in-Aid for Young Scientists (A) (17680008) from The Ministry of Education, Culture, Sports, Science and Technology, Japan.

6. REFERENCES

P Ekman and W V Friesen (1978), *Facial action coding system: A technique for the measurement of facial movement*, Consulting Psychologists Press, Palo Alto.

T Hanke (2004) HamNoSys; representing sign language data in language resources and language processing contexts, *Proc. LREC 2004*, pp.1-6.

T Kurokawa (1992) Gesture Coding and a Gesture Dictionary for a Nonverbal Interface, *IEICE Transactions of Fundamentals of Electoronics*, 75, pp.112-121.

T Kuroda, K Sato and K Chihara (1998), S-TEL an avatar based sign language telecommunication system, *Int, J. Virtual Reality*, **3**, *4*, pp. 21-27.

T Kuroda, Y Tabata, A Goto, H Ikuta and M Murakami (2004), Consumer price data-glove for sign language recognition, *Proc. ICDVRAT*, pp.253-258.

T Kuroda, W Makari, Y Manabe and K Chihara (2001), Transformation of Signing Avatar to Fit User's Figure, *Moroccan J. of Control Computer Science and Signal Processing*, 3, *1*.

ISO/IEC (2001), Information Technology – Coding of Audio-Visual Objects- Part2: Visual, ISO/IEC 14496-2.

Y Matsumoto, A Kitauchi, T Yamashita, Y Hirano, H Matsuda, K Takaoka and M Asahara (2001) Moriphological Analysis System ChaSen Manual, Nara Institute of Science and Technology.

Y Nagashima and K Kanda (2001), Present State and Issue of Sign Linguistic Engineering, *Proc. HCI International*, **3**, pp. 387-391.

C Schmaling, T Hanke, (2004), HamNoSys 4.0, http://www.sign-lang.uni-hamburg.de-/Projekte/HamNoSys/HNS4.0/englisch/HNS4.pdf.

W C Stokoe, D. Casterline and C Croneberg (1965), *A Dictionary of American Sign Language on Linguistic Princeples*, Gallaudet University Press, Washington DC.

V Sutton (1998), The SignWriting Literacy Project, *Proc. Cognition AREA Conf.*

T Takemura and S Ashida (2001), A Concept of Medical Corpus for Natural Language Processing in Medicine, *Proc. the 3rd International Workshop on Enterprise Networking and Computing in Health Care Industry*.

A Yonekawa (ed) (1997), *Japanese-Japanese Sign Language Dictionary*, Japanese Federation of the Deaf, Tokyo.