6—3 A Retrieval Method for Japanese Signs Using Japanese Verbal Descriptions

Hisahiro Adachi and Kazuo Kamata
Dept. of Information Science, Utsunomiya University
adachi@galileo.infor.utsunomiya-u.ac.jp

ABSTRACT

One of the inherent problems in constructing the sign language dictionary is how to make retrieval and comparison operations on the visual database of signs. This paper describes a retrieval method, especifically for Japanese signs. This method has a useful capability for flexible retrieval of the sign from a bilingual dictionary. Our method can retrieve similar signs to the given input. The retrieval mechanism is essentially based on similarity between the given verbal description and verbal descriptions in a retrieval database. The similarity measure of verbal descriptions can be considered as the approximations for the similarity of sign motion images. As a result of our experiment, the success ratio of the retrievals is 96 % in average.

1 Introduction

Sign language(SL) refers to a kind of visual languages used by many deaf people in deaf communities as their native language. Furthermore, it can be considered as an important communication path to communicate with the hearing people. Therefore, the hearing people learn to communicate by the SL. Several Japanese-to-Sign bilingual dictionaries are published in printed paper media or their CD-ROM formatted version. However, it is difficult for learners to look up the corresponding Japanese headword of an unknown sign by the motion of sign as search key. This problem is derived from the inherent characteristics of sign languages. That is, both the perception and production of their languages use visual representations. Learning sign languages as a second language, the learner usually wants to look up the following questions:

- Acquiring one sign, "what is the other sign with similar motion and is difference between the signs?"
- "What does a sign motion mean (that is, Jpanese headword)?"

However, existing sign dictionaries don't answer the above questions. While, a few researchers have provide a precedent in the use of a notation system devised by William Stokoe [1], an important pioneer in the phonological analysis of American

Sign Language(ASL). We may also note that his notation system was used for an important dictionary of ASL, the first of its kind in which signs could be looked up according to their formation components rather than according to the alphabetic order of their English translations. Most of previous retrieval methods for signs were based on phonological features[1]. These approaches are the **phonological feature-based** retrieval methods, which need the phonological analysis. However, it is unwieldy for novice users to specify several feature parameters. In this paper, we propose a new retrieval method for signs. The method is based on the similarity between the Japanese verbal descriptions of signs.

We describe how to retrieve signs using the Japanese verbal descriptions(JVD). Figure 1 shows the basic mechanism of the method. It consists of two parts: (1) the retrieval database is the set of the verbal descriptions extracted from a Japanese-to-Sign dictionary, and (2) the retrieval engine is to retrieve the appropriate verbal description that is similar to an given input JVD. In our previous work [3], we show how to classify signs using JVD[2]. The basic idea of the work is that the similarity of signs approximate to the similarity of their verbal descriptions. Results of the work also show the semantic similarities were often correlated with the articulatory similarities. The sign motions can be represented by a Japanese verbal description such as "右手を口の前で二度左右に動かす (in front of a mouth, sway the right hand from side to side)". By computing the character-based longest common subsequences against the Japanese verbal descriptions, the similar signs ("渋い (astringent)","辛い (salty)","苦い (bitter)", and "甘い (sweet)") are retrieved. Obviously, the most appropriate candidate is the sign "苦い (bitter)". However, the rest of candidates are very helpful for us to learn sign language. Because, the learner can understand simultaneously the commonality or distinctive feature among similar signs (e.g., the commonality is "taste concept").

Compared with the phonological feature-based methods, the **verbal description-based** method has the following advantages:

(1) Phonological or morphological feature analysis is unnecessary.

(2) It is convenient for learner to use natural language instead of the formal language "Structured Query Language(SQL)" as the user-interface.

(3) Simultaneously, the similar signs with motion or semantics can be retrieved and learned.

2 Similarity Measure

2.1 Verbal Descriptions

Typical Japanese sign dictionaries are divided into two parts. That is, the one is a visual description, the other is Japanese verbal descriptions of a sign (JVD) as shown in Figure 1. This JVDS can be considered as characteristic sequences derived from the series of sign motions. In

[1] e.g., the handshape, the location and the movement of hands

[2] the term is usually called as "Manual Motion Descriptions(MMD)"

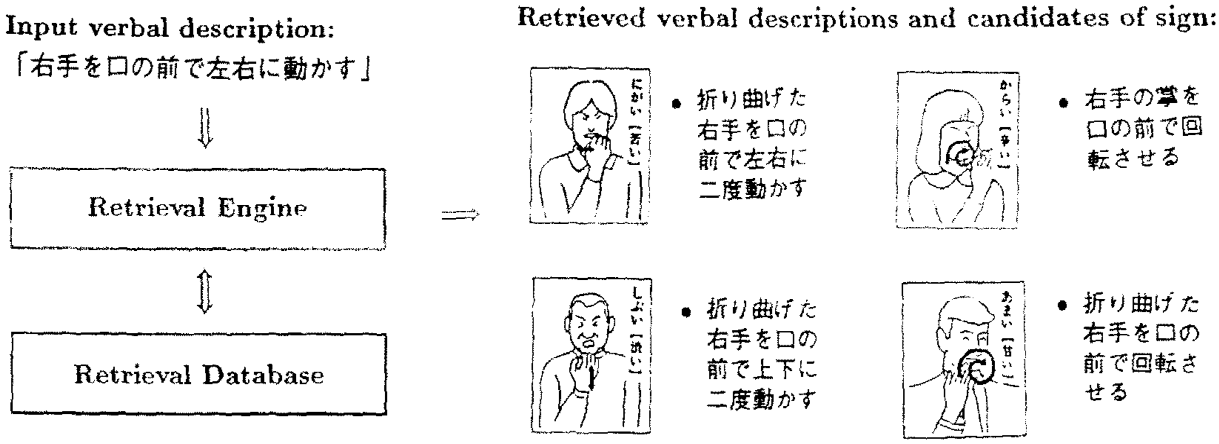

Figure 1: The Outline of A Retrieval Method Using JVDs

other words, the structure of sign images are mapped into the natural language sentence. Furthermore, JVD have the remarkable characteristics as follows.

- **constraints on kana- and kanji-characters**

  (1) indicate the locations of hands related to the body "鼻 (nose), 口 (mouth)", (2) indicate the movements of the hands "上げる (up), 下げる (down)", (3) indicate the directions of movements "右 (right), 左 (left)", (4) indicate the degrees of behavior "強く (strong), 弱く (weak)".

- **constraints on syntactic patterns**

  JVD can be considered as a kind of programming language to generate sign motion images. Therefore, the description form can be characterized according to the concatenate order of transition states in sign motions. That is, the arrangement order of case markers "を (WO), から (KARA), に (NI)" has a kind of constraints

Thus, JVD can be considered that their verbal descriptions have potentially more semantic and syntactic constraints than general Japanese sentences.

2.2 Character-Based Longest Common Subsequence

A subsequence of a given string is any string obtained by deleting zero or more symbols from the given string. An LCS(Longest Common Subsequence) of two strings is a subsequence of both that is as long as any other common subsequence. There is one further point that we must not ignore. An LCS means that the number of matching symbols is considering the symbol order constraint.

Let $A = a_1 a_2 \cdots a_m$ and $B = b_1 b_2 \cdots b_n$ be string sequences. For a given sequence $X = x_1 x_2 \cdots x_l$, we define the ith prefix of X, for $i = 0, 1, \cdots, l$, as $X_i = x_1 x_2 \cdots x_i$. For example, if $X = abcde$, then $X_3 = abc$ and X_0 is the empty sequence. Then, an LCS of A and B, denoted by $LCS(A, B)$, can be computed efficiently as the following recursive formula using Dynamic Programming Method [2].

$$LCS(A,B) = c(m,n) \quad (1)$$

$$c(i,j) = \begin{cases} c(i-1,j-1)+1 & \text{if } a_i = b_j \\ \max(c(i,j-1), c(i-1,j)) & \text{if } a_i \neq b_j \end{cases}$$

$$(1 \leq i \leq m \wedge 1 \leq j \leq n)$$

Then, similarity $S(A,B)$ can be defined as follows:

$$S(A,B) = \frac{LCS(A,B)^2}{mn} \quad (0 \leq S(A,B) \leq 1) \quad (2)$$

3 Retrieval Method

A binary relation that is reflexive, symmetric and transitive is called an equivalence relation. It is well known that an equivalence relation groups elements which are equivalent under the relation into disjoint classes. That is, the relation can be considered to effectively group elements into sets whose members are "similar" to each other to some specified degree. In contrast, the similarity measure $S(A,B)$ can be considered as a binary (similarity) relation, which clearly satisfies reflectivity ($S(A,A) = 1$) and symmetric($S(A,B) = S(B,A)$). However, $S(A,B)$ is not transitive. Then we introduce the following inequality, called "max-mini composition".

$$S(A,B) \geq \max\min(S(A,X), S(X,B)) \quad (3)$$

3.1 Retrieval Database

We describe how to create the retrieval database. Let $X = \{a,b,c,d,e\}$ be a finite set of signs.

1. a similarity relation $S(X,X)$ defined on $X = \{a,b,c,d,e\}$ is represented by a similarity matrix **M**. Since the matrix is symmetric and all elements on the main diagonal are equal to 1, the relation is clearly reflexive and symmetric.

$$M = \begin{pmatrix} & a & b & c & d & e \\ a & 1 & 0.2 & 0.5 & 0.3 & 0.8 \\ b & 0.2 & 1 & 0.3 & 0.5 & 0.3 \\ c & 0.5 & 0.3 & 1 & 0.2 & 0.7 \\ d & 0.3 & 0.5 & 0.2 & 1 & 0.2 \\ e & 0.8 & 0.3 & 0.7 & 0.2 & 1 \end{pmatrix}$$

2. The similarity matrix M can be transformed into a transitive matrix **T** by the equality (3).

3. The transitive matrix T can be transformed into the other matrix by a matrix sorting operation which rearrange the elements according to their correlation coefficients. Thus, a set of signs can be grouped using the partition T_α with the appropriate threshold α.

$$T = \left(\begin{array}{cccc|cc} & a & e & c & d & b \\ a & 1 & 0.8 & 0.7 & 0.3 & 0.3 \\ e & 0.8 & 1 & 0.7 & 0.3 & 0.3 \\ c & 0.7 & 0.7 & 1 & 0.3 & 0.3 \\ \hline d & 0.3 & 0.3 & 0.3 & 1 & 0.5 \\ b & 0.3 & 0.3 & 0.3 & 0.5 & 1 \end{array}\right)$$

3.2 Retrieval Mechanism

The matrix is convenient to represent the similarity relation. By the operation mentioned above, the transformed matrix is represented as several equivalent classes. As a result, candidates can be arranged in an equivalent class. For example, let F be an input and $S(f, X) = \{0.7, 0.2, 0.5, 0.2, 0.6\}$ be similarity grades between F and JVDS in the retrieval database. Then, $S(f, X)$ can be transformed into $T(f, X) = \{0.7, 0.3, 0.6, 0.3, 0.7\}$ by the above operations. The transitive matrix can be represented as follows:

$$T = \left(\begin{array}{c|cccc|cc} & a & e & c & f & d & b \\ a & 1 & 0.8 & 0.7 & 0.7 & 0.3 & 0.3 \\ e & 0.8 & 1 & 0.7 & 0.7 & 0.3 & 0.3 \\ c & 0.7 & 0.7 & 1 & 0.6 & 0.3 & 0.3 \\ f & 0.7 & 0.7 & 0.6 & 1 & 0.3 & 0.3 \\ \hline d & 0.3 & 0.3 & 0.3 & 0.3 & 1 & 0.5 \\ b & 0.3 & 0.3 & 0.3 & 0.3 & 0.5 & 1 \end{array} \right)$$

Let a threshold be $\alpha = 0.5$. The equivalence classes formed by the level of refinement of a similarity relation can be interpreted as grouping elements that are similar to each other. Thus, the retrievals $\{a, c, e\}$ of the input f are all similar to each other to a degree not less than α.

4 Experiments

To make discussions simpler, we extracted the JVD (129 entries) with a character "口 (mouth)" from a sign dictionary [4]. By merging the identical data ($S(A, B) = 1$), 129 entries merged into 101 entries. The total amount of the pairs of signs ($S(A, B) \geq 0.6$) are 25 pairs and a similarity matrix (31×31) is obtained. Then, this similarity matrix is transformed into a transitive matrix. We selected signs (10 entries) as shown in Figure 2 as test data. Test subjects (10 students) observed sign motions for these signs, and wrote verbal descriptions. By computing the similarities for each pair of JVD (31 entries) and them, the similarity matrix was transformed into a transitive matrix as shown in Figure 3. To make the distributions of equivalence classes clearer, the elements ($T(A, B) \geq 0.6$) are shown by '*', and the other points are masked. For example, an input description of "被験者.1(subject No.1)" belongs to the equivalence class "辛い (salty)" that includes "ソース (Worcestershire sauce), こしょう (pepper), 唐辛子 (red pepper), 渋い (astringent)". It is difficult to evaluate a retrieval method, because its effectiveness essentially depends on the learner's satisfaction. The detailed evaluation of our method is now in progress, and we show one result of experiments, here. The retrieval performance was evaluated by a binary judgment $\{0, 1\}$. That is, we evaluated whether a JVDS of subjects is the correct equivalence class or not. As a result of our experiment, the success ratio is 96 % in average as shown in Table 1.

subjects	1	2	3	4	5	6	7	8	9	10	success
sign.1	1	1	1	1	1	1	1	1	1	1	10
sign.2	1	1	1	1	1	1	1	1	1	1	10
sign.3	1	1	1	1	1	1	1	1	1	1	10
sign.4	0	1	1	1	1	1	1	1	0	1	8
sign.5	1	0	1	1	1	1	1	1	1	1	9
sign.6	1	1	1	1	1	1	1	1	1	1	10
sign.7	1	1	1	1	1	1	1	1	1	1	10
sign.8	1	1	1	1	1	1	1	1	1	1	10
sign.9	1	1	1	1	1	1	0	1	1	1	9
sign.10	1	1	1	1	1	1	1	1	1	1	10
success	9	9	10	10	10	10	9	10	9	10	96 %

Table 1: Results of Experiment

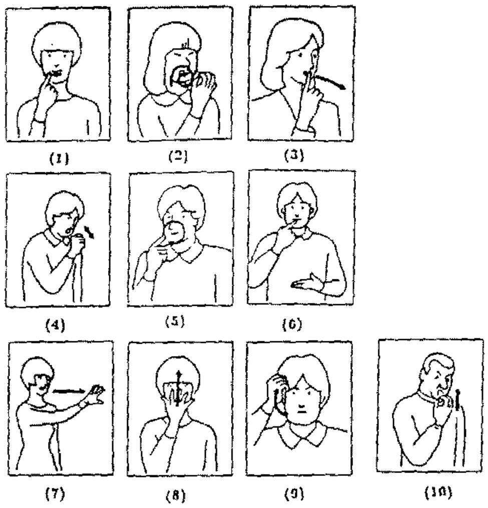

Figure 2: Samples of Sign Motion Images

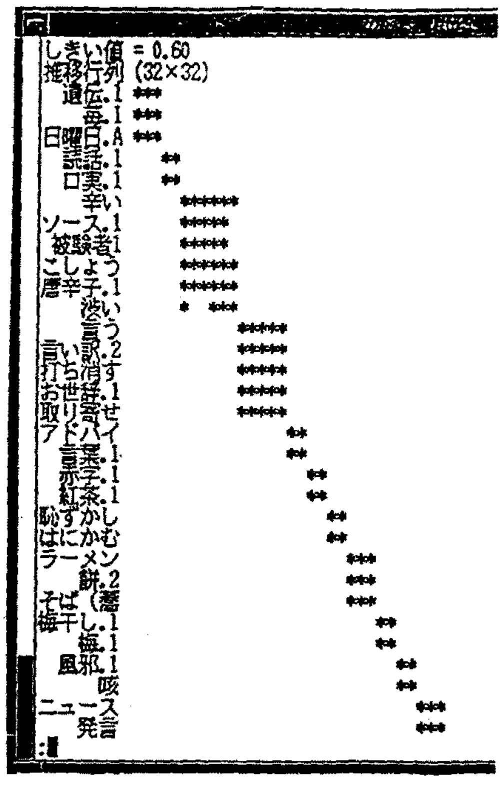

Figure 3: A Generated Retrieval Database

5 Concluding Remarks

In this paper, we have proposed a retrieval method for Japanese signs using Japanese verbal descriptions of sign (JVD). The similarity measure is derived from the character-based longest common subsequence (LCS) against JVD. As a retrieval method, we have introduced a finite set of signs divided into equivalence classes as a retrieval database on the equivalence relation. By computing the similarity between input and a JVD in the retrieval database, the equivalence classes are retrieved. Our method has the following advantages: (1) allow learns to use natural language which are familiar for novice learners as the user-interface, (2) signs with similar motions can be retrieved simultaneously, which including some kind of synonyms, antonyms and derivations.

References

[1] Stokoe, W.C: "A Dictionary of American Sign Language on Linguistic Principles", Linstok Press, 1965.

[2] Thomas H, et al: "Introduction to Algorithms", pp. 314-320, MIT Press, 1991.

[3] Hisahiro Adachi and Kazuo Kamata: "A Classification Method for Japanese Signs Using Manual Motion Descriptions", Vol. II, pp. 961-962, Proc. of COLING-94, 1994.

[4] Kouji Maruyama: "Illustrated Sign Dictionary", Dynamic Sellers, 1984.