Searching SignWriting Signs

Steven Aerts*, Bart Braem*,
Katrien Van Mulders†, Kristof De Weerdt†

*University of Antwerp
Campus Middelheim, Building G
Middelheimlaan 1
B 2020, Antwerp, Belgium
{bart.braem, steven.aerts}@ua.ac.be

†University of Ghent
Rozier 44
B 9000 Gent, Belgium
{katrien.vanmulders, kristof.deweerdt}@ugent.be

Abstract

At the moment the publication of the first written Flemish Sign Language (VGT) Dictionary is in progress. It consists of VGT glossaries and allows its users to lookup the signs for over 2000 Dutch words. The signs are written in SignWriting.
We have established an electronic representation of this sign language dictionary. Searching for signs starting from a Dutch word works straightforward. The opposite, receiving results ordered by relevance, has never been develloped before. In this paper we explain how we have worked out such a system.

1. Introduction

We have developed an online database driven dictionary system currently containing about 5000 signs(Aerts et al., 2003). These signs were (and are) collected by researchers of the university of Ghent and written down in SignWriting using SignWriter DOS(Gleaves,). Our system can convert these binary files to SWML, which is an XML-based SignWriting representation language(da Rocha Costa and Dimuro, 2003). The major advantage of SignWriting in general and SWML in particular is that it is very lightweight and thus ideal for the web. Our database is modelled on the SWML structure to contain exactly the same information.
The SignWriting system itself is a practical visual writing system for deaf sign languages, composed of a set of intuitive graphical-schematic symbols and simple rules for combining them to represent signs(da Rocha Costa and Dimuro, 2003). It was invented by Valerie Sutton inspired by her already created choreographical writing, called language DanceWritingSutton, a) Sutton, b). SignWriting symbols represent the body parts involved and the movements and face-expressions made when producing signs. This way in an electronic representation each symbol is stored along with its transformations: translation, mirror and rotation. In this paper we provide an outline to construct an intuitive, user-friendly, yet powerful search by sign system for SignWriting. All information can be extracted from the signs only, without external information e.g. the position of the hand or the dominant hand.

2. The manual approach

Searching using only the SignWriting signs has never been done before. Searching for the meaning of a sign in a database containing filmed signs manually enriched with extra semantic information, however, is common practice. It usually consists of selecting the type and direction of the movement, the location on the body where the sign is made and finally the hand form. This information needs to be added to each individual sign, causing a big slowdown in dictionary development due to tedious manual work. When creating a huge, constantly evolving dictionary this is highly undesirable.

3. Semantic view on SignWriting

The first question we should ask ourselves is whether the following common search information can be extracted from SignWriting signs:

- Type of movement: the movements are pretty well described as they are represented by different symbols, but whereas it is rather easy for human beings to find the only physiologically feasible possibility, it is difficult to find the matching moving body part with a computer.
- Direction of the movement: this is almost impossible to extract. Because of the two-dimensional representation in SignWriting the difference between horizontal and vertical movements is - again - only easily detectable by human beings.
- Location: a course grained distinction between zones of positions should be possible, but only when a body part is touched. When no body part is touched, the location can only be extracted from the SignWriting symbols by considering the most likely physiological positions.
- Hand form: is very accurately defined in SignWriting through the use of different symbols. Hand forms will obviously be the key feature to search by.

4. Elaboration

4.1. User input

The user first specifies the hand form. Then he specifies which body zones are touched (head, torso or legs) and the way in which they are touched (touch, grasp, in-between, strike, brush contact or rub contact).
It is also possible to specify the orientation of the hands (palm up, back of the hand up or side view). This does not change the essence of the search.

4.2. Processing

Selecting all signs that include the specified hand form(s) is obvious. The different types of touching are also very well depicted in SignWriting, but one type of touching has multiple variations. The body zones involved and the type of contacts are accurately specified. If multiple zones or contacts are involved, matching them is difficult.
It is however possible to parametrise the goodness of a match. The goodness-measure we would like to use is the product of the distances between the touch and the middle of the corresponding body zone. When measuring the goodness from a match of n contacts c_i $(i = 1..n)$ and the corresponding body zones z_i this results in the following function: $\mu(c_1, ..., c_n, z_1, ..., z_n) = \Pi_{i=1}^{n}((c_{ix} - z_{ix})^2 + (c_{iy} - z_{iy})^2)$

4.3. Results

We are able order matches, using this measure: the closer the match the lower the goodness-measure. Dropping very bad matches, which have a very high goodness-measure, is also possible. The ordering does not happen in a natural way (Butler, 2001; Sutton, 2004), because for that to be possible manually added information about e.g. the dominant hand would be necessary.

5. Issues

5.1. Precision & Recall

Precision is described as *the ratio between the relevant retrieved signs and the retrieved signs*, whereas recall stands for *the ratio between the relevant retrieved signs and the relevant signs*.
In this case: is the user able to clearly specify what he wants? If this is problematic, improving usability will be possible by broadening the search to closely matching symbols.

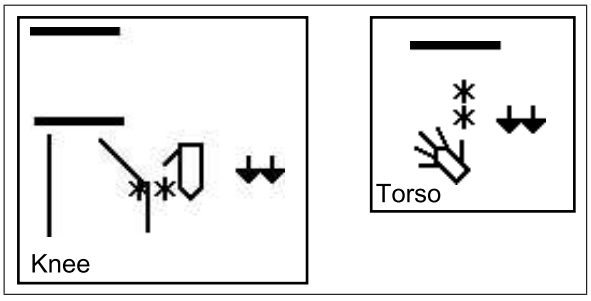

Figure 1: Determining the body zones

5.2. Performance

Selecting signs through the right symbols and contacts is a database issue where performance is not at stake. A well designed database will process those queries fast enough. Determining the body zones is rather straightforward and can also be done in the database. If we see one thick black horizontal line, for example, we know the chest is meant, whereas two lines depict the legs and hips. Figure 1 illustrates this principle.
Calculating the goodness-measure will be done over a very limited number of matching body-symbols and contacts: a sign containing four contacts is extremely rare. The number of comparisons will be low and will not affect the global performance by one order-of-magnitude.

6. Future work

We do not have an implementation of the search algorithm right now, which is the only missing part in our dictionary system. We expect to have a working system ready soon, because the method we have described is pretty straightforward to implement.
The structure of our system is built with the algorithm in mind. The exact mapping of SWML on the relational database prevents loss of information about the signs. Every single symbol can be traced back to its containing sign, allowing fast lookups of relevant signs.

7. Conclusion

The great advantage over existing systems is the fact that all information originates from the signs only.
This system is intuitive for a user with basic SignWriting knowledge. Its friendliness will largely depend on the used interface but can be improved with the goodness measure and a broadened search. The real strength of our system lies in the use of the very well specified SignWriting hand forms, which compensates for the vague movements.
Because of the use of databases, SWML and relatively simple calculations, this method is also straightforward to implement.
Most important, the Deaf Community and its researchers will benefit by this new search method since it allows for easier dictionary-searching. Moreover the system can be used as an online reference for the meaning of SignWriting signs.

8. References

Aerts, Steven, Bart Braem, Kristof De Weerdt, and Katrien Van Mulders, 2003. `http://plantijn.ruca.ua.ac.be/~aertsbraem/interface/` - Vlaams Gebarenwoordenboek.

Butler, Charles, 2001. An ordering system for SignWriting. *The SignWriting Journal*, 0(1).

da Rocha Costa, Antonio Carlos and Gracaliz Pereira Dimuro, 2003. SignWriting and SWML: Paving the way to sign language processing.

Gleaves, Richard. `http://signwriting.org/forums/software/sw44/` - signwriter dos computer program.

Sutton, Valerie, a. http://www.dancwriting.org/ - DanceWriting: Read and write dance.

Sutton, Valerie, b. http://www.signwriting.org/ - SignWriting: Read, write, type sign languages.

Sutton, Valerie, 2004. Sutton's SignSpelling guidelines 2004. SignWriting Library Database, DAC@signwriting.org.