A proposal for the recognition of handwritten SignWriting for SWORD project

Fabrizio Borgia\textsuperscript{1,2}, Maria De Marsico\textsuperscript{2}

\textsuperscript{1} Université Toulouse III - Paul Sabatier
\textsuperscript{2} Sapienza Università di Roma
Outline

- Introduction
- The SWORD project
- A new generation of SignWriting editors
- SW-OGR - Concept
- SW-OGR - Challenges
- SW-OGR - Design
- SW-OGR – In action
- Conclusions and Future

A proposal for the recognition of handwritten SignWriting for SWORD project
Introduction

- A pencil and a piece of paper are the only required items to produce signs using SignWriting.

- Since the early years the need to produce a digital version of the system was evident.

- The informatization of the system, started in 1986 with the SignWriter computer program, allowed SignWriting to achieve a wider diffusion through:
  - Newspaper
  - Books
  - Websites
  - Other digital resources
Introduction

- Digital version of the glyphs were created.
- The conceptual organization of glyph images was ensured by assigning to each of them a unique ISWA code.
Introduction

- SignWriting digital editors are the tools that enable the creation of digital resources written in SignWriting.

- In other words, they are critical for the informatization of SignWriting and therefore the “digital transcription” of SLs.

- Many applications have been produced by different teams, delivered in different ways, ranging from desktop to web applications.
Introduction

- Most SignWriting digital editors basically provide the same functionalities

- Despite differences in design and implementation existing from one editor to another, such functionalities are:
  - Search for (or type) glyphs which belong to the ISWA
  - Insert the chosen glyphs onto an area which is designated for the composition of the sign
  - Manage the glyphs on the sign composition area
  - Save the sign in one (or more) formats
  - See SignMaker and SWift

A proposal for the recognition of handwritten SignWriting for SWORD project
Introduction

Despite their increasing capabilities, SignWriting digital editors are still far from the simplicity of the handwriting approach.
Introduction

- Any SignWriting digital editor developed so far relies heavily on “Windows, Icons, Menus, Pointer” (WIMP) interfaces.

- They require a navigation/selection process in a huge amount of digital images that, though facilitated, may become long and cumbersome.

- SignWriting users are more fast and comfortable when using handwriting, rather than when dealing with the complex interaction style of a SignWriting digital editor.

- Writing is easier than understanding and using a User Interface.
The SWORD project

SignWriting Oriented Resources for the Deaf (SWORD)

- A project towards the full SignWriting-based integration of deaf people in the digital society

- The project aims at producing digital artifacts (applications, websites, etc.) featuring deaf-oriented accessibility via written Sign Language

- A SignWriting digital editor, namely SWift, was produced within SWORD
A new generation of SignWriting editors

SWift still carries most intrinsic shortcomings of its peers, including the strict dependency on WIMP interfaces.
We designed a new generation of SignWriting editors, able to partially overcome the concept of the WIMP interface and to move along the line of the so called “natural interfaces”.

- No click, drag, search and browse on the UI during the SW production process
Introducing a SW-OGR engine within an existing SW editor, such as SWift:

- The user will handwrite glyphs rather than searching them among thousands others
- Each handwritten symbol will be processed and replaced with its standard ISWA version by SW-OGR

We maintain the advantages of a digital editor

- Multiple save options: PNG, XML and remove save
- Each sign is save in a database with the list of its component glyphs: support for linguistic analysis
SW-OGR

Challenges

- WIMP is currently the most common interface style, it cannot be totally left behind
  - The interaction with most application is still firmly grounded on it

- Nevertheless, our aim is to limit its use, or to dismiss it whenever possible, during the SW production process
SW-OGR

Challenges

- Very high number of glyphs (tens of thousands)
- Multilinear/two-dimensional arrangement
- No rigid rules for composition
- Complex segmentation, particularly when handwritten

- OCR-like pattern-recognition approach not feasible
  - Large amount of patterns to recognize, and consequently very long machine training
SignWriting Optical Glyph Recognition (SW-OGR)

An application module designed to operate the electronic conversion (recognition) of user-produced images containing handwritten (or printed) SignWriting symbols into machine-encoded (ISWA) SW text.

A proposal for the recognition of handwritten SignWriting for SWORD project
SW-OGR Design

- The SW-OGR approach is based on:
  - Geometric features of the glyphs
  - Topologic information about the glyphs
  - Contextual information about the organization of the ISWA
SW-OGR
Design

A proposal for the recognition of handwritten SignWriting for SWORD project
SW-OGR Design

Image (SignWriting text) production by the User

Image acquisition interface

A proposal for the recognition of handwritten SignWriting for SWORD project
A proposal for the recognition of handwritten SignWriting for SWORD project

SW-OGR

Design

Optical glyph recognition by SW-OGR module
A proposal for the recognition of handwritten SignWriting for SWORD project

+ **SW-OGR Design**

Human-assisted review of the recognition
SW-OGR Design

- The recognition by SW-OGR is composed by different analyses of the image (SignWriting text):
  1. Image pre-processing and binarization
     - Noise reduction, thresholding
  2. Detection of geometric shapes within the image
     - Circles, rectangles, trapeziuses, etc...
  3. Category inference on geometric shapes
     - Topological and contextual information
  4. Glyph recognition
     - OGR coding to ISWA coding conversion
A proposal for the recognition of handwritten SignWriting for SWORD project

**SW-OGR Design**

- **OGR Coding**
  - An alternative coding for SignWriting symbols able to guide the recognition process
  - Only takes into account geometric/topological features of the glyphs
  - It does not substitute ISWA: it is a “process coding”

- **OGR Coding** of a handwritten glyph can be automatically calculated by SW-OGR

- **OGR to ISWA** mapping tables are required to complete the recognition

- **ISWA Bianchini** was employed for the conversion, but the recognition output can also be encoded in ISWA 2010
**SW-OGR Design**

- OGR to ISWA mapping table

<table>
<thead>
<tr>
<th>ISWA Bianchini codes</th>
<th>OGR codes</th>
</tr>
</thead>
<tbody>
<tr>
<td>09021-02-00000-0000-a</td>
<td>1.2.0.0.4.*</td>
</tr>
<tr>
<td>09021-02-00000-0000-a</td>
<td>1.2.4.0.4.*</td>
</tr>
<tr>
<td>09021-02-00000-0000-a</td>
<td>1.2.6.1.*</td>
</tr>
<tr>
<td>09021-05-00000-0000-a</td>
<td>1.2.0.3.*</td>
</tr>
<tr>
<td>09021-05-00000-0000-a</td>
<td>1.2.0.0.2.*</td>
</tr>
<tr>
<td>09021-05-00000-0000-a</td>
<td>1.2.4.0.2.*</td>
</tr>
<tr>
<td>09021-05-00000-0000-a</td>
<td>1.2.4.3.*</td>
</tr>
<tr>
<td>09022-02-00000-0000-a</td>
<td>1.4.0.0.*</td>
</tr>
<tr>
<td>09022-03-00000-0000-a</td>
<td>1.4.0.0.*</td>
</tr>
<tr>
<td>09024-04-00000-2000-a</td>
<td>2.1.0.4.0.1</td>
</tr>
<tr>
<td>09024-06-00000-1000-a</td>
<td>2.1.0.0.0.2</td>
</tr>
<tr>
<td>09021-01-00000-0000-a</td>
<td>2.2.0.0.*</td>
</tr>
<tr>
<td>09021-01-00000-0000-a</td>
<td>2.2.0.4.*</td>
</tr>
<tr>
<td>09021-01-00000-0000-a</td>
<td>2.2.4.0.*</td>
</tr>
<tr>
<td>09021-01-00000-0000-a</td>
<td>2.2.4.4.*</td>
</tr>
<tr>
<td>09021-06-00000-0000-a</td>
<td>2.2.0.1.*</td>
</tr>
<tr>
<td>09021-04-00000-0000-a</td>
<td>2.2.4.1.*</td>
</tr>
<tr>
<td>09021-03-00000-0000-a</td>
<td>2.2.0.3.*</td>
</tr>
<tr>
<td>09021-03-00000-0000-a</td>
<td>2.2.0.4.*</td>
</tr>
<tr>
<td>09022-01-00000-0000-a</td>
<td>2.2.1.4.*</td>
</tr>
<tr>
<td>09022-01-00000-0000-a</td>
<td>2.2.2.2.*</td>
</tr>
<tr>
<td>09024-01-00000-2000-a</td>
<td>2.2.1.0.4.1</td>
</tr>
<tr>
<td>09024-01-00000-2000-a</td>
<td>2.2.1.2.4.1</td>
</tr>
<tr>
<td>09024-01-00000-2000-a</td>
<td>2.2.2.0.4.1</td>
</tr>
<tr>
<td>09024-05-00000-2000-a</td>
<td>2.2.2.2.4.1</td>
</tr>
</tbody>
</table>

A proposal for the recognition of handwritten SignWriting for SWORD project.
SW-OGR
In action

Image pre-processing and binarization

A proposal for the recognition of handwritten SignWriting for SWORD project
Detection of geometric shapes within the image

A proposal for the recognition of handwritten SignWriting for SWORD project
SW-OGR
In action

Recognition output (in development)

A proposal for the recognition of handwritten SignWriting for SWORD project
Conclusions and Future

“IT... COULD... WORK!”

(Dr. Frederick Frankenstein)

A proposal for the recognition of handwritten SignWriting for SWORD project
Conclusions and Future

- SignWriting Optical Glyph Recognition is possible
- SW-OGR makes digital handwriting of SignWriting texts real
- Overcoming WIMP interface in digital editors will make the sign composition process real
- SW-OGR needs to be expanded adding support for more and more ISWA categories and groups
- Real-time text recognition and batch corpora digitalization
- Next step: transcription from videos

A proposal for the recognition of handwritten SignWriting for SWORD project