

PC BASED DICTIONARY OF PHOTOGRAMMETRY AND REMOTE SENSING

Stilla Dunkel, inpho GmbH, D - 70174 Stuttgart, Germany
Karlheinz Güthner, ABAKUS GmbH, D - 88250 Weingarten, Germany
Jürgen Peipe, Universität der Bundeswehr München, D - 85577 Neubiberg, Germany

Commission VI, Working Group 3

KEY WORDS: Terminology, Computer, Software, Machine Translation, PC Based Dictionary

ABSTRACT

Machine translation is a useful tool to generate draft translations of photogrammetric texts provided that subject-specific (photogrammetric) terms are included in the dictionary of the translation system. In the paper, the building of a microdictionary for photogrammetry and remote sensing within the Globalink translation software is outlined.

KURZFASSUNG

Automatische Übersetzung kann zur Herstellung von Rohübersetzungen photogrammetrischer Texte eingesetzt werden, vorausgesetzt daß die fachspezifischen (photogrammetrischen) Begriffe im Wörterbuch des Übersetzungssystems enthalten sind. In diesem Bericht wird ein Überblick über die Generierung eines Wörterbuchs für Photogrammetrie und Fernerkundung innerhalb des Globalink Übersetzungssystems gegeben.

1. INTRODUCTION

Machine translation of photogrammetric texts from the English into the German language and vice versa can be accomplished by means of a PC based translation system suggested by the authors a few years ago (Güthner and Peipe, 1993). The Globalink translation software is the core of the system (Globalink, 1990).

The automatic features of the program support the human interpreter and may increase his productivity. Literary masterpieces are not appropriate to automatic translation, of course, but precisely formulated technical documents lead to satisfactory draft translations representing the general meaning of the text in the target language. The software offers a comprehensive text analysis according to syntax and grammar instead of a simple word-by-word transfer. In the interactive mode, the text is translated sentence by sentence. Two windows are displayed on the PC monitor showing the text in both languages. The user is able to modify the source and the target text on-line to produce an optimized target document (see examples given in Güthner and Peipe, 1993).

Essential part of the software are dictionaries as lexical databases. A general dictionary includes approximately 65000 terms. Subject-specific user-defined microdictionaries can be added to improve substantially the quality of translation. Such a microdictionary contains the terminology of a special subject area, e.g. photogrammetry and remote sensing. If a term exists in both the general dictionary and a microdictionary, the subject-specific term has priority. All the dictionaries provide two sections, a single word data base and a semantic unit data base. Word implies the vocable itself

and its grammatical form (gender, inflection, part of speech etc.). Semantic units are phrases consisting of a sequence of single words. They correspond to semantic units in the target language and must not be translated word-for-word.

Building and enhancing the dictionaries is crucial to obtain a satisfactory quality of translation. In the Globalink software, the dictionaries can be modified using an update program which enables the user to check, add, change and eliminate terms. As an important part of the procedure, the grammatical function of any term is entered in coded form (e.g. part of speech, gender, tense, singularity or plurality of a term, etc.). When a noun or verb is to be added, information can be necessary to create correct inflection codes (Sect. 2).

In this paper, the generation of a microdictionary of photogrammetry and remote sensing is outlined. The product can be used within the Globalink translation system.

2. A MICRODICTIONARY OF PHOTOGRAMMETRY AND REMOTE SENSING

To give an impression of the building and the content of the microdictionary, the input of nouns, verbs and semantic units is briefly described in Sect. 2.1 to 2.3.

The terms have to be entered in both the English and the German dictionary because the dictionaries are monolingual and separate from one another. The connection between the source dictionary (e.g. German) and the target dictionary (e.g. English) is

Field	Input
Term	Fehler
Main Term	
GRAMMAR	
Part of speech	N
Gender	M
Person	
Number	
Tense	
Parset code	
Object case	
Reflexive	
Transitive	
TRANSLATIONS	
1) Part of speech	N
Term	error

Table 1 Input of a German single word (noun)

Field	Input
Term	observe
Main Term	
GRAMMAR	
Part of speech	V
Person	
Number	
Tense	
Parset code	200
Reflexive	
Transitive	
TRANSLATIONS	
1) Part of speech	V
Term	beobachten

Table 4 Input of an English single word (verb)

Field	Input
Term	error
Main Term	
GRAMMAR	
Part of speech	N
Person	
Number	
Tense	
Parset code	
Reflexive	
Transitive	
TRANSLATIONS	
1) Part of speech	N
Term	Fehler

Table 2 Input of an English single word (noun)

Field	Input
Semantic unit	mittlerer Fehler
Part of speech	N
Gender	M
Person	
Number	
Case/Tense	
Parset code	
Inflecting word	2
Synonym	
Translation	mean square error
Preferred translation	

Table 5 Input of a German semantic unit

Field	Input
Term	beobachten
Main Term	
GRAMMAR	
Part of speech	V
Gender	
Person	
Number	
Tense	7
Parset code	200
Object case	1) A
Reflexive	
Transitive	
TRANSLATIONS	
1) Part of speech	V
Term	observe

Table 3 Input of a German single word (verb)

Field	Input
Semantic unit	mean square error
Part of speech	N
Gender	
Person	
Number	
Case/Tense	
Parset code	
Inflecting word	3
Synonym	
Translation	mittlerer Fehler
Preferred translation	

Table 6 Input of an English semantic unit

effected by introducing the target translation term in the source dictionary (Sect. 2.4). During the input of a term alternative translations and synonyms can be entered.

2.1 Input of Nouns

Tab. 1 presents the steps necessary to input a noun in the German dictionary, in this case the word FEHLER. The table corresponds fairly with the display on the PC monitor generated by the "Dictionary Update Program" of the translation software to show or change terms already available and to add new ones.

In Tab. 1 the *Part of speech* field has code N for noun, the *Gender* field has M for masculine. The following fields need no information to be entered in this case. Under *Translations* the part of speech and the term in the target language are displayed. In addition, it can be necessary to input information on the German noun's pattern of case endings.

A similar screen is provided for the input of an English noun (e.g. ERROR) in the English dictionary (Tab. 2).

2.2 Input of Verbs

The entry of the German verb BEOBACHTEN is shown in Tab. 3. The *Part of speech* field has the code V for verb. The *Tense* field has 7 for Infinitive. Additional information can be necessary on the inflection, transitivity and reflexive nature of verbs and on some specific qualities (e.g. *Parset code* 200 means the verb can be followed by a new clause; *Object case* 1)A stands for the verb takes one object (A for accusative)). It has to be noticed that leaving an input field blank may also give an information to the update program.

The input of the English verb TO OBSERVE is displayed in Tab. 4.

2.3 Input of Semantic Units

The "Semantic Unit Dictionary Update Program" is used to enter new semantic units in the German and the English dictionary. The displays (Tab. 5 and 6) are slightly different in comparison with the single word dictionary update screens (Sect. 2.1 and 2.2). The *Inflecting word* field, for instance, contains a code indicating the position of the "head" word of the semantic unit.

2.4 Linking

The input of new terms in the dictionaries of Globalink requires a linking step. Linking stands for the process of connecting a pair of monolingual dictionaries (e.g. English and German), i.e. connecting the translation term in the target dictionary to a term in the source dictionary.

The procedure is quite simple, if a single word in the source dictionary (e.g. English) corresponds to a single word in the target dictionary (e.g. German). In this case, first the new German word is entered in the German dictionary without translation. Then the English word is entered in the English dictionary with translation and the linking is performed automatically by the program.

However, the process is more complex, if an English single word corresponds to a semantic unit in the German language (e.g. ISOCENTRES and WINKELTREUE PUNKTE). In this case, the first step is to input the English word in the "English Single Word Dictionary" together with the information that the translation is a semantic unit. Then the German term is entered in the "German Semantic Unit Dictionary" without translation. Finally, the English single word is entered in the "English Semantic Unit Dictionary" with the German semantic unit as translation.

3. TEST

The translation system and the microdictionary of photogrammetry and remote sensing have to be checked in practice by translating a great number of subject-specific texts. As an example, Tab. 7 shows the first paragraph of a close range photogrammetry text in English (Baltsavias, 1995). The German translation generated automatically in the batch mode of Globalink (Tab. 8) contains some errors but reproduces more or less understandable the content of the source text. Interactive processing and post-editing may improve the translation furthermore.

<p>During the last years, opto-electronic sensors have gained importance in areas such as photogrammetry, remote sensing, machine, robot and computer vision. They provide an effective way of fast image acquisition and the potential of automated image analysis. Digital images, in combination with on-line data evaluation, are a valuable tool for automating quality control and are of particular importance for recording kinematic processes, promising to revolutionize a wide range of measurement tasks.</p>
--

Table 7 Source text

<p>Während der letzten Jahre, opto - elektronische Sensoren haben Wichtigkeit in Bereichen wie z.B. Photogrammetrie, Fernerkundung, Maschine, Roboter und Computer-Sehvermögen gewonnen. Sie stellen einen wirksamen Weg schneller Bild-Erwerbung und das Potential automatisierter Bildanalyse bereit. Digital-Bilder, in Verbindung mit onlinen Daten sind Bewertung, ein wertvolles Werkzeug für Automatisieren von Qualitätskontrolle und sind besonderer Wichtigkeit für Aufzeichnung von kinematic Prozessen, Versprechen, um einen weiten Bereich von Messung-Tasks zu revolutionieren.</p>
--

Table 8 Target text

REFERENCES

- Baltsavias, E.P., 1995. From Pixels to Sequences, Introduction. Int. Arch. Photogrammetry and Remote Sensing, Vol. 30, Part 5W1, p. IX
- Globalink, 1990. The Globalink Translation System, German to English, English to German, Version 2.0. User's Guide and Reference Manual. Globalink, Inc., Fairfax/USA
- Güthner, K., Peipe, J., 1993. PC-Assisted Translation of Photogrammetric Papers. ISPRS Journal of Photogrammetry and Remote Sensing, 48 (4), pp. 35-40