



Deliverable Report

Ref.: DR_FRESH_WP2_2.4.4

Vers.: DL_FRESH_1.0

Date : 01/03/07

Page : 1 /3

Client : European Commission

Project : FRESH

Project N°: FP6-516059

Project Number: FP6-516059
Document number: DR_FRESH_WP2_2.4.4
Document Title: Semantic knowledge database populated with relevant information
Document status: Confidential

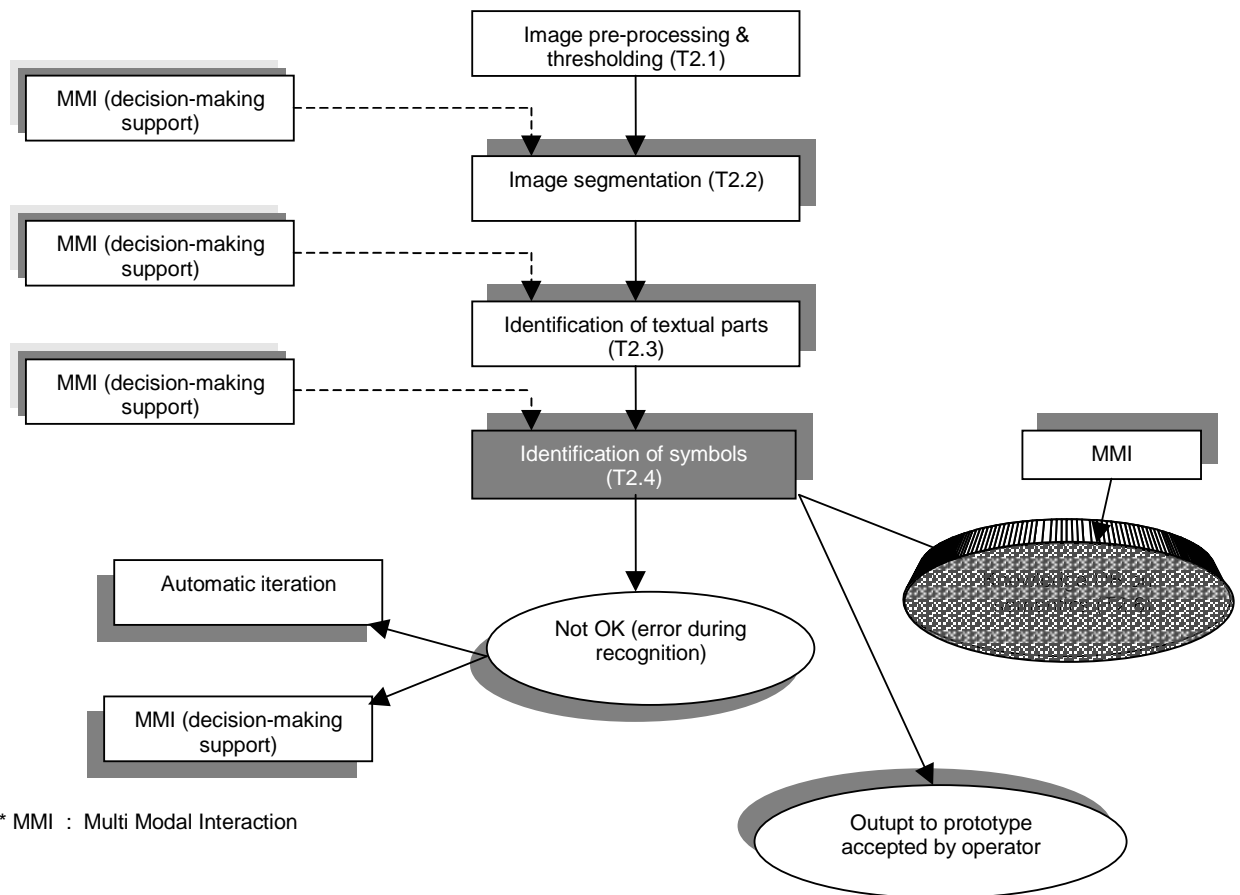
Abstract **This report contains the description of the semantic knowledge database populated with relevant information.**

Keyword List WP2, symbol, recognition, integration, specifications, documentation, semantic knowledge database.

1. PURPOSE, SCOPE AND OVERVIEW OF THIS DOCUMENT

This document contains and describes the content of the Semantic Knowledge base, needed by the WP 2 Symbol Recognition Process. It is to be noted that the current state of the Knowledge base only takes into account needs of Task 2.4 within WP 2. Further work (as described in D2.6.2 and D2.6.3 of WP 2.6) may ultimately enrich this knowledge base, as need arises. This document partially completes document D2.4.2 "Final report on symbol recognition with evaluation of performances" in the sense that it gives an exhaustive overview of known and used symbols.

Figure 1 : Task 2.4 in WP2



The Identification of symbols (or Symbol Recognition Task) has as main inputs:

- Small parts of scanned wiring plans, called Source Images, resulting from Image Segmentation Task (T2.2). These symbols are given in Win32 BMP format.
- A family of models, called Reference Images, which can be attached to some parts of Semantic Knowledge DB (T2.6). Same format as previous item: Win32 BMP. The main objective of our software is that the source images (unknown images) are recognized among the reference images.

This document details the content of the reference image database, required for recognition.



2. GENERAL REPRESENTATION AND DATA DESCRIPTION

The reference database format is the one described in D2.4.3. Please also refer to document D2.4.2 "Final report on symbol recognition with evaluation of performances" for details on descriptor computation and distance measurements.

2.1. IMAGE SIGNATURES

An image signature is a vector where each coordinate is the result of a descriptor as described in D2.4.2. These geometric image signatures are vectors of non-negative integers. The order of descriptors used inside the vector is the same for all image signatures. The set of vectors describing the set of reference images (and thus composing the Knowledge Base) are stored in a CSV file, which format given below:

```
Image 1 name; Descriptor 1 name; Descriptor 1 value; ...; Descriptor n name; Descriptor n value  
Image 2 name; Descriptor 1 name; Descriptor 1 value; ...; Descriptor n name; Descriptor n value  
...  
Image k name; Descriptor 1 name; Descriptor 1 value; ...; Descriptor n name; Descriptor n value
```

Since the descriptor name is given each time inside the signature file, the order of them is free.

2.2. REFERENCE IMAGES

The image name, mentioned in the previous paragraph, refers to a Win32 BMP in the same directory as the signature file, as described in RefDir (this is a parameter containing the path of the reference images as well as their signatures. This path is either an absolute path or a relative path from the execution directory, as per D2.4.3 section 3.2.1). Currently used reference images contain only ATA symbols, but there is no restriction on expanding the knowledge base to other symbol conventions.