FROM GEOMETRIC TO SEMANTIC INFORMATION PROCESSING IN PHOTOGRAMMETRY

B. Makarovic, Enschede

1. INTRODUCTION

Recent developments in photogrammetry are signified by the transition from fragmentary subsystems or components to larger integrated systems, and from analogue, via analytical, to digital image processing techniques. Parallel with the great advent of microprocessor technology and the consequent emergence of expert systems, a transition is taking place in photogrammetric research from the geometric to the semantic domain.

Fully automatic image interpretation is still beyond our reach, but there is a definite trend towards more automatic and fewer manual operations, i.e., to increase speed and/or to reduce the cost of geo-information (GI) production. Digital image processing techniques tend increasingly to support the human operator, and partly to replace him.

The aim of this paper is to present some thoughts on automatic image interpretation in the perspective of the changing GI-related photogrammetric technology. A state-of-the-art review of automatic image analysis and understanding is beyond our scope, but we would like to present some general considerations on image interpretation, including some basic relationships between the physical and engineering domains, the stages of progressive abstraction of semantic information (SI), and a generalized description of the photogrammetric restitution process.

Image interpretation comprises the SI specifications, descriptions of the input and output, and the interpretation procedure. The optimization of the SI specifications is essential; the specified SI items permit formulation of a datum of expectations. This datum and the additional external knowledge are linked via a suitable inference mechanism with the evidence from images. The outcomes of the inference process are the hypotheses (interpretations) on the specified SI items.

The quality of interpretation depends on the completeness and certainty of the evidence in the images, the comprehensiveness and accuracy of the a priori knowledge, and the adequacy of the interference mechanism.

2. GENERAL CONSIDERATIONS

Some basic contextual factors of a broader frame of reference for image interpretation include the physical and engineering domains and their interactions, the stages of a progressive process of SI abstraction, and the photogrammetric restitution process with its main stages in both geometric and semantic domains.

a) Interactions of domains

The items of the physical and engineering domains, represented in figure 1, are very general; they are not restricted to GI and thus to SI, nor to the corresponding photogrammetric processes. The items of these domains interact very strongly.

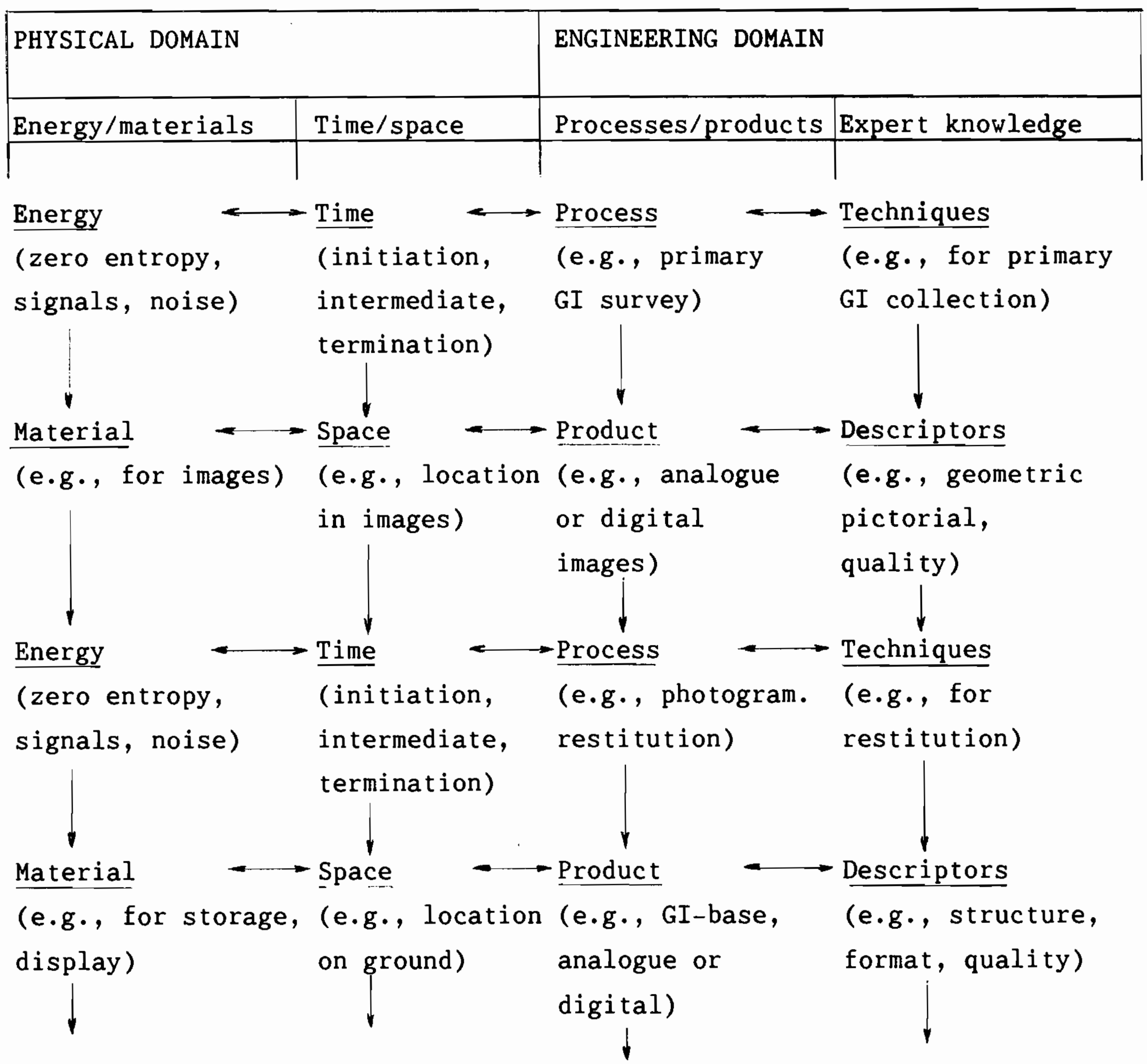

Fig 1: Physical and engineering domains and their interactions.

Expert knowledge, energy and time are involved in each process step. Descriptors (e.g., structure, form, quality), location (with metric attributes), and storage (and/or display) medium can be similarly associated with each product entity (e.g., SI item).

b) Progressive abstraction

Image interpretation proceeds gradually from low-level to high-level processes in several successive stages. This usually leads to increasing abstraction of SI, progressively from one stage to the next (figure 2).

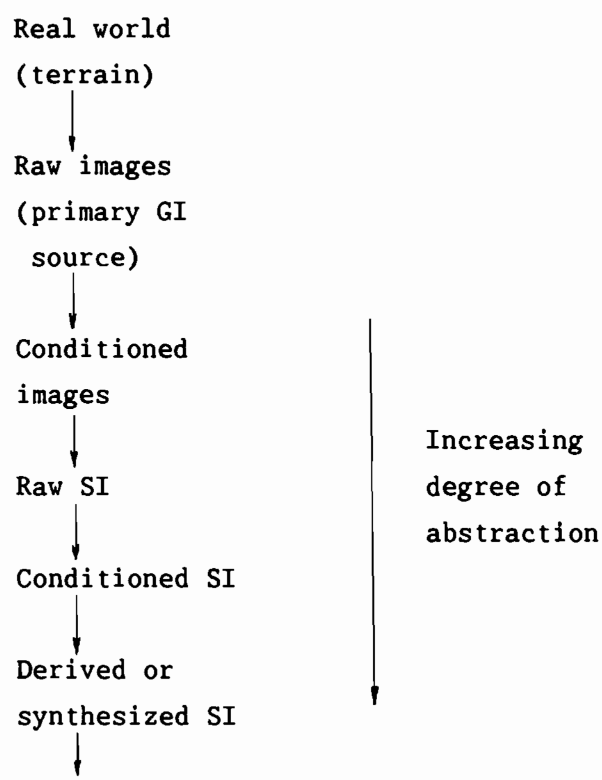

Fig 2: Stages of progressive SI abstraction.

The successive process stages commonly imply changes of the SI items, their context and some other attributes, and consequently their structure. The degree of SI abstraction is determined by the SI users' specifications.

c) Generalized procedure for restitution

Image interpretation takes a central place in most photogrammetric restitution processes. It is especially important in modelling terrain features (figure 3).

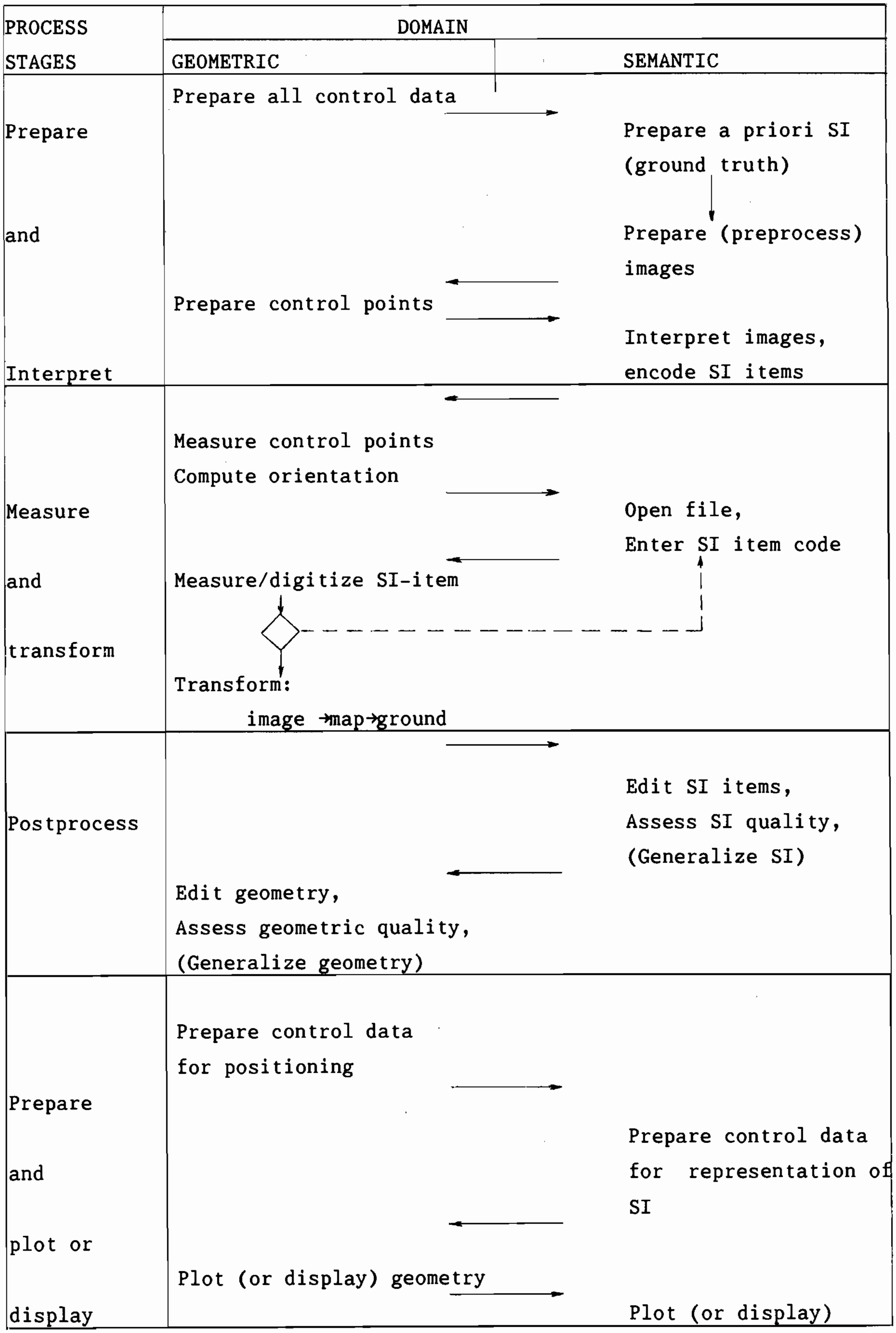

Fig 3: Generalized procedure for photogrammetric restitution.

Each main process stage contains operations in both domains. In the process of photogrammetric modelling of GI, image interpretation and geometric positioning (measurement) are usually applied alternately; thus they are unified in a common process.

d) Comparison of variants for SI collection

For a comparative study, the following three variants for acquisition of the SI are most significant:

a) Visual perception and manual interpretation directly in terrain,
b) Visual perception and manual interpretation of analogue images,
c) Automatic perception and interpretation of digital images.

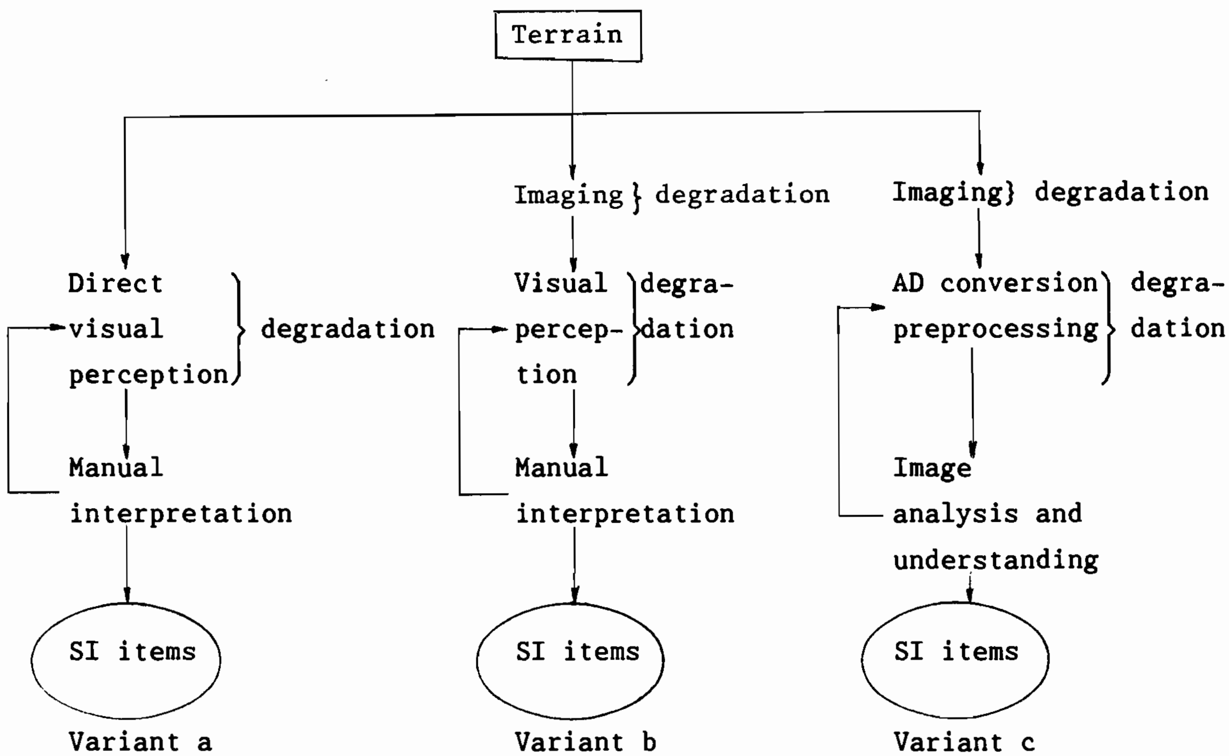

Fig 4: Variants for SI collection.

The difference between variants a and b concerns the quantity and quality of the perceived information of the same terrain area (figure 4). Hence, to collect identical SI (as specified) in both, the deficit in image information in b has to be compensated by some external (a priori) knowledge. The amount of required external knowledge is very substantial.

Variant c can be regarded as an automated version of b. An additional reduction of image information is caused by the AD conversion. This, however, requires more external knowledge to compensate for the loss. The key problem in variant c is to devise effective procedures for automatic image analysis and understanding in conjunction with the external knowledge. The feedback loops in figure 4 indicate the iterative and adaptive nature of these processes. Further considerations on image interpretation are addressed below.

3. IMAGE INTERPRETATION

The objective of image interpretation is to generate hypotheses on the specified SI. Image interpretation is the most essential and yet the least understood part of photogrammetric restitution, despite the great progress made recently, especially in the areas of neurobiology, physiology and psychology.

Image interpretation can be viewed from the perspective of photogrammetric systems engineering. Image interpretation of aerial survey images has been regarded as a routine or a skill of the human interpreter, usually a photogrammetric operator. After some training, he develops the required skill and can perform the task satisfactorily. The visual perception and manual interpretation, followed by a selective extraction of the relevant SI, are apparently unified in a single complex process.

For the design of an automated counterpart, the individual operations and their interactions in manual interpretation need to be identified and understood. At present, our understanding is still rather modest. As we gain more insight, the immense complexity and the extent of the interpretation process become apparent. Hence it is useful to address some selected aspects of the total area, including the specifications for SI (which have a strong influence on the interpretation process), the input and output of the process, and--the central issue-- the procedure for image interpretation.

a) SI specifications

The specifications should, in principle, be stated by the GI users. In practice, however, the GI suppliers and suppliers of other relevant information should be involved to attain a balance between the information supply and its use. For effective problem-solving (PS) in a user's domain, the SI should be sufficient (not excessively redundant) and the PS procedure adequate (figure 5).

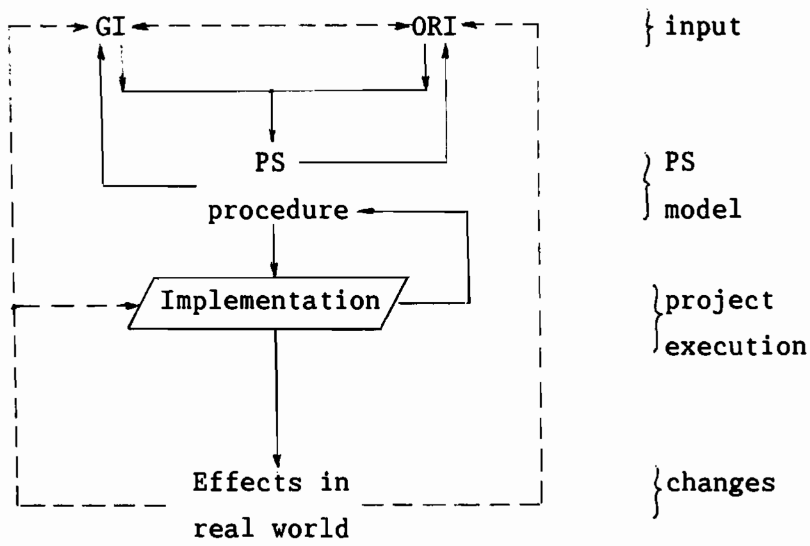

Fig 5: Information, its use and effect.

The input for problem solving comprises GI and other relevant information (ORI), whereas the outcome of problem solving is commonly decisions to be implemented in the real world. The real world, however, may change or be changed by these decisions. Thus these changes should be surveyed and entered into the GI-base and/or ORI-base for updating. The new information can affect changes in PS outcome, and subsequently the real world.

Increased knowledge can lead to upgraded problem solving procedures, which in turn may require a re-specification of the GI and/or other relevant information. To attain a state of balance, both the information input and the PS procedure should be specified as a whole (figure 6).

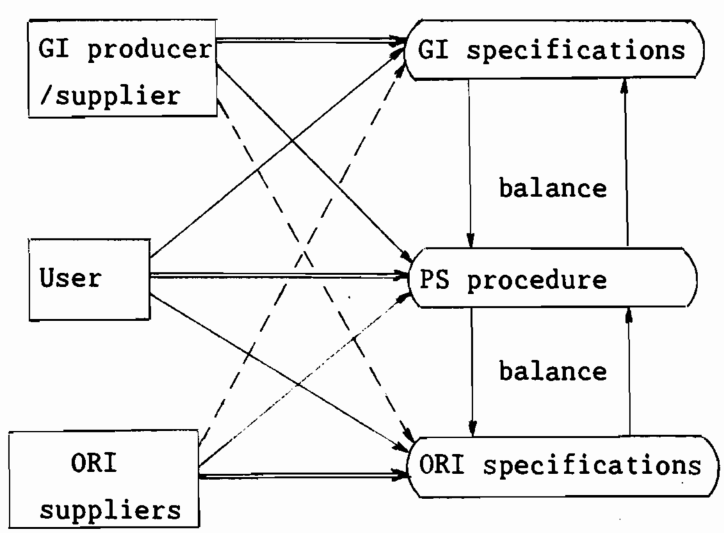

Fig 6: Balance between information input and PS procedure.

The double, full and interrupted lines indicate the different degrees of involvement of the cooperating parties. Each is supposed to be sufficiently competent in the two other fields. The optimizing process is usually iterative "from coarse to fine".

The GI specifications address both semantic and geometric domains (figure 3). (The geometric domain is beyond our scope here.) The SI specifications pertain to the SI items, their attributes, and the corresponding structure. The structures also depend on the photogrammetric restitution procedure (figure 7).

It seems important to realise that the SI structures depend on both the procedure for image interpretation and on the user's problem solving model. The link between them is established by the SI specifications. For optimizing these specifications, several feedback loops are required.

b) Input and output

Before addressing the image interpretation process itself, it seems useful to consider the corresponding input and output. The input should comprise all pertinent knowledge, including the project description and the SI specifications, and the terrain images. The knowledge is differentiated according to general (broad) and specific parts. The latter can be further subdivided into external and internal. A knowledge base contains the a priori GI base and the control data (such as rules, frames, and control mechanisms); it provides the support and controls the processes. Internal knowledge (specific) can be generated in the course of a process itself, and it can be used for testing and/or directing the further process.

A project description is required to select the specific expert knowledge (EK) and to specify the SI (key items, attributes) and its structure (figure 7). The expert knowledge about the specified SI and the imaging process permits modelling and thus prediction of the corresponding evidence in images. Such modelling provides a datum of expectations (DEX) for the specified SI items in a form compatible with the evidence in images.

The sources of the SI evidence are images of terrain and the existing GI-base. The control data include the "ground truth" of the SI which support the interpretation process and permit testing of the hypotheses. The SI evidence in images is inherent in the two-dimensional variation of the intensity (or gray level) and/or of spectral characteristics. The evidence

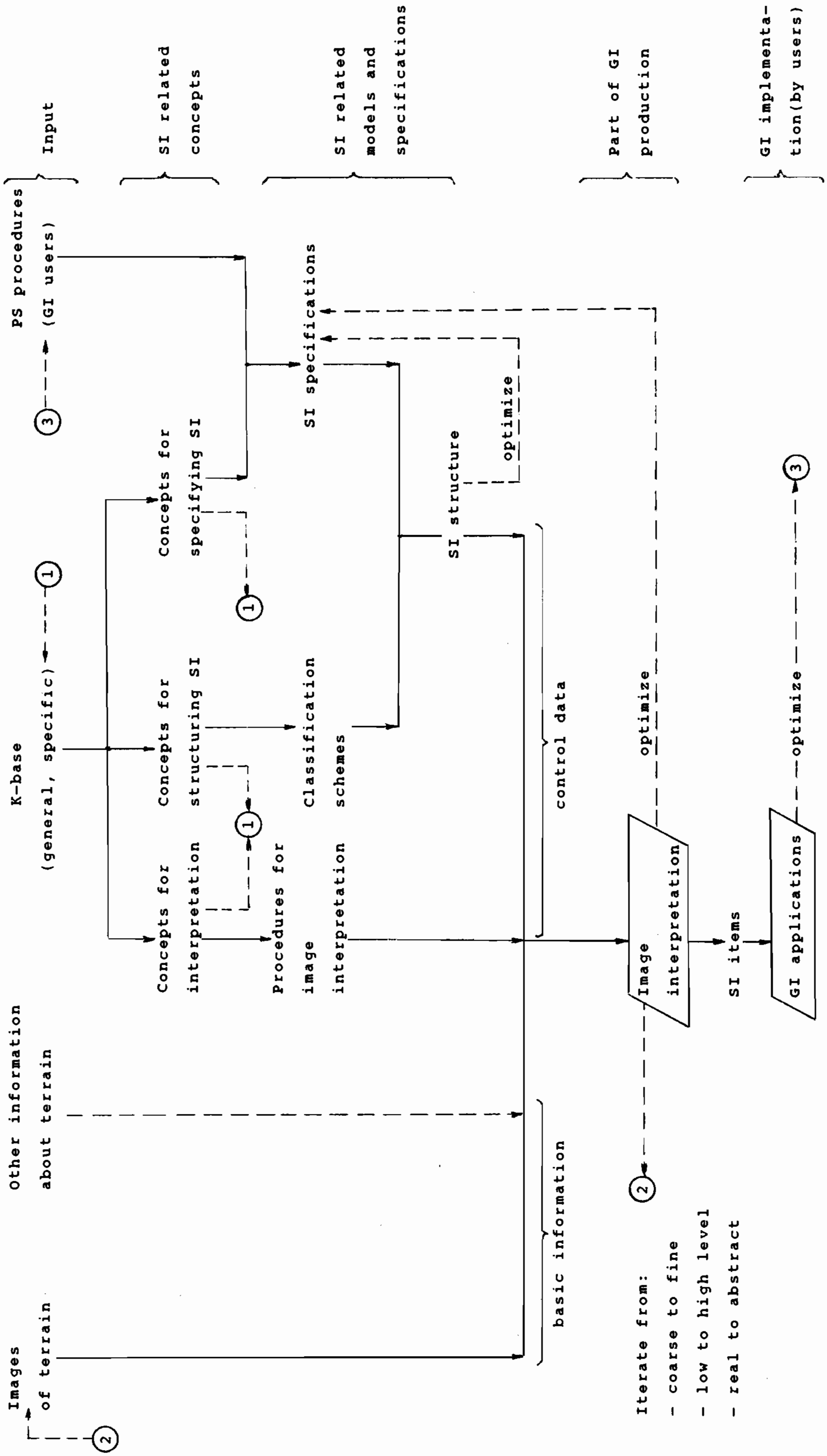

Fig 7: Optimization of SI specifications.

in images reflects both the specified SI items themselves and their context. The context represents the co-occurrence of some other features in the neighborhood of the SI item concerned. A generalized description of the context for each SI item should be included in the expert knowledge-base, and thus it should be reflected in the datum of expectations.

The output of image interpretation is hypotheses about the SI items, associated with some descriptors. One important descriptor is the quality estimate in terms of the certainty of interpretation. As a byproduct, some additional control data may emerge. All GI and the pertinent control data should be represented in a structured knowledge-base for further use.

c) Interpretation procedure

A fundamental requirement for the design of a system for automatic interpretation is sufficient understanding of the corresponding manual process. At present, however, this process is not that well understood. The hypotheses on SI items emerge from the image analysis and under- standing, which establish the link between the datum of expectations and the evidence in images. The process is supported by the external GI injected from the knowledge-base, and controlled by the specific expert knowledge.

The interpretation process comprises three main stages (figure 8): pre-process, image analysis and understanding, and post-process (the imaging process is not described here).

Each main process stage is supported and controlled by the general K-base and the specific expert knowledge (parts 1, 2, 3). The core of the interpretation process is image analysis and understanding, which implies a suitable inference mechanism. This mechanism links the evidence in images with the datum of expectations and other relevant knowledge from both the K-base and the expert knowledge-bases. The inference process proceeds usually from low-level to high-level operations and is iterative.

Raw images, and thus the preprocessed images, contain a relatively small amount of evidence pertaining to the specified SI. The deficit in evidence in images has therefore to be compensated by the information from the K-base. The latter contains by far the largest part of the information; image analysis and understanding are dominated by the information and control data from the K-base.

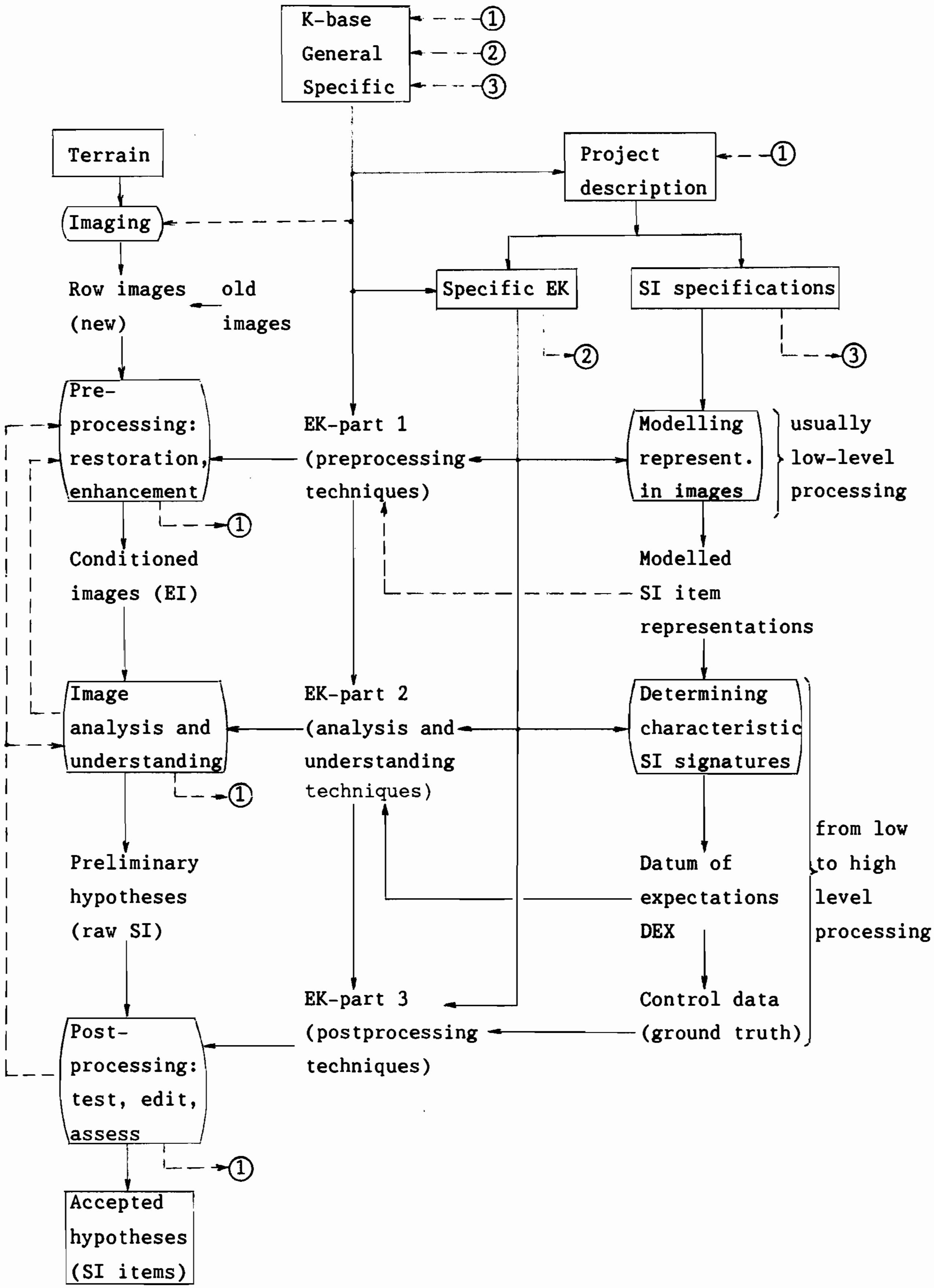

Fig 8: Image interpretation procedure

Image analysis implies a decomposition of the image into some significant components (or segments). These can be arranged hierarchically and mutually related. The processing mechanism is selected from the expert knowledge-base (part 2) in conjunction with the datum of expectations. The outcome of the analysis is entered in the process of understanding. Both processes, i.e., analysis and understanding, are intertwined.

Image understanding aims at the hypotheses about the specified SI items. The hypotheses emerge from the inference process. The corresponding mechanism is selected from the expert knowledge-base (part 2) in accordance with the DEX. The choice of the external information from the K-base and its injection into the inference process are also essential. The inference mechanism implies a number of decision stages which can be arranged into serial chains, e.g., in forward or backword sequence, a combination of both, or in networks of various degrees of complexity. As a principle, the inference mechanism can be simpler if the EI and the available knowledge are comprehensive enough and of good quality, and vice versa.

d) Postprocessing

Postprocessing is required to verify the preliminary hypotheses on the SI, to edit the accepted items, and assess the certainty and reliability of the interpretation. Verification of the hypotheses can rely on logical tests, e.g., by associating the items with their context, or on the comparison with the ground truth.

Editing implies removal of the incorrect and non-relevant SI, upgrading, sorting and structuring SI, and its merging with other information. Assessment of the certainty and reliability of image interpretation requires adequate quality control. The certainty of a hypothesis depends on the completeness and quality of the evidence in images, the injected external knowledge, and the suitability of the inference mechanism. To each entity of evidence in images a certainty estimate (or probability) can be assigned. These estimates can then propagate throughout the inference process, resulting in a certainty estimate for each hypothesis. The approach is analogous to that of error propagation for the assessment of the accuracy in the geometric domain. Reliability of interpretation is tangible in terms of completeness and failure rate. The completeness is quantifable by the ratio of the correct interpretations against the number of the factual items in the terrain (for a certain test area). The failure rate is similarly determined by the ratio of the failed interpretations against the number of factual items in terrain.

4. CONCLUSION

The emerging digital photogrammetry implies increasingly a transition in research from geometric to semantic information. Automatic image interpretation thereby thereby a key problem of great complexity. Image interpretation is involved in most photogrammetric restitution processes. It integrates external knowledge with the SI specifications, and links them with the evidence from images. The knowledge base is by far the largest part of an interpretation system. The existing means for storage, communication and processing are insufficient for the representation and handling of such extensive knowledge.

Automatic interpretation applies image analysis and understanding in the context of given SI specifications. These specifications should first be properly conditioned, i.e., to form a datum of expectations that is comparable with the evidence in images. The core of an interpretation system is the inference mechanism. For its modelling, a good understanding of the manual interpretation process is required, which is at present still rather modest.
The performance of image interpretation depends on the completeness and quality of the knowledge base, the amount and certainty of the evidence in images, the adequacy of the inference mechanism, and their mutual adjustment.
In GI photogrammetry (in contrast to close-range photogrammetry), success in automation of image interpretation has been hardly significant. Most potential applications are associated with aerial triangulation, for automatic selection of tie-points, and with automatic modelling of the terrain relief (for DTMs) to provide support in critical situations. Nevertheless, the progress made recently in research contributes to a better understanding and a deeper insight into the process of manual image interpretation. The border between the manual-interactive and automatic interpretation is expected to be gradually displaced towards more automatic, i.e., to simplify and accelerate manual interpretation.