Vision by Range and Intensity for Model-Based 3-D Object Recognition

Heinz Hügli, Philippe Gingins
University of Neuchâtel, Institute for Microtechnology
Rue Tivoli 28, CH-2003 Neuchâtel, Switzerland
E-mail: hugli@imt.unine.ch

Abstract We present a hybrid model-based 3-D object recognition approach that takes advantage of range and intensity vision. It does so in a specific way, by using range images to perform pose estimation and intensity images to perform subsequent hypothesis verification. In a first step, the recognition system takes advantage of the geometrical nature of range images to generate a number of hypothetical poses of objects which are then used, together with already available object models, to reconstruct hypothetical synthetic views of the scene. In the final step, these 2-D synthetic views are then compared to the real intensity image for verification.

Key Words 3-D vision, range images, model-based vision, vision system

1. Introduction

The goal of vision is to recognise and locate objects. In 3-D vision, two fundamental problems make the recognition difficult. The first is related to the semantic gap between signal and symbol: often symbols do not provide the needed intrinsic features in the image space. The second problem is related to the computational complexity of the recognition process.

Recent 3-D recognition approaches have tried to solve the first problem by using range images. Range images are preferred to intensity images because of their intrinsic capability to measure the true geometry of objects [1]. Despite the successes, vision systems based on geometrical information only, like range images, still are subject to a number of problems. A fundamental problem often encountered is the ambiguity of the sole geometrical description of objects that hinders to differentiate objects with identical shape, for example to differentiate candy from shoe boxes. Another problem, related to the practical work with erroneous range data, is the lack of robustness in recognition.

A tempting approach to solve these problems is to use conjointly range and intensity images. So much the more the development of fast devices that combine both range and intensity vision makes this approach very attractive [6].

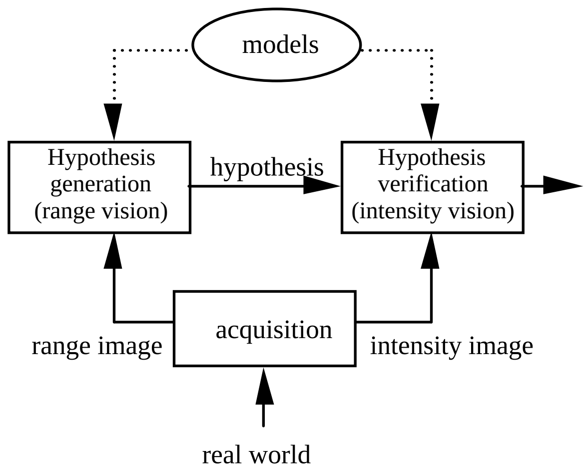

Fig. 1 Range and intensity imaging in the hypothesis generation and verification paradigm

Heinz Hügli & Philippe Gingins, "Vision by Range and Intensity for Model-Based 3-D Object Recognition", Proc. 2nd Japan-French Congress on Mechatronics, Takmatsu, Japan, Nov 1-3, 1994

In this paper, we propose a particular form of this approach where range and intensity information are combined in a hypothesis generation and verification paradigm. We consider the particular case where hypothesis generation is associated to range vision and hypothesis generation to intensity vision. Further sections develop this approach.

We see four advantages to our approach. First, the generation-verification approach brings robustness. Secondly, as a consequence of robustness, it provides simplified computation and possible speedup of pose estimation without affecting the quality of the final recognition. Third, combining range and intensity images allows to disambiguate objects which differ solely by shape or texture. Finally, new capabilities, usually considered to be complex, are offered, like recovering objects from complex backgrounds [7].

The problem of computational complexity, mentioned above, emerges when the recognition process is heavily dominated by complex search processes while trying to compare the observed scene with available models [4][5]. To reduce this difficulty, it is important that vision takes best advantage of available knowledge by using models from which many features of the observed scene can be predicted directly from the model. Our approach therefore specially addresses application fields like robot control, telemanipulation and telepresence, all fields of activity where the models of the environment are very sophisticated and the knowledge of the environment is nearly complete and certain [3].

Recognition system

Figure 1 illustrates the paradigm behind our recognition system. The acquisition module provides information from the real world as both intensity and range images. The range information is used in the hypothesis generation module which is responsible for providing hypotheses for recognised objects, in the form of an

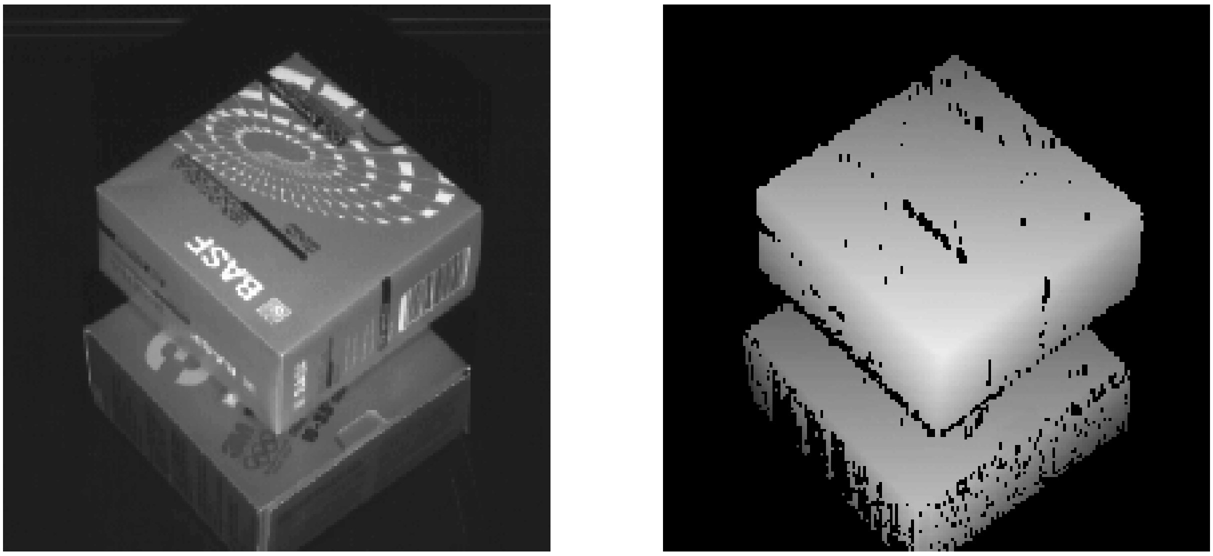

Fig. 2 a) Intensity image b) Range image estimate of pose and class of object. The module generates hypotheses in the sense that the solutions it provides are not necessarily unique nor do we expect they are fully correct. This assumption alleviates the constraints on the 3-D recognition and permits to make the respective recognition module simple and fast enough.

Further on the processing path we find the verification module. Its purpose is to verify the validity of each emitted hypothesis. It does so by comparing the real intensity image with a synthetic image generated according to the hypothesis interpretation. This interpretation involves the knowledge about objects and scene, shortly labeled as models in figure 1. The result is a set of verified and compatible hypotheses.

Functionally, the hypothesis generation module is a 3-D shape recognition system which processes range images. It performs two main steps. First a segmentation method finds smooth surface patches in the range image. We use a direct and simple method that involves only local operations [2]. Then, patches are grouped into sets of few patches to be used for generating hypothesis: estimated class and pose of detected objects [8].

The hypothesis verification module is composed mainly of a renderer and an image correlator. The renderer has access to the scene model and is in a position to generate arbitrary views of it. According to each given hypothesis, the renderer generates the synthetic image to be compared to the current intensity image by a fast correlator. Acceptance or rejection of a hypothesis is decided according to a score evaluated by the correlation coefficient.

Experiments

The described system has been set up and used so far in experiments involving boxes of different shapes and textures. Boxes are among the common objects whose shape and texture information is necessary in order to recognise them.

Figure 2 is a typical example of a pair of range and intensity images as they result from the acquisition module. We recognise overlapping boxes. Subsequent analysis of the range image gives rise to three hypotheses for boxes as shown in figure 3. Verification involving correlation tests for all hypotheses and for all compatible object models, finally accepts the two correct hypotheses and rejects the third.

Notice here that in order to recover from ambiguous shapes, the test with a single object model may include several image correlations. In the case of the floppy box shown, eight different images can possibly be mapped on the recognised shape.

Figure 4 illustrates the result of another example involving boxes of arbitrary sizes

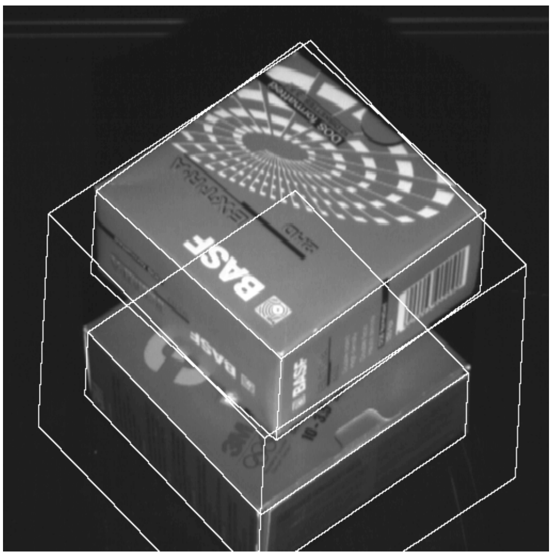

Fig. 3 Hypothetical poses

Heinz Hügli & Philippe Gingins, "Vision by Range and Intensity for Model-Based 3-D Object Recognition", Proc. 2nd Japan-French Congress on Mechatronics, Takmatsu, Japan, Nov 1-3, 1994

and orientations. Three hypotheses are first generated and two of them are accepted correctly after verification.

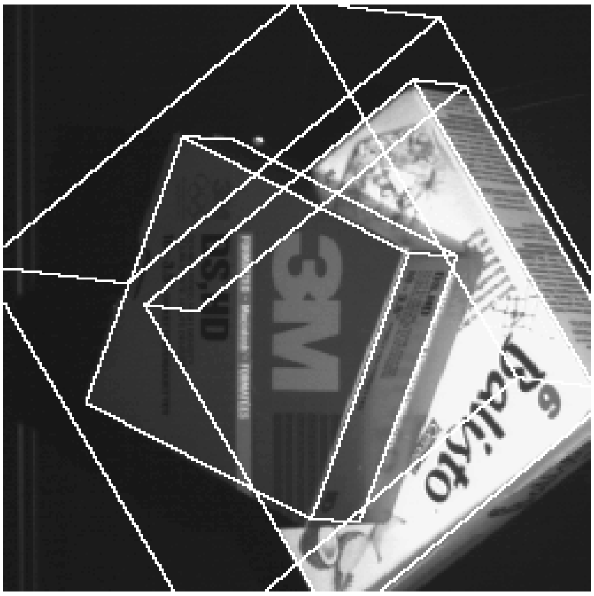

Fig. 4 Hypothetical poses

Figure 5 illustrates the performance of the system to recognise boxes of arbitrary sizes and in arbitrary positions, and more specifically to recognise objects seen under degenerated views.

In the presented experiments, all objects were recognised correctly. False hypotheses were rejected. The size estimation error of the objects is in the order of 5%. The experiments further demonstrated the advantage of the combined range and intensity vision by recognising correctly objects which are ambiguous either in shape or texture.

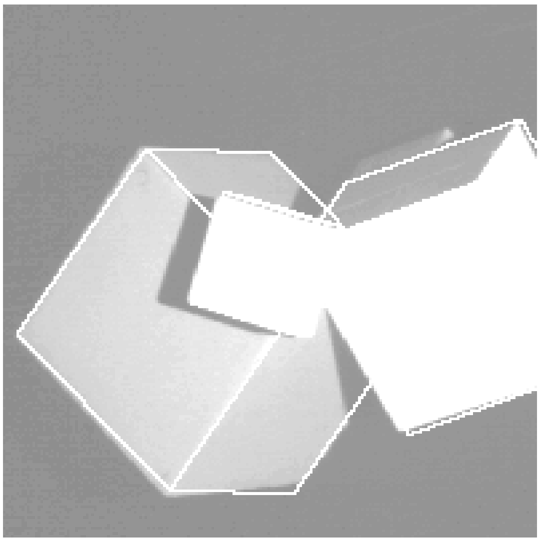

Fig. 5 Final poses

Conclusions

The presented recognition approach combines range and intensity images in a hypothesis generation and verification paradigm. It is simple, precise and robust. It can detect false poses resulting from wrong groupings, recognise simple objects that are ambiguous from their sole shape and recognise uniquely the pose of shape-symmetrical objects. The presented hypothesis generation and verification approach is particularly adequate in vision for robotics, teleoperation, telepresence, which usually have available the necessary world knowledge.

References

[1] Jain R.C. & Jain A.K., "Analysis and interpretation of Range Images", Springer Verlag, N.Y., 1990

[2] Maître G. & Hügli H., " Range image segmentation by controlled-continuity spline approximation for parallel computation", Proc. SPIE Vol 1610, Boston, 1991

[3] Natonek E., L. Flueckiger & Baur C., "Virtual reality: an intuitive apprach to robotics", Proc. SPIE 2351, Boston, Oct. 1994

[4] H.P. Amann & H. Hügli, " An algorithm for the inexact matching of high level 3D polyhedric representations", Selected papers on model-based vision, Ed. Hatem Nasr, SPIE Milestone Series, Vol. MS 72, SPIE, pp. 313-321, 1993

[5] F. Arman & J.K. Aggarwal, "CAD-based object recognition in range images using pre-compiles strategy trees", in Three-

dimensional object recognition systems, A.K. Jain & P.J. Flynn, Elsevier, 1993

[6] Beraldin J.A. & al.," Registrered intensity and range imaging at 10 mega-samples per second, Optical Engineering, Jan 1992, Vol. 31, No. 1

[7] Natonek E. & Baur C., "Recognition of 3-D objects on complex backgrounds using model based vision and range images", Proc. IEEE Southwest Symposium on Image Analysis and Interpretation, Dallas, April 1994

[8] P. Gingins & H. Hügli, "Model-based 3D object recognition by a hybrid hypothesis generation and verification approach", SPIE vol 2353, Boston, Oct. 1994