Object Recognition by a Robot Dog Connected to a Wide-Area Grid System*

J.M. Geusebroek and F.J. Seinstra
ISLA, Informatics Institute, University of Amsterdam,
Kruislaan 403, 1098 SJ Amsterdam, The Netherlands
{mark, fjseins}@science.uva.nl

Abstract

We will demonstrate object recognition performed by a Sony Aibo robot dog. The dog is connected to a wide-area Grid system consisting of hundreds of computers located at several institutes in Europe.

Object recognition is obtained by matching local histograms of color invariant features against a learned database. We effectively decompose object appearance recognition into a view based (learned) part and an appearance (invariant) part. Invariance deals with lighting conditions, color constancy, and robustness against shading effects and cast shadows. A learned set of object views guarantees recognition of different aspects of the object.

As such, we show state-of-the-art in object recognition in images, as well as state-of-the-art in multimedia Grid computing, merged together into a single application.

1 Introduction

The problem of object recognition is to determine which, if any, of a given repository of objects appears in an image or video stream. It is a computationally demanding problem that involves a non-trivial tradeoff between specificity of recognition (e.g., discriminating between different faces) and invariance (e.g., to different lighting conditions). Due to the rapid increase in the size of multimedia repositories consisting of 'known' objects, see e.g. [1], state-of-the-art sequential computers no longer can live up to the computational demands. As a consequence, high-performance distributed and Grid architectures are indispensable commodities in future multimedia computing.

In this demonstration we show state-of-the-art in object recognition applied on a wide-area Grid system, whose constituent components are scattered all over Europe. Apart from the quality of recognition, we demonstrate the effectiveness of Grid usage in multimedia computing. Moreover, we show the ease with which multimedia applications can be integrated in Grid computing.

*This work is sponsored by the Netherlands Organization for Scientific Research (NWO) and by the DELOS NoE on Digital Libraries (EU FP6).

2 Color Invariant Object Recognition

It is well known that color is a powerful cue in the distinction and recognition of objects. Recognition based on color, rather than just intensity, provides a broader class of discrimination between objects. However, the use of RGB values do not directly increase recognition performance, certainly not when variations in imaging conditions have to be counteracted. Differences in intensity, direction, and color of the illumination, as well as shading and cast-shadow significantly effect the appearance of an object. Therefore, it is meaningful to transform the RGB values to invariant properties, which relate to object (surface) properties rather than to object appearance. We previously derived a broad class of invariants [2, 3], which are shown to be robust under noisy conditions. Furthermore, these invariants can be scaled to the size of the object structure. Recognition with these invariants boils down to learning an invariant representation of the object, rather than learning every possible appearance of a single view of the object.

We decompose the recognition of object appearance into two schemes. First, we have different views or aspects of an object, each of which has to be learned. Secondly, there is the illumination, drastically influencing object appearance. For this class of appearance effects, we demonstrate invariants to be very effective. Object recognition may be based on a weak description of the important features in the scene, as long as mutual correspondence between observation and objects in the world is maintained. Therefore, we learn local histograms of invariant features for each aspect of an object.

3 Multimedia Grid Computing

In recent years, Grid computing has become an active area of research, with tremendous efforts in the installation of wide-area systems, as well as in the development of Grid middleware and Grid-enabled applications. Essentially, all of this work is performed in pursuit of the single, foremost goal of the Grid: to provide inexpensive and easy-to-use 'wall socket' computing over a distributed set of resources.

Despite the progress, Grid computing is still far from be-

0-7803-9332-5/05/$20.00 ©2005 IEEE

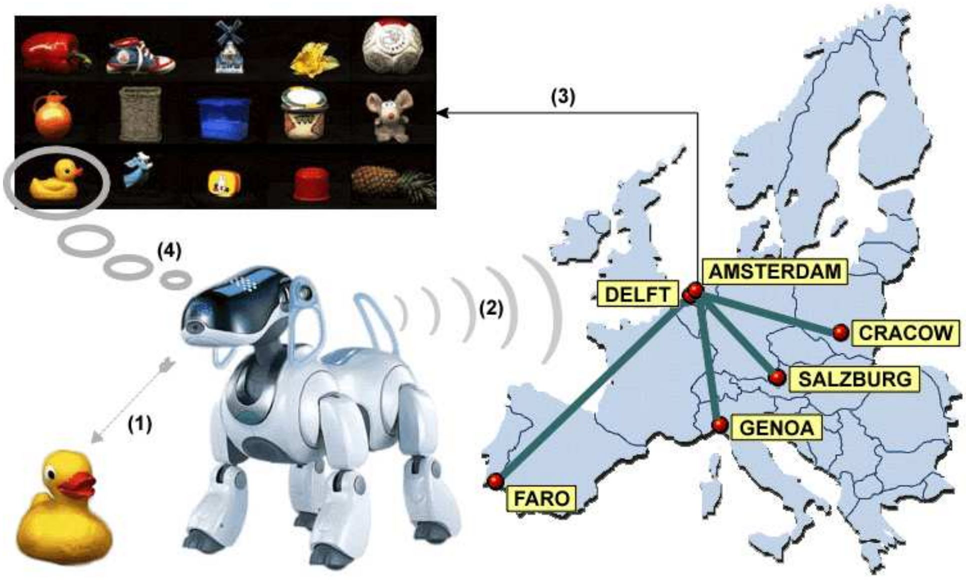

Figure 1. *Object recognition by our robot dog connected to a wide-area Grid system: (1) an object is held in front of the dog's camera; (2) the video data is scattered and processed on the Grid; (3) based on the video processing results, a database of known objects is searched; (4) in case the dog recognizes the object, it reacts accordingly.*

ing more than an academic concept. From the user's perspective, this is because Grids do not yet have the full basic functionality needed for extensive use. Consequently, as long as programming and usage is hard, most researchers in multimedia computing will not regard Grids as a viable alternative to more traditional hardware architectures.

The Parallel-Horus project [4, 5, 6] attempts to overcome this problem by providing a software architecture that allows multimedia researchers to implement *fully sequential* applications for efficient execution on Grid systems. To enhance sustainability of the developed framework without compromising on the efficiency of parallel execution, focus is on the integration of algorithmic patterns for parallel image and video processing [6], automatic parallelization and optimization [5], and domain specific performance modeling [6]. Obtained results have shown the feasibility of the Parallel-Horus approach, with close-to-optimal parallel performance for many realistic multimedia applications [4, 5].

4 Demonstration

We will demonstrate color invariant object recognition from a video stream obtained from a Sony Aibo robot dog, which is connected to a wide-area Grid system consisting of several hundreds of computers (see Figure 1). All low level image processing functionality is performed in parallel, on a per-frame basis. The resulting invariant representations are compared with 'known' objects from a database, which is accessible through the Grid.

Acknowledgements

The authors are grateful to Michiel van Liempt for his excellent efforts in implementing the original sequential object recognition code. Thanks go out to Edwin Steffens and Arnoud Visser for providing us with a robot dog.

References

[1] J. M. Geusebroek, G. J. Burghouts, and A. W. M. Smeulders. The Amsterdam library of object images. *Int. J. Comput. Vision*, 61(1):103–112, January 2005.

[2] J. M. Geusebroek, R. van den Boomgaard, A. W. M. Smeulders, and H. Geerts. Color invariance. *IEEE Trans. Pattern Anal. Machine Intell.*, 23(12):1338–1350, 2001.

[3] J. M. Geusebroek, R. van den Boomgaard, A. W. M. Smeulders, and T. Gevers. Color constancy from physical principles. *Pat. Rec. Let.*, 24(11):1653–1662, 2003.

[4] F. J. Seinstra and D. Koelma. User Transparency: A Fully Sequential Programming Model for Efficient Data Parallel Image Processing. *Concurrency and Computation: Practice and Experience*, 16(6):611–644, May 2004.

[5] F. J. Seinstra, D. Koelma, and A. D. Bagdanov. Finite State Machine Based Optimization of Data Parallel Regular Domain Problems Applied in Low Level Image Processing. *IEEE Transactions on Parallel and Distributed Systems*, 15(10):865–877, Oct. 2004.

[6] F. J. Seinstra, D. Koelma, and J. M. Geusebroek. A Software Architecture for User Transparent Parallel Image Processing. *Parallel Computing*, 28(7–8):967–993, Aug. 2002.