Color-Based Object Recognition by a Grid-Connected Robot Dog

Frank J. Seinstra and Jan-Mark Geusebroek
ISLA, Informatics Institute, University of Amsterdam, The Netherlands
Email: {fjseins, mark}@science.uva.nl

Description

Multimedia data is rapidly gaining importance along with recent developments such as the increasing deployment of surveillance cameras in public locations. In a few years time, analyzing the content of multimedia data will be a problem of phenomenal proportions, as digital video may produce data at rates beyond 100 Mb/s, and multimedia archives steadily run into Petabytes of storage space. Consequently, for urgent problems in multimedia content analysis, Grid computing is rapidly becoming indispensable.

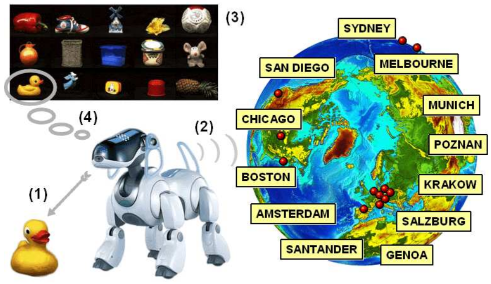
In this video demonstration we show the viability of wide-area Grid systems in adhering to the heavy demands of a real-time object recognition task. Specifically, we show a Sony Aibo robot dog, capable of recognizing objects from a set of learned objects, while connected to a Grid system comprising of cluster computers located in Europe, the United States, and Australia. As such, we demonstrate the effective integration of state-of-the-art results from two largely distinct research fields: multimedia content analysis and Grid computing. See also: http://www.science.uva.nl/˜fjseins/aibo.html.

Color-Based Object Recognition

Color is a powerful cue in the recognition of objects. Recognition based on color, rather than just intensity, provides a broader class of discrimination between objects. The use of RGB values, however, does not directly increase recognition performance, certainly not under varying imaging conditions. Differences in intensity, direction, and color of the illumination, as well as shading and cast-shadow significantly affect the appearance of an object. Therefore, it is meaningful to transform the RGB values to invariant properties, which relate to surface properties rather than to object appearance. We previously derived a broad class of invariants [1], [2], 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.

In our demonstration, we learn a set of local histograms of invariant features for each aspect of an object. Each local histogram is modeled by an integrated Weibull type distribution. The Weibull density parameters β and γ indicate the (local) edge contrast, and the (local) textureness, respectively.

In the first 'learning' phase, we present one object under a single visual setting. The obtained Weibull parameters are



stored in a database. In the second 'recognition' phase, we validate the learning step by showing the object again, under varying lighting conditions, lighting color, and viewing position. In this manner, our dog has learned 1,000 objects [3] from one example, while being capable of recognizing more than 300 objects of these under a diversity of imaging conditions. Interestingly, this recognition rate is higher than the recognition rate of around 200 objects reported for a real dog [4].

Multimedia Grid Computing

Previously [5], [6], [7] we have developed a software architecture that allows multimedia researchers to implement fully sequential applications for efficient parallel execution on cluster systems. This sequential programming model is supported by an easy-to-use execution model based on *Wide-Area Multimedia Services*, i.e. high-performance multimedia functionality that can be invoked transparently from sequential applications running on a desktop machine. As such, dynamic systems of distributed multimedia services, in which clients and servers can participate at will, can be created without any parallelization and distribution effort. Our robot dog application indeed constitutes a dynamic system of this kind.

References

[1] J.M. Geusebroek et al., "Color Invariance," *IEEE Trans. Pattern Anal. Mach. Intell.*, 23(12):1338–1350, 2001.
[2] J.M. Geusebroek et al., "Color Constancy from Physical Principles," *Pat. Rec. Let.*, 24(11):1653–1662, 2003.
[3] J.M. Geusebroek et al, "The Amsterdam Library of Object Images," *Int. J. Comput. Vis.*, 61(1):103–112, 2005.
[4] J. Kaminski et al., "Word Learning in a Domestic Dog: Evidence for "Fast Mapping"," *Science*, 304, June 2004.
[5] F.J. Seinstra et al., "User Transparency: A Fully Sequential Programming Model for Efficient Data Parallel Image Processing," *Concurr. Comput.-Pract. Exp.*, 16(6):611-644, 2004.
[6] F.J. Seinstra et al., "Finite State Machine Based Optimization of Data Parallel Regular Domain Problems Applied in Low Level Image Processing," *IEEE Trans. Par. Dist. Syst.*, 15(10):865–877, 2004,
[7] F.J. Seinstra et al., "A Software Architecture for User Transparent Parallel Image Processing," *Parallel Comput.*, 28(7–8):967-993, 2002.