Depth Information by Stage Classification

Vladimir Nedović¹, Arnold Smeulders¹, André Redert⁶, and Jan-Mark Geusebroek¹

¹Intelligent Systems Lab Amsterdam (ISLA), University of Amsterdam, Kruislaan 403, 1098 SJ Amsterdam, The Netherlands
²Philips Research Laboratories Eindhoven, High Tech Campus 36, 5656 AE Eindhoven, The Netherlands

vnedovic@science.uva.nl

Problem Statement

GOAL:
• exploit the inherent constraints of the 3D world to reduce the problem of scene geometry estimation from single images

APPROACH:
• observe the structure in real-world images to define geometric scene categories - stages
• to derive scene geometry of an image, classify it first into one of the stages
• the stage label provides rough background depth profile, used as a prior for more precise estimation

The Stages

We observe the structure present in real-world images in order to arrive at a limited number of geometric scene types. The structure in visual space is imposed by three crucial constraints:
• natural image statistics results in statistical regularities
• 3D viewpoint constraints limit the possibilities with respect to perspective [Holm et al. ICCV 2005]
• film rules ensure for the orthogonality of relevant lines and angles

Stage Classification

Hierarchical stage classification: Modeling of geometric, as opposed to semantic scene classes, bypasses the top-level division into indoor and outdoor scenes.

We use a Gaussian scale-space framework to extract features. Histograms of gradient magnitude are modeled by an integrated Weibull distribution, with a single visual surface observed, gradient histogram typically follows a decaying power-law distribution.

Natural Image Statistics

• There exists a direct relation between image statistics, scene structure and depth pattern [Torralba and Oliva, PAMI 2002]
• With a single visual surface observed, gradient histogram typically follows a decaying power-law distribution
• With increased depth and multiple structures present, integration over various power-laws results in a Weibull distribution [Geusebroek and Smeulders, IJCV 2005]
• Spatial image statistics will conform to Weibull pdf until depth increases to the point at which the observed samples become completely uncorrelated, resulting in a Gaussian histogram

Weibull parameters as a function of depth for textures of grass and bricks: β decreases from the point of fixation, whereas γ increases with depth.

Classification Results

For evaluation, we have used the key-frames of the 2006 TRECVID video benchmark.

Conclusions & Future Work

• We describe how the problem of scene geometry estimation from single images can be approached by first performing scene classification
• Inherent structure of the visual world, resulting from natural image statistics and viewpoint constraints, leads to only 15 typical 3D scene geometries - stages - each with a unique depth pattern
• Scene classification results on news video data indicate that stages without much variation or object clutter can be detected with up to 60% success rate; average classification rate is 28% at the stage level, and 40% at the stage-group level, yielding a baseline performance for stage classification in depth estimation
• Stage information is just a prior; next phase of the work is on stage parameter estimation that results in a more precise background depth profile