Multi-Scale Natural Images: A database and some statistics

Gustafsson, David Karl John; Pedersen, Kim Steenstrup; Nielsen, Mads

Published in:
Proceedings of 16'th Danish   Conference on Pattern Recognition and Image Analysis (DSAGM) 2008

Publication date:
2008

Document version
Peer reviewed version

Citation for published version (APA):
Gustafsson, D. K. J., Pedersen, K. S., & Nielsen, M. (2008). Multi-Scale Natural Images: A database and some statistics. In S. I. Olsen (Ed.), *Proceedings of 16'th Danish Conference on Pattern Recognition and Image Analysis (DSAGM) 2008* Department of Computer Science, University of Copenhagen. DIKU Technical Report, No. 08-10

MULTI-SCALE NATURAL IMAGES: A DATABASE AND SOME STATISTICS

David Gustavsson, Kim Steenstrup Pedersen and Mads Nielsen

DIKU, University of Copenhagen
{davidg, kimstp, madsn}@diku.dk

1. INTRODUCTION

Images contain different types of information, from highly stochastic textures, such as grass and fur, to highly geometric structures, such as houses and cars. Furthermore, most images contain a mix of geometric structures and stochastic textures.

It is well known from scale space theory that the image contents does not only depends on the objects in an image but also on the scale that the image has been captured ([7, 4]). At a coarse scale finer details are suppressed while the coarse scale structure are brought out. At a finer scale the coarse scale geometric structures are suppressed while the finder scale details are brought out.

Different image processing tools are suitable for different type of image contents. A tool suitable for a type of image content maybe useless for another type of image content. Most tools in image processing are very image content dependent. Segmentation of an image containing geometric structures calls for edge-based methods, while segmentation of an image containing texture calls for texture based segmentation methods (or pre-processing that transform the image textures to geometric structures).

A database containing an ensemble of image sequences containing the same scene captured at different scales is presented.

The main objectives (and applications) for collecting the database are:

Geometric Structure and Texture

The image contents depends on the scale that it has been captured at. By capture the same scene at different scales the image contents will differ - geometric structures will be transformed in to texture and texture will be transformed in to geometric structure. How does the image contents changes over the scales? How can the image contents be characterized in terms of geometric structure and texture? How can an image complexity measure be constructed that capture the image contents in terms of geometric structure and texture?

On what scale is a brick wall a wall and on what scale does it decompose in to a set of bricks, on what scale is a scrub a scrub and on what scale does it decompose in to a set of twigs? How can this transitions be measured using an image complexity measure.

Zoom-In and Zoom-Out

The zoom-in and zoom-out problems have received a lot of research interest in the recent years ([2]). A high resolution image should be shown on a small display in a mobile phone or camera, and a low resolution image should be shown on a large display or using a projector. Zoom-out (or sub-sampling) is also a common pre-processing step, motivated solely by computation time, in many image processing application.

Zoom-in: creating an image with higher spatial resolution from an image with lower spatial resolution.

Zoom-out: creating an image with lower spatial resolution from an image with higher spatial resolution.

Zoom-in is related to image interpolation, inpainting and texture synthesis. Common method used for zoom-out is low-pass filtering followed by sub-sampling, and block average.

What is the objective for zoom-in and zoom-out? Should the zoomed image be similar to the scene capture at the corresponding scale or should it just be visual appealing?

Segmentation - Cue Integration

Images contain a lot of edges, some of the edges are object boundaries will other edges are part of a texture. When is an edge an object boundary that can be used directly in the segmentation and the is it part of a texture (that can only be used indirectly in the segmentation)? An image complexity measure that characterize the image contents with respect to geometric structures and textures is informative for deciding if an edge is a boundary or port of a texture.

2. A MULTI-SCALE GEOMETRIC STRUCTURES AND TEXTURE DATABASE

The database contains images of the same scene captured at different scale. The camera that has been used is a Nikon D40X and three different objectives: 18-55 mm, 55-200 mm and 70-300 mm. The camera has been placed on a tripod stand facing the scene. A region of interest in the scene of such a size that it is present at all scales has been selected. The scene, with the region of interest approximately in the center, is captured at different scale by adjusting/changing the objective. The scene is captured at 15 different scale, the focal length is from 18 mm to 300 mm - roughly 4 octaves and 16 times magnification. A 1×1 regions in the least zoomed image corresponds to a 16×16 region in the most zoomed image. The image resolution is 2592×3872.

The scenes selected for the database are mostly natural images containing both man-made environments - mostly buildings - and natural environments - trees, tree trunks and bushes. In many cases the same type of scenes has been captured but with different distance between the camera and the scene, which are change the image contents captured. By varying the distance between the camera and the scene, each set of images captured using a fixed focal length will be an ensemble of natural images.

The region present in all images in a sequence has been extracted, resulting in sequences of regions containing the same part of the scene captured at different scales.

3. NATURAL IMAGE STATISTICS

To verify the soundness of the database content, a number of well known statistical properties, with some extensions, of natural images is verified on the database. The soundness of the image database is verified on the ensemble of images in the database (i.e. using all images in the database), and on the ensemble of images captured using the same focal length (i.e. on sets containing one image from each sequence).

One of the earliest result in the area of characterization of natural images is the scaling property ([5, 6, 3]). The scaling property was first formulated as power spectra of a large ensemble of natural images follow a power law

$$S(\omega) = \frac{A}{|\omega|^{2-\eta}} \tag{1}$$

where ω is the spatial frequency, and A is constant that depends on the overall contrast in the image. η is usually a small value and values close to 0.2 has been reported ([6, 3]). It should also be noted that η depends on the type of images ([8]) and that small image databases with specific contents - for example beaches and blue skies - may have η far from 0.2. The scale invariant property of natural images can also be expressed in the spatial domain using the correlation function.

It has been reported, [3], that the distribution of the partial derivatives of an ensemble of natural images can be modeled by an Generalized Laplacian Distribution

$$p(x) = \frac{1}{Z} e^{-|\frac{x}{s}|^{\alpha}} \tag{2}$$

where α and s are parameters estimated from the ensemble of natural images. The parameters s and α are related to the variance and kurtosis.

Compared with the Gaussian distribution, the Generalized Laplacian distribution (usually) has a sharper peak at zero and 'heavy tails'. Most natural images contain homogenous regions, objects under similar illumination, with similar or smoothly varying intensities which corresponds to the sharp peak at zero, at the object boundary the intensities changes rapidly which corresponds to the 'heavy tails'.

It is natural to consider how the size of homogenous regions in natural are distributed. Alvarez et. al. ([1]) analyze the size distribution of homogenous regions in natural images, in terms of area and perimeter, and they show that the size distribution of homogenous regions in natural images follow a power law

$$f(s) = \frac{A}{s^{\alpha}} \tag{3}$$

where s is the area, A and α are an image dependent parameters. The parameters α and A can be estimated by log-regression.

For ensembles of natural images $\alpha \approx 2$, for individual images the α varies. For image containing larger geometric structures α is often smaller around 1.5, while for image containing small scale texture α is often larger around 3.0.

The statistics computed on the database and on the sequences are consistent with the result previous reported, but not identical.

4. REFERENCES

[1] L. Alvarez, Y. Gousseau, and J.-M. Morel. Scales in natural images and a consequence on their bounded variation norm. pages 247–258, 1999.

[2] Tony F Chan and Jianhong Shen. *Image Processing and Analysis - variational, PDE, wavelet, and stochastic methods*. SIAM, 2006.

[3] Jinggang Huang and David Mumford. Statistics of natural images and models. *cvpr*, 01:1541, 1999.

[4] Jan J Koenderink. The structure of images. *Biological Cybernetics*, 50:363–370, 1984.

[5] D. L. Ruderman and W. Bialek. Statistics of natural images: Scaling in the woods. *Physical Review Letters*, 73(6):814–817, August 1994.

[6] Daniel L. Ruderman. Origins of scaling in natural images. *Vision Research*, 37(23):3385–3398, 1997.

[7] Bart M. ter Haar Romeny. *Front-End Vision and Multi-Scale Image Analysis: Multi-Scale Computer Vision Theory and Applications, written in Mathematica*, volume 27 of *Computional Imaging and Vision*. Kluwer Academic Publishers, 2003.

[8] Antonio Torralba and Aude Oliva. Statistics of natural image categories. *Network: Computation in Neural Systems*, 14(3):391 – 412, August 2003.