1-Motivation

The analysis of the statistical properties of natural images has yielded a great wealth of knowledge about the operation of the human visual system (HVS). To be useful in many visual science applications, images need to be acquired by a calibrated device, each pixel converted into a device-independent representation (such as the CIE1931 X,Y,Z system) and then converted into cone activation (LMS, for long, medium and short wavelength) representations.

However, to produce a truly representative and reliable dataset of cone-activation calibrated natural scenes has proven remarkably difficult.

There are currently several techniques that have produced calibrated databases, which can be summarised as:

1) Hyperspectral: exceptionally accurate images, but too few to be called "representative" e.g. publicly available hyperspectral datasets.

2) Trichromatic: more representative datasets (i.e. more image samples) based on inaccurate matrix transformations of the camera’s sensors into LMS representations.

We have generated a dataset of digital natural images which allows us to predict the camera output for a given combination of spectral reflectance and illumination. Once we calculate the sensor’s RGB and the corresponding LMS values for our set of spectral radiances of choice, we run a fitting algorithm to find the optimal polynomial transformation between the two of them. The advantage of this solution is that it can be "customised" for any "visual environment" of our choice, be it natural scenes, northern European natural reflectances, man-made pigments, or genericistic environments.

The figure shows an example for a set of 270 Munsell chip reflectances illuminated by D65 illumination.

3-Optimisation

The usual transformation from camera colour space to LMS cone activation is done by one of several 3x3 matrix transformations. This transformation is only accurate for monochromatic light (i.e. only approximated in most visual environments).

Our solution was to take advantage of the fact that we already know the spectral sensitivity of the camera sensors, which allows us to predict the camera output for a given combination of spectral reflectance and illumination.

4-Database

We have generated a dataset of digital natural images where each colour plane corresponds to the human LMS long, medium, short-wavelength cone activations. The images were chosen to represent five different visual environments (e.g. forest, seaside, mountain snow, urban, motorways) and were taken under natural illumination at different times of the day. At the bottom-left corner of each picture there was a matte grey ball of approximately constant spectral reflectance (across the camera’s response spectrum) and nearly Lambertian reflective properties, which allows to compute (and remove, if necessary) the illuminant.

The database is also available in CIEXYZ colours and is a 6 x 3 matrix to be determined.

http://www.cat.uab.cat/Datasets/