

Neil Robertson Image and Video Analysis

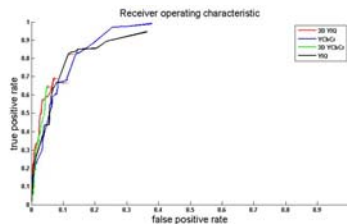
Person Detection

We detect skin regions in low-resolution images non-parametrically using an approximation to the Battacharrya coefficient to compare each pixel to a skin histogram.



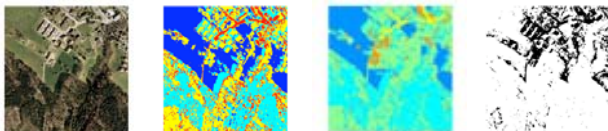
$$\rho(\hat{p}(y_i), \hat{q}) = \sum_{k=1}^K \sqrt{\hat{p}_k(y_i) \cdot \hat{q}_k} \quad \omega_{i0} = \sum_{k=1}^K \sqrt{\frac{\hat{q}_k}{\hat{p}_k(y_i)}} \delta [h(x_{i0}) - d_i]$$

A comparison of various colour-spaces with histogram size and scale shows that the 3D YIQ skin model gives best per-pixel performance.



Faces are segmented out of the total skin probability image by transforming the probability map into a binary image, detecting connected pixels and using an aspect-ratio to pick out faces.

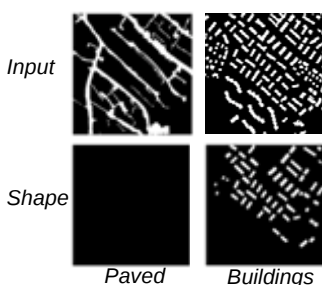
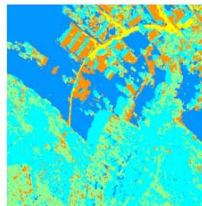
Aerial Image Segmentation for Flood Risk Analysis



We propose a new 3-step algorithm for image segmentation using a probabilistic combination of (a) non-parametric colour-similarity (b) entropy (c) gradient filter.

	Topsoil	Till	Paved	Building	Total
Mainly Topsoil	63.3931	85.0439	56.3944	59.6453	66.0898
Mainly Tilled	63.2274	89.2494	38.4452	64.8347	84.1771
Mainly urban	63.5007	79.0975	50.8026	63.4272	61.6417
Even Distribution	80.3663	92.6072	41.3953	64.0512	84.3636

Final classification percentages for the Zurich dataset as compared to ground truth data. The automatic classification is used as input to an hydrological model to compute water run-off and peak flow given a certain amount of rainfall.

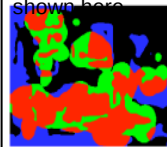


The main error in the process is due to buildings having neither suitably distinct appearance or texture. Shape can be used to distinguish between known structures.

Here the paved areas give rise to no detected predefined "building" shapes.

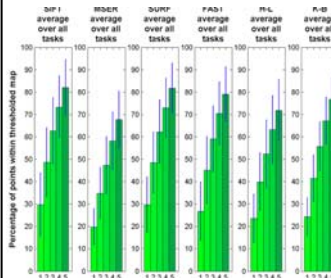
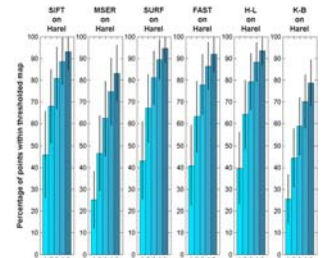
Visual Saliency

To assess the impact of image processing algorithms (e.g. compression) on the information content in an image we propose to track the changes in stable features. Some of these are shown here.



Robust features do not necessarily correlate to visually significant regions. When a human is given a visual search task these salient regions change.

We compare 2 models of saliency (Itti and Harel) with the 6 most robust feature detectors and compare the overlap. SIFT and SURF give best performance. Eye-tracker data is used to create a task-saliency map to evaluate the effect of diverting attention from the instantaneous response.



The addition of a search task reduces the overlap but the Kadir-Brady detector is clearly finding different regions from the other detectors. This indicates SURF and K-B could be used as a basis for rapidly tracking changes in information content when an image is processed.

New Projects

- ★ *Active Control of Zoom* (The Royal Society)
- ★ *Context-driven Scene Segmentation* (EPSRC & Thales)
- ★ *Embedded processing: image scene analysis* (EPSRC and Thales)
- ★ *Detection of Unusual Behaviour* (EPSRC and Roke Manor)

Publications

- *Automatic Human Behaviour Recognition*, Security Journal, 2008
- *On the Detection of Low-Resolution Skin Regions*, Proc. IEEE Int. Wkshp. Visual Surveillance, Marseille, 2008
- *Metrics for Measuring the Impact of Image Processing*, Proc. Remote Sensing, Cardiff, 2008
- *Detecting Low-Resolution Faces in Video*, Proc. Comp. Vis. & Intelligent Robotics, San Jose, 2008.

Students

- Toni Linares (PhD, Heriot-Watt)
- Jonathan Letham (EngD with Thales Optronics)
- Patrick Harding (EngD with Thales Optronics)
- Tak Chan (City University, Hong Kong)
- Nils Janssen (Darmstadt University)