Stochastic visibility in point-sampled scenes

Miles Hansard
http://www.eecs.qmul.ac.uk/~milesh/

School of EECS
Queen Mary, University of London
London E1 4NS

This work introduces a new statistical model of visibility in point-sampled scenes, such as those constructed from multiple lidar or depth-camera scans. A *visibility density* is defined along each optical ray, which gives the probability that a non-occluded scene-point exists at any particular location, with respect to a given camera. The new approach avoids any commitment to a surface-mesh, in order to develop a more data-driven probabilistic model. Furthermore, this approach naturally allows for the existence of gaps and uncertainty in the point-cloud (see fig. 2).

The new model has potential applications to multi-view stereo problems, in which visibility is an essential component of the photometric reprojection error [3]. There are other potential applications to the graphical rendering of point-cloud data [2].

Figure 2: A typical 3-D point-cloud, constructed from multi-view RGB-D scans [1], as used in the experimental evaluation.

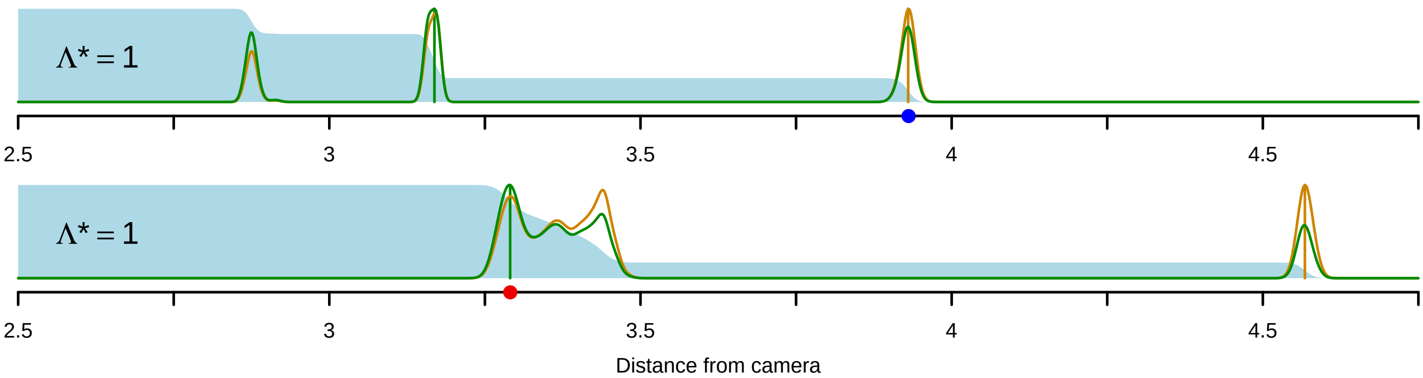

Figure 1: The visibility density (3 & 7) shown in green, for two rays in the evaluation data-set. The target-point of each ray is indicated by a dot, colour-coded by its true state; blue means occluded, and red means visible. Orange curves show the occupancy density (2 & 4), while the blue polygon indicates the attenuating vacancy density (1 & 5). Vertical lines indicate global maxima along each ray. **Top**: the distant occluded target generates the occupancy maximum (orange), but not the visibility maximum (green). **Bottom**: the nearby visible target generates the visibility maximum (green), but not the occupancy maximum (orange).

It is essential to make a careful definition of visibility, for point-sampled scenes. Let $\mathbf{c}$ be the camera-centre, and let $\mathbf{u}$ be the direction of visual ray $\mathcal{R}$. Then the *event* that point $\mathbf{p}(t) = \mathbf{c} + t\,\mathbf{u}$ is visible will be denoted $(\multimap t \,|\, \mathcal{R})$, where the dot is intended to suggest the end of a visual ray. Visibility will now be defined as the conjunction of two events. Firstly, the ray segment $\mathbf{c} + r\mathbf{u}$, where $0 \le r < t$, should be *vacant*, i.e. free of occluders. Secondly, the point $\mathbf{c} + t\,\mathbf{u}$ should be *occupied*. The following notation will be used for these events:

$$\left(\varnothing_t \,|\, \mathcal{R}\right) \;\Leftrightarrow\; \text{The ray-segment from } \mathbf{c} \text{ to } \mathbf{c} + t\,\mathbf{u} \text{ is vacant} \tag{1}$$

$$\left(\mathbb{1}_t \,|\, \mathcal{R}\right) \;\Leftrightarrow\; \text{The point } \mathbf{p} = \mathbf{c} + t\,\mathbf{u} \text{ is occupied} \tag{2}$$

Hence the *probability* of the point $\mathbf{p}(t) = \mathbf{c} + t\,\mathbf{u}$ on ray $\mathcal{R}$ being visible in scene $\mathcal{S}$ is the *product* of the vacancy and occupancy probabilities:

$$\mathrm{pr}\left(\multimap t \,|\, \mathcal{R},\mathcal{S}\right) \propto \mathrm{pr}\left(\varnothing_t \,|\, \mathcal{R},\mathcal{S}\right) \times \mathrm{pr}\left(\mathbb{1}_t \,|\, \mathcal{R},\mathcal{S}\right). \tag{3}$$

This density, which is a function of distance t, can now be developed in relation to the point-sampled scene model.

The scene $\mathcal{S}$ is represented as a mixture of N Gaussian surface-patches, the positions and orientations of which are readily estimated from the data. The 'intersection' of ray $\mathcal{R}$ with 3-D patch k defines a 1-D Gaussian, and so the **occupancy density** (2) is a 1-D mixture, corresponding to the **orange curves** in fig. 1:

$$\mathrm{pr}\left(\mathbb{1}_t \,|\, \mathcal{R},\mathcal{S}\right) = \frac{1}{N}\sum_k^N w_k\, G\big((t-\mu_k)^2/\sigma_k^2\big). \tag{4}$$

Here $G(x) = (2\pi)^{-\frac{1}{2}}\exp(-x/2)$, and the parameters (μ_k, σ_k) and weights w_k can all be computed directly from the known geometry.

If the scene consisted of *randomly* distributed particles, of radius ε, then it could be modelled as a volumetric Poisson distribution. It would follow that the probability of point $\mathbf{p}(t)$ being non-occluded would be $\mathrm{pr}(\varnothing_{\mathcal{C}}) \propto \exp(-\lambda|\mathcal{C}|)$, where $|\mathcal{C}| = \pi\varepsilon^2 t$ is the volume of a cylinder, which must be *empty*, between $\mathbf{p}(t)$ and the optical centre. This can be generalized to the case of non-uniformly distributed points, leading to a **vacancy density** corresponding to (1) and the **blue polygons** in fig. 1:

$$\mathrm{pr}\left(\varnothing_t \,|\, \mathcal{R},\mathcal{S}\right) = \exp\big(-\eta\Lambda(t)\big) \tag{5}$$

where η is a free parameter, and $\Lambda(t)$ is a generalized volume, which increases in dense regions of the point cloud. This can be defined as:

$$\Lambda(t) = \int_0^t \mathrm{pr}\left(\mathbb{1}_r \,|\, \mathcal{R},\mathcal{S}\right) \mathrm{d}r. \tag{6}$$

It can be shown that $\Lambda(t)$ is a *product* of generalized step-functions. These steps are probabilistic representations of potential occluders, lying between $\mathbf{p}(t)$ and the optical centre.

Substituting (4) and (5) into (3) gives the final **visibility density**, corresponding to the **green curves** in fig. 1, along ray $\mathcal{R}$:

$$\mathrm{pr}\left(\multimap t \,|\, \mathcal{R},\mathcal{S}\right) = \frac{\exp\big(-\eta\Lambda(t)\big)}{|\mathcal{R}\cap\mathcal{S}|}\sum_k^N w_k\, G\big((t-\mu_k)^2/\sigma_k^2\big). \tag{7}$$

The scalar $|\mathcal{R}\cap\mathcal{S}|$ is a normalizing constant, representing the 'total intersection' of the ray with the probabilistic scene-model, which can be computed numerically.

The model was evaluated by computing reference visibilities in high-resolution point-clouds [2], then decimating these clouds, and re-estimating the visibility of a large number of test-points. An ROC analysis was performed on the estimated vs. true states (visible/occluded) of the test points. The results indicate that the new model outperforms naive visibility tests, in addition to the theoretical contributions outlined here.

[1] K. Lai, L. Bo, and D. Fox. Unsupervised feature learning for 3D scene labelling. In *Proc. ICRA*, 2014.

[2] G. Schaufler and H.W. Jensen. Ray-tracing point-sampled geometry. In *Proc. Eurographics Workshop on Rendering*, 2000.

[3] A. Yezzi and S. Soatto. Stereoscopic segmentation. *IJCV*, 53(1), 2003.