Ambient Occlusion via Compressive Visibility Estimation

Wei Yang1, Yu Ji1, Haiting Lin1, Yang Yang1, Sing Bing Kang2, Jingyi Yu1
1University of Delaware, Newark, DE, USA. 2Microsoft Research, Redmond, WA, USA.

The problem of recovering intrinsic properties of a scene/object from images has attracted much attention in the past decade. Tremendous efforts have been focused on intrinsic properties related to shading and reflectance [1, 2]. In this paper, we explore a challenging type of intrinsic properties called the ambient occlusion map. Ambient Occlusion (AO) characterizes the visibility of a surface point due to local geometry occlusions. Given a scene point x, its AO measures the occlusion of ambient light caused by local surface geometry:

\[ A(x) = \frac{1}{4\pi} \int_{\Omega} \nu(x, \hat{w}) (\hat{w} \cdot \hat{n}) d\hat{w} \]  
(1)

where \( \hat{w} \) is the direction of incident light; \( \hat{n} \) is the normal of \( x \); and \( (\cdot) \) refers to the dot product. \( \nu(x, \hat{w}) \) is the local visibility function and is equal to 0 if the light ray from \( \hat{w} \) is occluded from \( x \).

Intuitively, we can illuminate the object using a dense set of uniform directional lights \( \hat{w}_i \) and sum up images captured from all directions.

\[ \sum_{i=1}^{N} I_i = \rho \sum_{i=1}^{N} \nu_i(\hat{w}_i, \hat{n}) = \rho \hat{A}c \]  
(2)

AO term \( \hat{A} \) cannot be resolved since the albedo \( \rho \) is also unknown. Hauagge et.al [3] assume the visibility function follows cone-shaped distribution centered at the normal as \( \alpha = \pi \sin^2 \alpha \), \( \alpha \) is the cone’s half angle. Under uniformly distributed lighting, they show that computing \( \kappa = E[I]^2/E[I^2] \) (\( E[\cdot] \) stands for expectation) directly cancels the albedo. For their assumption to work, densely distributed light sources will be needed.

Instead of capturing one lighting direction at a time, we aim to enable multiple lighting directions in one shot. A downside though is that we cannot use the \( \kappa \) statistics to cancel out the albedo. Instead, we build our solution on compressive signal reconstruction. We use a binary vector \( b = [1_1...1_N] \) to represent the status of \( N \) lighting directions, where \( I_i \) = 1 or 0 corresponds to if the lighting direction \( \hat{w}_i \) is enabled or disabled. We have:

\[ I = \rho \sum_{i=1}^{N} l_i \nu_i(\hat{w}_i, \hat{n}) \]  
(3)

We can now use a set of \( M \) strategically coded directional lighting patterns. For each pattern \( b^j, j = 1...M \), we capture an image \( I^j \). This results in an \( M \times N \) measurement matrix \( B = [b^1, b^2...b^M]^T \). Rewrite Eqn. 3 as

\[ I = \rho B[W \ast \hat{n}] \]  
(4)

where \( W(\hat{n}) = [(\nu_1 \cdot \hat{n}), (\nu_2 \cdot \hat{n}), ..., (\nu_N \cdot \hat{n})] \) and \( [\ast] \) refers to the pairwise element-wise product. Given the measurements, we aim to solve for \( \rho, V \) and \( \hat{n} \). Our solution is to reduce the problem to two sub-problems and solve them using iterative optimization.

Visibility Recovery Sub-problem. \( V \) is a binary pattern and solving \( V \) is NP-hard. We reduce this problem to an \( \ell_1 \) regularized \( \ell_1 \) minimization:

\[ \hat{\rho}, \hat{V} \leftarrow \arg\min_{\rho, V} \left\{ \| \rho B(W_0 \ast V) - I\|^2_2 + \lambda_1 \| V \|_1 \right\} \]  
(5)

where \( \lambda_1, \lambda_2 \) and \( \lambda_3 \) are weighting factors. The new objective function consists of four terms: 1) \( \| \rho B(W_0 \ast V) - I\|^2_2 \) corresponds to the fidelity term where the estimated \( V \) should be consistent with the observed pixel intensities \( I \); 2) \( \| V \|_1 \) is the sparse prior term that forces the visibility of negligible light directions should be zero. With this term, the solution would favor a sparse set of visible light directions; 3) \( \| V - 0.5 \|_\infty \) is the binary prior term. It is used to clamp the elements of \( V \) with high values to 1 and lows values to 0. Combining \( \| V \|_1 \) and \( \| V - 0.5 \|_\infty \) with weighting factors allows us to obtain an approximate binary solution; and 4) \( \| V \|_1 \) is the total variation term, i.e., to bias towards a solution with compact visible areas.

Normal Recovery Sub-problem We then threshold the \( \hat{V} \) to get a binary visibility vector \( V \). Now that we have both the visibility vector \( V \) and albedo, we can refine the estimation of normal \( \hat{n} \) by solving for the following least square problem:

\[ \begin{align*}
\hat{\rho} \cdot \hat{n} & \leftarrow \arg\min_{\rho, \hat{n}} \| \hat{\rho} B[W(\hat{n}) \ast \hat{V}] - I\|_2 \\
\text{Subject to} & \| \hat{n} \|_2 = 1
\end{align*} \]  
(6)

Specifically, we relax the constraint to \( \| \hat{n} \|_2 \leq 1 \) and solve it via constrained least square minimization. Next, we use the result \( \hat{n} \) to update \( W \). We repeat the process to iteratively improve the visibility and normal estimation.

We construct an encodable directional light source using the light field probe [4] and validate our approach. Experiments show that our scheme produces AO estimation at comparable accuracy to [3] but with a much smaller set of images. In addition, we can recover more general visibility functions beyond the normal-centered cone-shaped models.