Rebuttal: A Lighting-Invariant Point Processor for Shading

We really appreciate the reviewers' positive comments and their thoughtful suggestions for improvement. Our theory suggests many different directions to pursue, and we hope to inspire others to help us explore them.

Clarity of Exposition

We worked hard to make the submission understandable to a vision audience, and the reviewers' comments will help make it even more accessible, both to experts and non-experts alike. We will move the proof of Theorem 1 to a supplemental and use the vacated space to:

- Introduce our work with a "big picture" statement about how the analysis is organized
- Insert brief, intuitive descriptions of the core mathematical concepts, including algebraic varieties, automorphisms, and the implicit function theorem (For example, an algebraic variety can be thought of as the zero-set of a system of polynomial equations.)
- Clarify the derivation of the polynomials in Eqns 3-5
- Provide additional details about the data that is used to train the coupled pair of neural networks

Individual Comments

R3: *How do you arrive at the polynomials in Eqns 3-5?*
The most important step is introducing the dummy variable $w = 1/||\mathbf{N}||$ that allows conversion of the Lambertian shading equation (1) to a pair of polynomials s, r_0. Then it's just standard differentiation and manipulation of those equations. We will clarify this in the final paper.

R2: *Is the network trained and tested on single direction light sources L^*?*
Our training set $\{\mathbf{I}, \mathbf{f}\}$ uses measurements $\mathbf{I}$ that are each generated using one single-direction light vector L^* that is uniformly chosen at random from the upper unit hemisphere S^2_+. Note that at runtime the approach only requires the light to be single-directional in a small receptive field U, and that the direction is allowed to change from one receptive field to the next.

It's also worth mentioning that another interesting future direction would be to generalize the lighting model in equation (1) to include first or second-order spherical harmonic lighting (an increase from our 3 lighting parameters to 4 or 9, respectively) and then to repeat the approach of eliminating lighting parameters and analyzing the varieties that emerge.

R3: *The 2-jet representation will presumably be highly sensitive to noise.*
The second derivatives' sensitivity to noise is indeed an important consideration, and we accounted for this when generating Figs. 6 and 7. The results in these figures were generated by rendering discrete images, adding per-pixel gaussian noise, and then robustly estimating the first and second derivatives using gaussian derivative filters. We have found that the results are unchanged for noise levels with standard deviation up to 1% of maximum image intensity. We will add these details to the final paper.

R2: *Why is the surface patch solely quadratic? What if the input data has higher-order surface derivatives?*
At any smooth surface point, regardless of shape, the point processor can be expected to work well only when the receptive field size is properly chosen, meaning that a second-order Taylor approximation to the surface is accurate within the receptive field.

Of course, in a realistic setting, one does not know *a priori* which receptive field size is "right" for each image point. One way to deal with this might be to apply the point processor to many different receptive field sizes at each point (e.g., using different gaussian derivative filters) and then to reason about which sizes to use based on global consistency. This strategy has been successful, for example, in a prior patch-based method using quadratic patches [17], and we are hopeful that it will also work with our point processor.

A different future direction would be to derive an analogous statement to Theorem 1 based on third-order or higher derivatives. One will obtain a statement analogous to Theorem 1, but with higher order polynomials that include additional measurement and shape variables $(I_{xxx}, I_{xxy}, \ldots, f_{xxx}, f_{xxy}, \ldots)$.

R3: *[I would like to see] a fully worked out, robust algorithm, and evaluation on real data.*
We would certainly like that too, and we are grateful that the reviewers were willing to give a theory paper a chance without a second systems paper to go with it. One of the reasons we want to publish this paper is that the theory suggests too many different algorithms for us to explore on our own.

As mentioned above, the multi-scale strategy of [17] is one possibility. Reviewer 2 proposes another based on discretizing the orientation space and using a discrete MRF. Yet another would be to exploit the fact that the learned maps $\hat{\phi}_{\mathbf{I}}$ are differentiable, so objectives for spatial consistency like equation (13) can leverage continuous optimization methods such as gradient descent. Another still would be to consider spatial regularization schemes that aim at producing a curvature field for a surface instead of an orientation field; perhaps this will provide a computational tool that helps resolve the long-standing speculation that humans directly infer curvature from shading instead of orientation (e.g., Perception 23:169-189, 1994.)