Non-appearance-preserving Image Smoothing

We would like to thank all reviewers for their constructive comments and all suggested discussions will be revised.

Major Response

Motivation Many image smoothing applications decompose images into independent editable maps, and achieving satisfactory decomposition is a fundamental goal of these applications. Nevertheless, the full-image appearance preservation is resistant to this goal. Our motivation is to explicitly identify the undesired pattern pixels (*e.g.*, texture, shadow, specularity, *etc.*) and discard the appearance preservation on those pixels. This enables more thorough smoothing/decomposition and can benefit related applications.

Modeling The undesired pattern pixels are identified by a knapsack model based on two observations: (1) The image smoothing energy (*e.g.*, texture removal energy) tends to penalize undesired patterns (*e.g.*, texture). During image smoothing, the more a pixel is penalized, the more its color changes. Therefore, the input-output color change is our knapsack value (Eq.(8)) because it reflects how each pixel is penalized and undesired. (2) After smoothing (*e.g.*, texture removal), the undesired (*e.g.*, texture) pixels tend to have less salient contours than the desired (*e.g.*, structure) pixels. These salient contours are measured as our knapsack weight (Eq.(9)) to represent how each pixel is preserved and desired. These two observations are verified by statistics (Supp. §1). Finally, the knapsack identifies as much undesired patterns (knapsack value) as possible while preserving a limited amount of desired ones (knapsack weight).

Pattern Identification As in the first example of Fig. 2, the wood textures are undesired patterns. After knapsack solving, all wood textures are identified with green marks successfully. Solid evidences (line 357-368 and Supp. Fig. 44-49) also validate such pattern identification capability.

Detailed Response

Reviewer 1:

Thank for your constructive comments. We will also revise and add examples of the bilateral texture filter.

Q: Why choose points instead of other structures?
A: Point is flexible to use and can represent many structures. Our modeling is not limited to using sparse control points, and it can also use other structures like super-pixels.

Reviewer 2:

Firstly, we would like to amicably convey that the weight in Eq.(9) is computed on smoothed matrix $\boldsymbol{Y}$ instead of the input images, and the input noise may not interfere the weight.

Q: The usage of sparse control points has been studied in the previous literature.
A: Thank you very much for pointing out the related literature. We have read the JMIV and EMMCVPR papers thoroughly. We will narrow down the claims to avoid attributing the novelty of NIS to the usage of sparse control points. Our fundamental novelty is to identify undesired patterns so as to discard their appearance preservation. Although the mentioned papers also use points to process images (*e.g.*, reconstruct image from a set of color points for image compression), it remains a novel direction to identify patterns like specularity and texture using sparse control points. The effectiveness of NIS comes mainly from achieving such identification, and this is also a contribution to point-based imaging literature.

Q: The choice of the fidelity term needs discussion.
A: Thank you for pointing out with helpful references. We will revise and attach importance to the fidelity choices.

Reviewer 3:

Q: NIS introduces more hyper-parameters.
A: Although NIS introduces the extra parameters, the upcoming benefits significantly outweigh the cost in parameter tuning. This is because: (1) the parameter adjustment is not difficult as we do not need to determine the amount of salient constituents accurately. The knapsack solver is self-adaptive and can solve the accurate quantity of undesired pixels from coarse parameters as discussed in line 368-373. (2) NIS is robust enough to handle diverse inputs even with default parameters, *i.e.*, in-the-wild examples in Supp. Fig. 5-17 are with the same default settings.

Q: NIS should be compared to iterative smoothing.
A: We will include comparisons to iterative smoothing for better exposition. Meanwhile, iterative image smoothing may not achieve satisfactory results as ours, supported by two evidences: (1) Iterative smoothing with existing methods often causes desaturated/low-contrast artifacts (visually similar to the third column in Supp. Fig. 18), whereas NIS does not suffer from this symptom. (2) Repeating the image smoothing process may not address the challenging open problems like specularity removal, whereas NIS shows promising results (Supp. Fig. 19-23,29-39). Furthermore, we attribute the effectiveness of NIS to the pattern identification capability within our knapsack modeling.

Q: In §1.2, the ground truth structure is unknown so that we cannot judge whether it is perfectly reconstructed.
A: The derivation of §1.2 actually comes from real-life applications. Many computer graphic applications assume that objects have measurable reflectance/albedo/diffuse intensity, and in these cases, it is appropriate to introduce a target ground truth for image smoothing problems. We will revise the exposition by adding such assumption explicitly. Experiments validate that this assumption is applicable to a variety of applications.