Coherent Intrinsic Images from Photo Collections
Supplemental document

Pierre-Yves Laffont[1] Adrien Bousseau[1] Sylvain Paris[2] Frédo Durand[3] George Drettakis[1]

[1] REVES / INRIA Sophia Antipolis [2] Adobe Systems [3] MIT CSAIL

Description of this document

Selection of constrained pairs. Fig. 1 illustrates the steps of our sampling algorithm for selecting candidate pairs (Sec. 4.2).

Ambient occlusion. Fig. 2 shows an example of reconstructed geometry proxy, estimated ambient occlusion, and the effect of the correction described in Sec. 4.1 on the intrinsic decomposition.

Image-guided decomposition. Fig. 3 shows the influence of our pairwise reflectance constraints (Sec. 5.1) on a Flickr image. Fig. 4 compares the two image-based smoothing priors described in Sec. 5.2. Fig. 5 illustrates the influence of grayscale regularization (Sec. 5.2).

Results. Fig. 6 illustrates the influence of reconstructed point cloud density on our decomposition, on the synthetic dataset. Fig. 7 shows a comparison between our decomposition on the Doll scene, and the results of previous approaches. Fig. 8 compares our results with those of a user-assisted method [Bousseau et al. 2009]. Fig. 9 shows a comparison between our decomposition and that obtained with a single-image method. It illustrates the limitations of a common assumption for such techniques, which enforces pixels with similar chrominance to share similar reflectance.

Accompanying files

In the accompanying video, we show image-based view transitions [Roberts 2009] between photographs with harmonized lighting, as described in Sec. 6.3, as well as artificial timelapse sequences synthesized by transferring all illumination conditions on a single viewpoint. We also provide HTML files which list our intrinsic decompositions on 9 datasets; input, ground truth and results of all compared methods on our synthetic benchmark; and evaluation of our results when varying the density of the reconstruction on the synthetic dataset. Lastly, we provide the Matlab sampling code which was used to generate Fig. 1.

Acknowledgments

We thank Don Chesnut (Fig. 2) and the following Flickr users for permission to use their pictures in this supplemental document: *ChihPing (Fig. 8)*, Fulvia Giannessi (Figs. 3, 8 and 9).

References

BOUSSEAU, A., PARIS, S., AND DURAND, F. 2009. User-assisted intrinsic images. *ACM Trans. Graph. 28*, 5.

LEVIN, A., LISCHINSKI, D., AND WEISS, Y. 2008. A closed-form solution to natural image matting. *IEEE Trans. PAMI*.

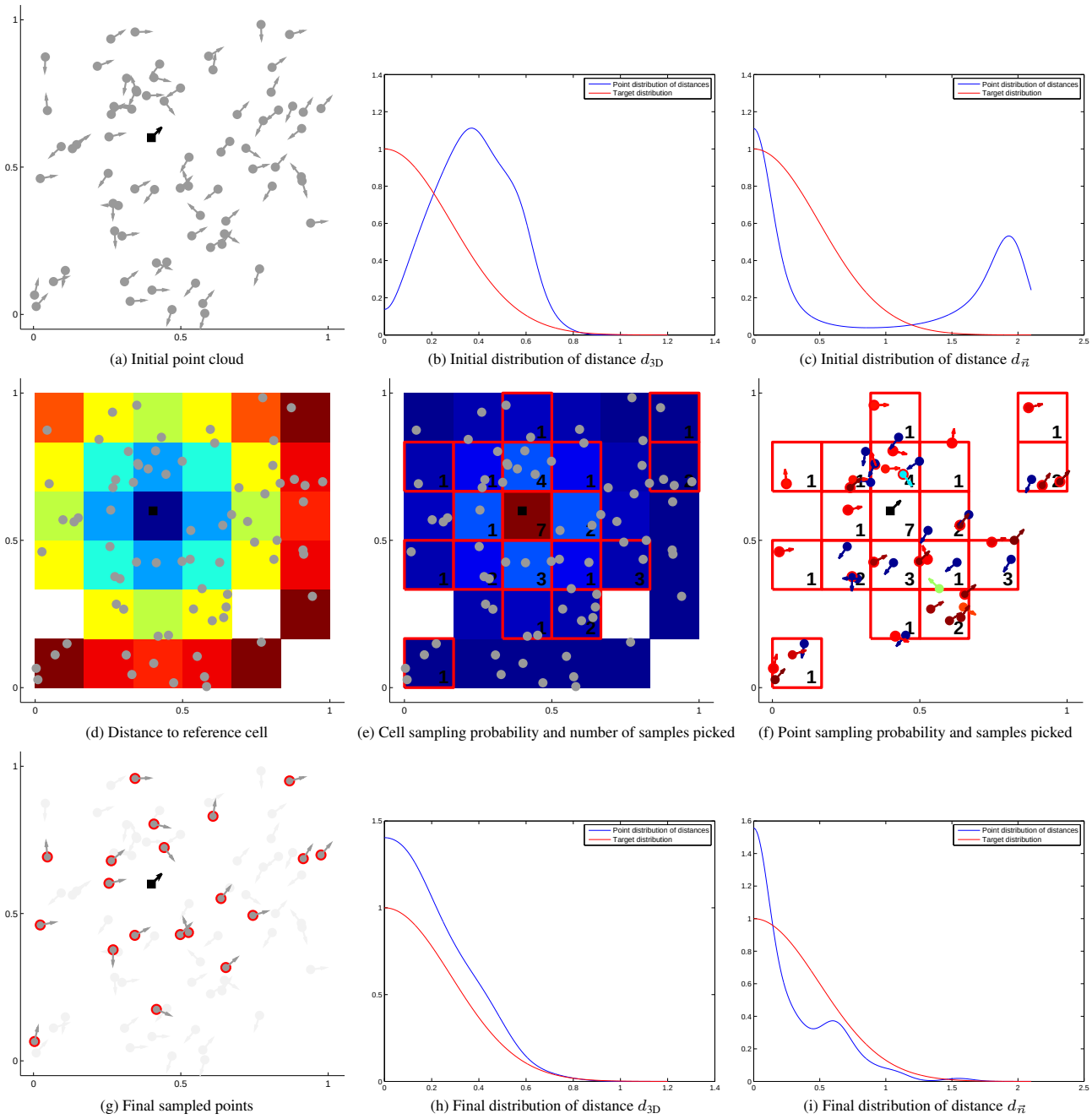
ROBERTS, D. A., 2009. Pixelstruct, an opensource tool for visualizing 3d scenes reconstructed from photographs.

SHEN, L., TAN, P., AND LIN, S. 2008. Intrinsic image decomposition with non-local texture cues. In *Proc. IEEE CVPR*.

TAPPEN, M. F., FREEMAN, W. T., AND ADELSON, E. H. 2005. Recovering intrinsic images from a single image. *IEEE Trans. PAMI 27*, 9.

WEISS, Y. 2001. Deriving intrinsic images from image sequences. In *IEEE ICCV*, vol. 2, 68.

ZHAO, Q., TAN, P., DAI, Q., SHEN, L., WU, E., AND LIN, S. 2012. A closed-form solution to retinex with nonlocal texture constraints. *IEEE Trans. PAMI 34*.

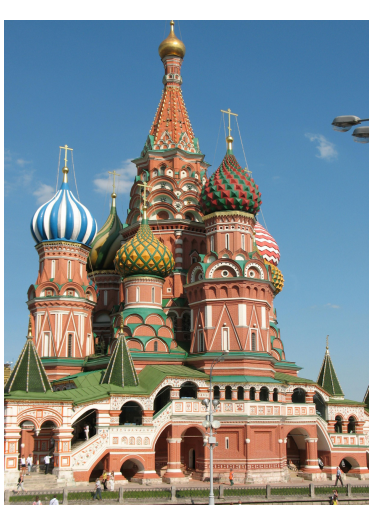


(a) Initial point cloud (b) Initial distribution of distance d_{3D} (c) Initial distribution of distance $d_{\vec{n}}$

(d) Distance to reference cell (e) Cell sampling probability and number of samples picked (f) Point sampling probability and samples picked

(g) Final sampled points (h) Final distribution of distance d_{3D} (i) Final distribution of distance $d_{\vec{n}}$

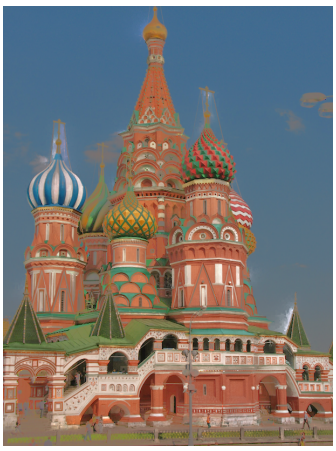
Figure 1: *2D illustration of our algorithm for sampling candidate pairs for a single point. (a) Given an oriented point cloud, we wish to select N points so that their distances d_{3D} and $d_{\vec{n}}$ to a reference point (black square) follow normal distributions. (b-c) The point cloud is irregularly sampled, and the distribution of distances of all points (blue curves) is very different from the target normal distributions (red curves). (d) We first embed the point cloud in a grid and compute the Euclidean distance d_{3D} to the cell containing the reference point: the distance is color-coded (blue: small distance; dark red: large distance). (e) We infer a sampling probability for each cell based on d_{3D} as described in Algorithm 1; this sampling probability is shown color-coded for each cell (blue: low sampling probability; dark red: large sampling probability). From these probabilities, we draw N samples to choose the number of points to select in each cell, shown as black numbers in the corresponding highlighted cells. We discard all points contained in cells for which no sample has been drawn. (f) For all the points within sampled cells, we infer a sampling probability based on $d_{\vec{n}}$ (shown color-coded; blue: low sampling probability; dark red: large sampling probability). We draw samples in each cell from these probabilities; the number of samples drawn in each cell corresponds to the result of (e). (g) The final samples are distributed so that many points are nearby and have similar normals compared to the reference point, while a few are further away or with different normals to produce a well-connected graph of constraints. (h-i) The distribution of distances of sampled points (blue curves) is closer to the desired normal distributions (red curves). We provide the Matlab sampling code used to generate this figure with the following parameters: 150 points in the point cloud and 35 samples drawn, for the illustrations (a, d-g); 500000 points in the point cloud, and 100 samples drawn, for the distributions estimated with the Matlab* `ksdensity` *function (b-c, h-i).*



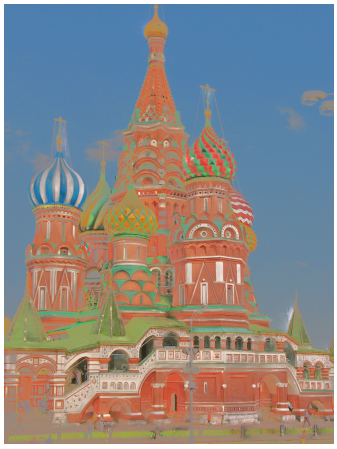
(a) View of St. Basil

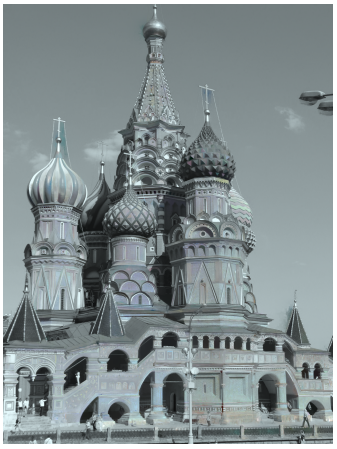
(b) Approximate proxy

(c) Ambient occlusion at 3D points



(d) Decomposition without correction





(e) Decomposition with correction

Figure 2: *Ambient occlusion estimation on the St. Basil scene, downloaded from Flickr. An approximate proxy created with Poisson reconstruction (b) is used to estimate ambient occlusion at sparse 3D points (c). Correcting pairwise constraints with the ratios of ambient occlusion yields a better decomposition (e) in regions systematically in shadow, such as the arches near the ground.*

Input | Without pairwise constraints | With pairwise constraints

Figure 3: *Influence of the pairwise relative constraints on an image of the Moldovita scene. Despite the intricate texture patterns on the painted façade, these constraints enable the separation of reflectance from the illumination.*

Input image

With prior of [Bousseau et al. 2009]

With our smoothness prior based on [Levin et al. 2008]

Figure 4: *Influence of the smoothing prior on an image of the Doll scene. We compare results of our decomposition using the image-guided prior of Bousseau et al. (middle), with ours based on the Matting prior of Levin et al. (right). Our prior better disambiguates texture from shading in complex regions and recovers a smoother illumination layer.*

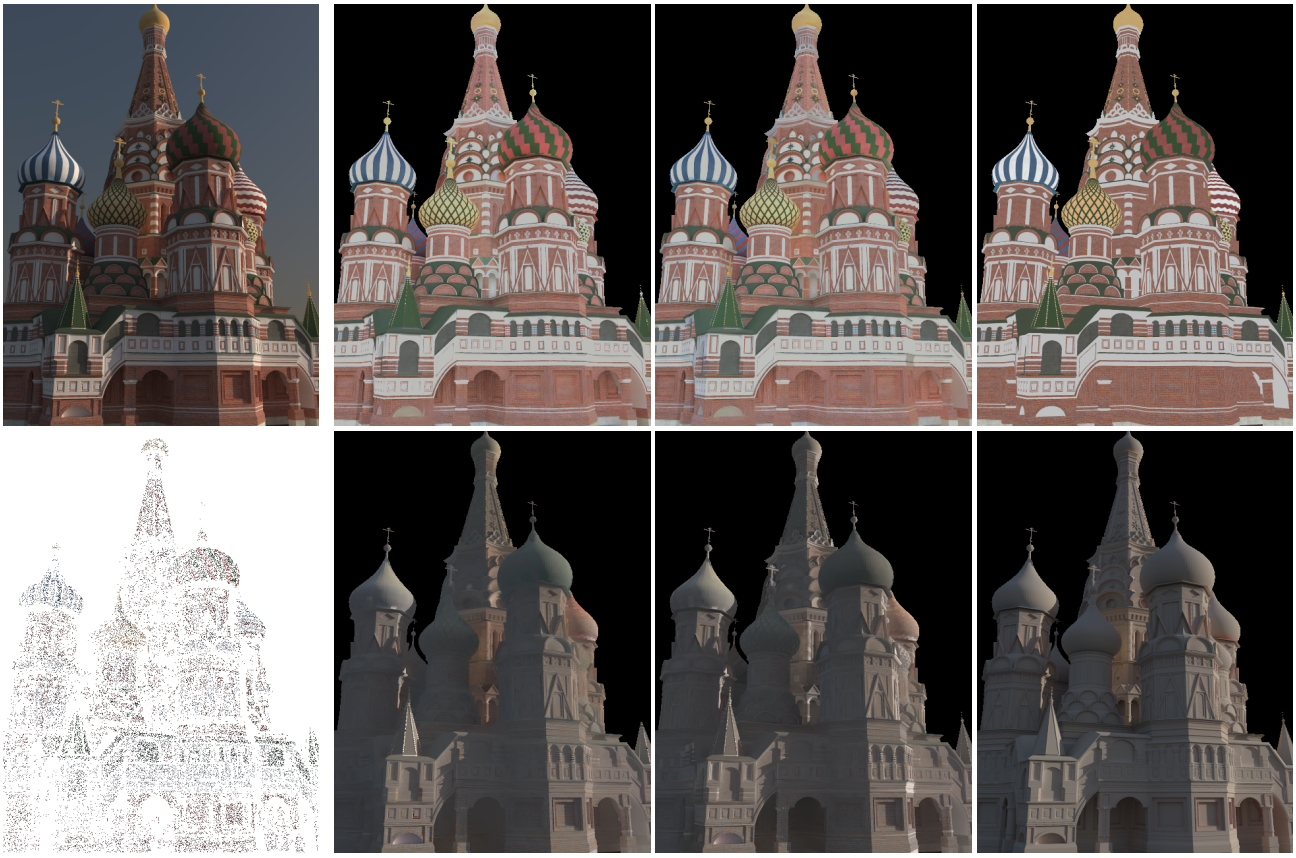


Figure 5: *Influence of grayscale regularization on an image of our synthetic dataset. While our method produces a high quality decomposition in most regions of the image, adding the grayscale regularization further improves the results in regions with ambient occlusion. The regularization helps to capture the shadowing effects in areas where only few 3D points are reconstructed.*

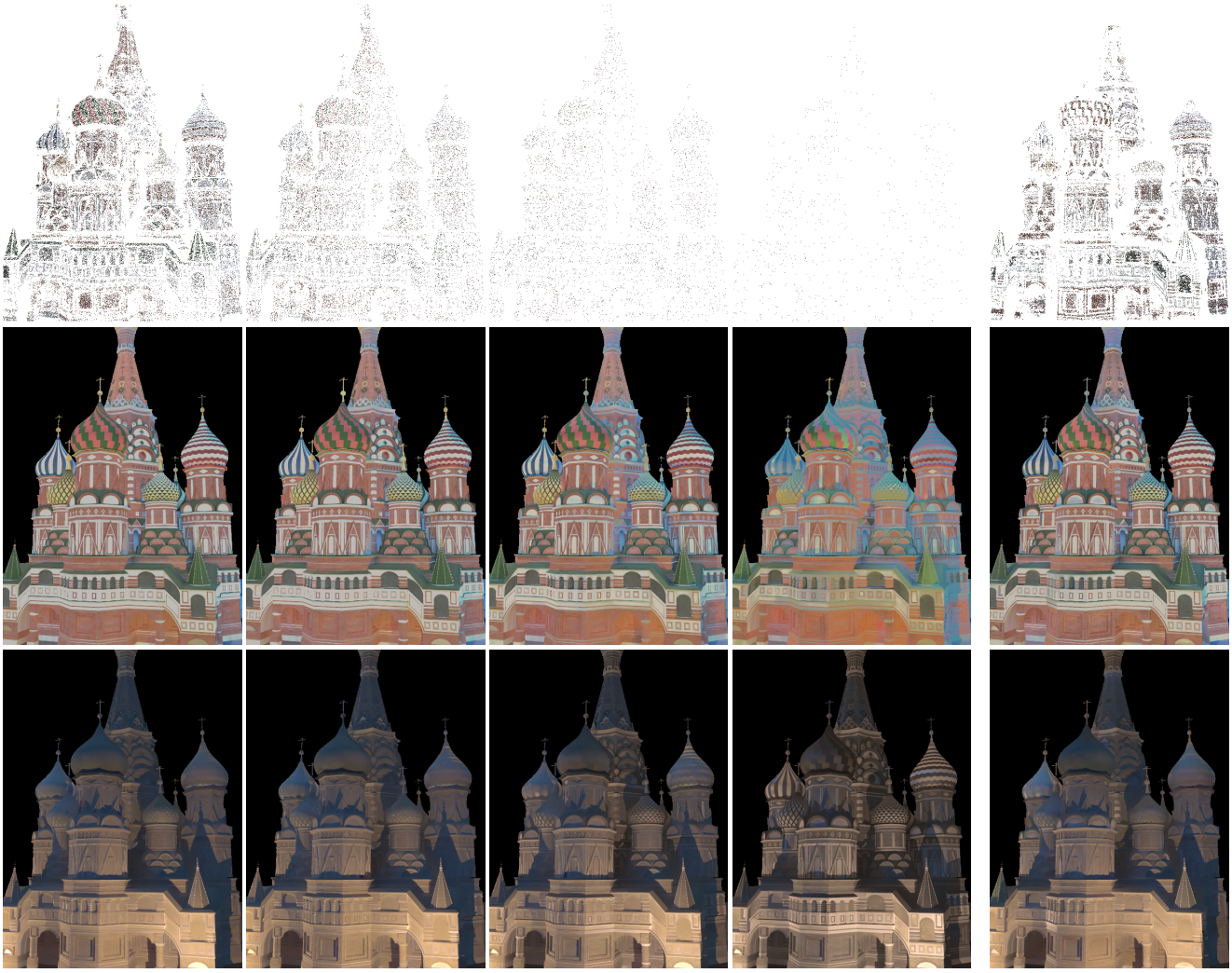
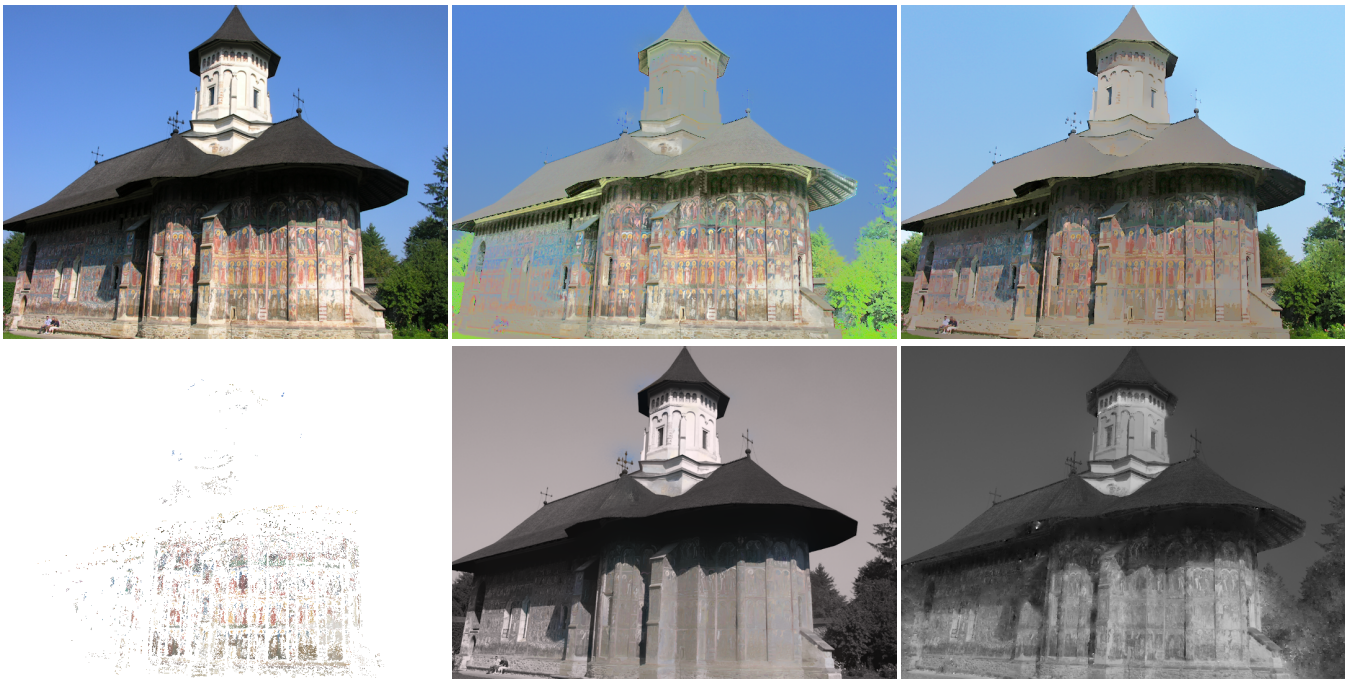


Figure 6: *Influence of the point cloud density and reconstruction method on an image of our synthetic dataset. Top row: constrained 3D points and their estimated reflectance. Middle row: estimated reflectance. Bottom row: estimated illumination. The right column corresponds to the PMVS reconstruction with ground truth camera parameters (instead of the output of structure from motion, which fails on synthetic images); note the irregular distribution of reconstructed points. For each setting, we report the LMS error on this view.*

Figure 7: *Comparison between our approach and existing single-image methods on a picture captured with a flash. We captured our own version of a similar doll from different viewpoints with a moving light source (flash) and compare with the results shown in previous papers. Although our input is more challenging due to the background texture and shadows cast on the doll, our automatic method successfully recovers a smooth illumination layer and a shading-free reflectance layer.*

Figure 8: *Comparison to the user-assisted approach of Bousseau et al. Our coherence constraints ensure that the reflectance is similar in every view and allows the recovery of reflectance values even in shadowed areas where the single image approach of produces noisy results. In addition, we recover a smoother illumination in textured planar regions.*



(a) Input image and constrained points

(b) Our decomposition

(c) Result from [Zhao et al. 2012]

Figure 9: *Comparison to a single-image method on the Moldovita scene. Our approach successfully separates the complex painted texture from the smooth illumination (b), in regions which are well reconstructed (a). In the absence of 3D points (e.g., steeple and rool, left part of the façade), our decomposition relies on the image-guided smoothness prior. In comparison, the method by Zhao et al. shares similar artifacts on the steeple and roof due to their assumption on chrominance, but does not extract the shadow cast on the façade (c).*