

Seamless Image-Based Texture Atlases using Multi-band Blending

Cédric Allène, Jean-Philippe Pons and Renaud Keriven
CERTIS, École des ponts, Paris-Est, France
{allene,pons,keriven}@certis.enpc.fr

Abstract

In this paper, we propose a novel method for creating a high-quality texture atlas from a 3D model and a set of calibrated images. Our method focuses on avoiding visual artifacts such as color discontinuities, ghosting or blurring, which typically arise from photometric and geometric inaccuracies. We first compute a partition of mesh faces which realizes a good trade-off between visual detail and color continuity at patch boundaries: we efficiently obtain a close-to-optimal seam placement using graph cuts optimization. We then apply a pixel-wise color correction in the vicinity of patch boundaries with a principled 3D extension of multi-band image blending: we achieve faultless color continuity while avoiding ghosting artifacts. We demonstrate the effectiveness of our method on two real-world large-scale scenes.

1. Introduction

Capturing, processing and displaying the visual attributes, such as color, of a 3D model typically requires to map its surface to a two-dimensional domain. Generally, it is impossible to find a global such parameterization with acceptable distortion, so an *atlas* structure is adopted: it consists of a partition of the surface into connected parts (*patches*) and of a piecewise 2D parameterization (*cf* for example [16] and references therein).

The instantiation of this problem in the context of image-based 3D modeling, *i.e.* when extracting geometry and/or visual attributes from digital photographs, has received much attention in the computer vision and computer graphics communities.

On the one hand, the image-based case alleviates the parameterization problem: the projective transformations from the surface to the input images constitute natural and optimal mappings [11, 17, 19]. They avoid image resampling and loss of visual detail, contrarily to approaches based on other parameterizations [1, 2, 7].

On the other hand, in the image-based modeling con-

text, color discontinuities at patch boundaries (*seams*) are a crucial issue, due to photometric and geometric inaccuracies: varying lighting conditions and camera response, non-Lambertian reflectance, imperfect camera calibration, approximate shape, etc. Weighted averaging of images in overlapping regions [2, 6, 12, 14, 17] is not sufficient. It causes unsatisfactory ghosting and blurring, unless the 3D model is highly accurate (*e.g.* obtained by laser range scanning) and camera calibration is tightened using image-based registration [2, 8, 12].

Two main axes of research have been explored in order to reduce seam visibility. The first approach is the optimization of patch layout. Some works force patch boundaries into regions of high negative curvature [13, 18]. Others use an image fidelity term [11, 20] to explicitly look for a partition of the surface inducing minimal color discontinuities. Of particular interest is the formulation of this problem as a *Markov random field* optimization [11], for it brings powerful algorithmic tools into play. However, these works suffer from the absence of per-pixel processing: they are unable to achieve perfect color continuity.

The second improvement path is precisely pixel-wise color correction in the vicinity of patch boundaries. A notable work in this category is a tentative extension of 2D multi-band image splining [5] to textured 3D surfaces [1]. However, this work misses the importance of an optimal patch layout, and fails to define transition zones of distinct and adapted width for the different frequency bands. As a result, it has to keep to a two-band frequency decomposition to limit ghosting artifacts.

In this paper, we propose a novel method for creating a high-quality seamless texture atlas from a 3D model and a set of calibrated images. Our method upgrades the Markov random field approach of [11] with a principled 3D extension of multi-band image blending, thereby achieving both close-to-optimal seam placement and faultless color continuity. We demonstrate the effectiveness of our method on two real-world large-scale scenes reconstructed from high-resolution images using recent

multi-view stereovision techniques.

2. Methods

In the following, we note I_1, \dots, I_n the input calibrated images, and Π_i the projection from 3D space to image I_i . We assume that the surface is represented by a polygonal mesh \mathcal{M} with faces $\mathcal{F} = \{f_1, \dots, f_m\}$.

2.1. Patchwork optimization using graph cuts

The first step of our method mainly follows [11], with a few clarifications and improvements: it computes a partition of the surface which realizes a good trade-off between visual detail and color continuity at patch boundaries, using graph cuts optimization.

In practice, it consists in assigning each face of the mesh to one of the input views in which it is visible. After discarding faces that are not visible in any view, this can be encoded by a labeling vector $\vec{\ell} = \{\ell_1, \dots, \ell_m\} \in \{1, \dots, n\}^m$. We denote by \mathcal{L} the set of *admissible* labeling vectors, *i.e.* that fulfill the aforementioned visibility constraints.

The optimality of a labeling is quantified by an energy function. It is a weighted sum of two terms. The first term measures *visual detail*. Rather than the heuristic combination of image resolution, viewing distance and angle between viewing direction and face normals proposed in [11], we adopt a measure both simpler to compute and easier to interpret: the total number of *texels* (texture elements) on the mesh. Given an admissible labeling vector $\vec{\ell} \in \mathcal{L}$, this writes as a sum over faces:

$$E_{detail}(\vec{\ell}) = - \sum_{j=1}^m \text{area} [\Pi_{\ell_j}(f_j)] \quad (1)$$

The second term measures *color continuity* at edges between faces assigned to different views. Let us note $e_{j,k}$ a non-border edge of the mesh, adjacent to faces f_j and f_k . If these two faces are assigned to different images, *i.e.* $\ell_j \neq \ell_k$, then color is very likely to be discontinuous across the edge, and $e_{j,k}$ is a *seam edge*. In order to minimize seam visibility, the second energy term is defined as the integral along the seams of color discrepancy between bordering images. This can be written as a sum over the set \mathcal{E} of all non-border edges:

$$E_{seams}(\vec{\ell}) = \sum_{e_{j,k} \in \mathcal{E}} g_{e_{j,k}}(\ell_j, \ell_k) \quad (2)$$

$$g_e(\ell, \ell') = \int_e \|I_\ell(\Pi_\ell(x)) - I_{\ell'}(\Pi_{\ell'}(x))\| dx \quad (3)$$

It must be noted that g_e respects the following regularity condition $\forall \ell, \ell', \ell'' \in \{1, \dots, n\}$:

$$g_e(\ell, \ell) + g_e(\ell', \ell'') \leq g_e(\ell, \ell'') + g_e(\ell', \ell) \quad (4)$$

This still holds with any metric on colors instead of the usual Euclidean distance in RGB color space.

This regularity property is not underlined in [11], whereas it has a considerable practical consequence: it allows to minimize the energy functional with α -*expansion* [4, 9]. It consists in translating the labeling problem, which is generally NP-hard, to a succession of binary minimum cut problems. Efficient solutions to these min-cut problems are described in [3]. The whole process monotonically decreases the energy and is guaranteed to converge to a *strong local minimum*, thereby ensuring a close-to-optimal seam placement.

For sake of completeness, let us mention that minimizing E_{detail} only, or equivalently, setting the weighting factor of E_{seams} to zero, results to independently map each face to the highest quality image. This naive approach is experimentally evaluated in Section 3.

2.2. Multi-band blending

The texture mosaic output by the first step of our method minimizes color discontinuities. But as later demonstrated in Section 3, it does not suppress them altogether, mainly due to differences in lighting conditions in the original views. This makes pixel-wise color correction in the vicinity of patch boundaries mandatory. In 2D, the work of Burt and Adelson [5] on multi-band blending has prove particularly effective for image mosaicing without blurring and ghosting artifacts. We extend this work to our case, *i.e.*: (i) more than two images, and (ii) image-based texture maps.

Following [5], we use the Laplacian pyramids of the input images as the multi-band decomposition, and we approximate them by differences of Gaussians. In the sequel, frequency bands are indexed by $b \in \{1, \dots, B\}$, with higher values represent lower frequencies. First, let us denote by $\mathcal{G}_b(I)$ the b^{th} level of the *Gaussian pyramid* of some image I . The first level, $\mathcal{G}_1(I)$, is a copy of I : $\mathcal{G}_1(I) = I$. Any higher level is calculated from its previous level: $\mathcal{G}_b(I)$ is obtained by convolving $\mathcal{G}_{b-1}(I)$ by some fixed Gaussian kernel G_σ : $\mathcal{G}_b(I) = \mathcal{G}_{b-1}(I) * G_\sigma$.

We can now define the b^{th} level of the *Laplacian pyramid* of image I , denoted by $\mathcal{L}_b(I)$. The last level, $\mathcal{L}_B(I)$ is a copy of the last level of the Gaussian pyramid: $\mathcal{L}_B(I) = \mathcal{G}_B(I)$. Any lower level is calculated as: $\mathcal{L}_b(I) = \mathcal{G}_b(I) - \mathcal{G}_{b-1}(I)$.

The main difficulty is now to define the color $C(x)$ of a point x of the surface as a proper combination of

the frequency bands of the different images:

$$C(x) = \sum_{b=1}^B \frac{\sum_{i=1}^n w_{i,b}(x) \mathcal{L}_b(I_i)(\Pi_i(x))}{\sum_{i=1}^n w_{i,b}(x)} \quad (5)$$

where $w_{i,b} : \mathcal{M} \rightarrow \mathbb{R}$ are some weighting functions on the surface to be determined.

These weighting functions must have several properties. First, they must be continuous and smooth over \mathcal{M} . Second, $w_{i,b}$ must have high values in the patches associated to image I_i , and must cancel in regions of the surface not visible in I_i . Fourth, the transitions of $w_{i,b}$ from zero to high values must be of proper width: narrow enough to avoid ghosting, wide enough to avoid a visual discontinuity. Moreover, this transition width must be consistent with the scale of the considered frequency band.

We compute $w_{i,b}$ as follows. First, we consider binary masks representing the useful regions in the different images: M_i is the projection in I_i of the associated 3D patches. We generate smoothed versions $M_{i,b}$ of these masks by solving a *2D heat equation* with initial conditions M_i , until some time instant which matches the b^{th} level of the Gaussian pyramid. We impose zero Dirichlet boundary conditions on exterior and occluding contours of the surface, as well as on image borders. Note that these complex boundary conditions dismiss the use of simple Gaussian convolution. Finally $w_{i,b}$ is defined by:

$$w_{i,b}(x) = \begin{cases} M_{i,b}(\Pi_i(x)) & \text{if } x \text{ is visible in } I_i, \\ 0 & \text{otherwise.} \end{cases} \quad (6)$$

The interested reader can easily check that these weight functions have all the required properties.

From a practical point of view, we discretize C in the image domain, by iterating over the pixels of each mask M_i . We define blended images I'_i that will be used instead of the original I_i to create the final texture. Then, for $x \in f_j$, $C(x)$ is stored in I'_{ℓ_j} at pixel $\Pi_{\ell_j}(x)$.

2.3. Texture atlas creation

We could render novel views from the blended images using multiple passes of projective texture mapping [6]. However, the creation of a single rectangular texture map is very desirable: it increases the rendering efficiency and allows to output portable 3D formats.

To build such a texture atlas, we first apply a morphological dilation to the masks M_i with a square structural element of a few pixels, in order to provision for automatic texture minifying during rendering. We then compute a connected component decomposition, yielding a list of texture fragments. We pack the latter using a

classical *first-fit decreasing* strategy: we place the fragments in decreasing order of size, at the first available slot found along a scanline search in the atlas. Finally, we set the texture coordinates of mesh vertices accordingly. Thus, the final output of our algorithm is compatible with standard 3D viewers.

3. Results and discussion

In this section, we demonstrate our method on two challenging datasets:

- "Aiguille du Midi": 37 images (1000×1332) of a famous French peak (Chamonix, Mont-Blanc), copyright Bernard Vallet (www.bvallet.com).
- "Ettlingen Castle": 19 images (1536×1024) of Ettlingen Castle (Germany), courtesy Christoph Strecha, EPFL (<http://cvlab.epfl.ch/~strecha/multiview/>).

To obtain accurate geometric models of these large-scale scenes automatically, we first applied a recent multi-view stereovision technique based on interest points, 3D Delaunay triangulation and global optimization with graph cuts [10]. The obtained models were later refined with a deformable mesh by gradient descent over some multi-view matching score [15].

Our code for computing the seamless texture atlases resorts to depth buffering on graphics hardware to efficiently compute the visibility of facets in the different views. Also, it uses the graph cuts minimization software by O. Veksler (<http://www.csd.uwo.ca/~olga/code.html>) and by V. Kolmogorov (<http://www.adastral.ucl.ac.uk/~vladkolm/software.html>).

Figure 1 displays a visual comparison of the views synthesized following three different approaches:

- a "naive" approach where each face is independently mapped to the highest quality image, without color correction.
- an "optimized patchwork" approach where graph cut optimization is used to minimize the length and the visibility of seams, but without color correction.
- our approach combining patchwork optimization with graph cut and color correction with multi-band blending.

Although patchwork optimization consistently improves visual quality, it cannot cope with varying lighting conditions and camera response. As a result, marked intensity discontinuities are visible on the foreground side of "Aiguille du Midi" (Figure 1, top middle), and on the ground and on the background wall of "Ettlingen Castle" (Figure 1, bottom middle). In contrast, no such artifact is apparent with our approach.

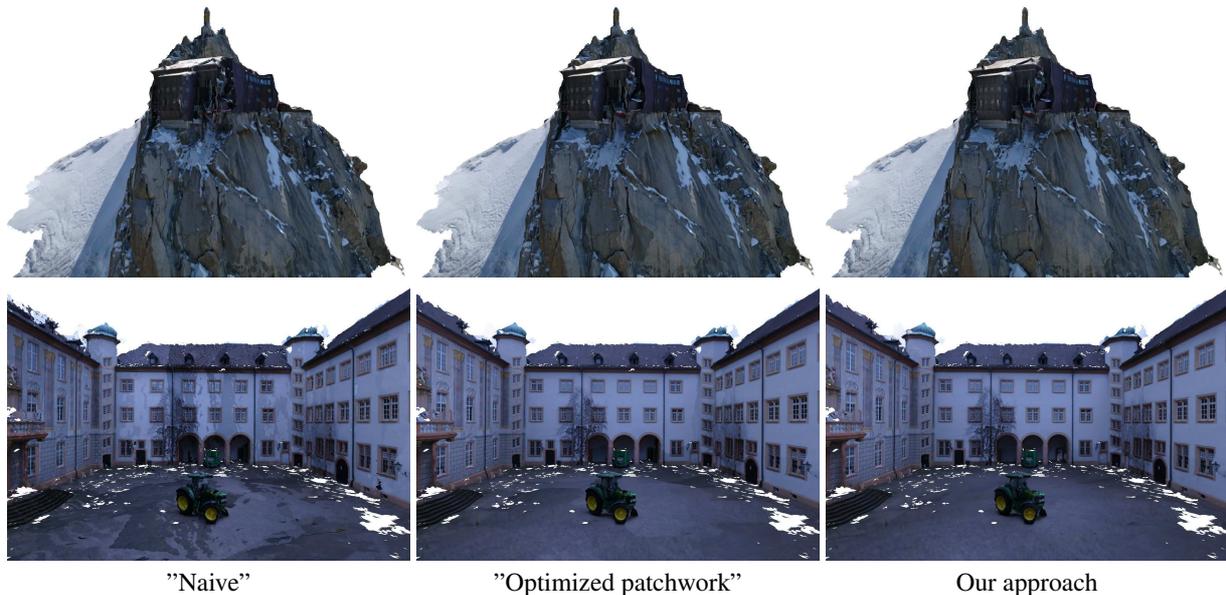


Figure 1: Comparison of three different approaches on two real-world large-scale scenes. See text for details.

4. Conclusion

We proposed a method to create seamless image-based texture atlases successfully handling real-world large-scale scene reconstructions. Full-size images, movies and interactive 3D views of these results are available online at <http://certis.enpc.fr/~allene/research-3Dblending.html>.

References

- [1] A. Baumberg. Blending images for texturing 3D models. In *BMVC*, 2002.
- [2] F. Bernardini, I. Martin, and H. Rushmeier. High-quality texture reconstruction from multiple scans. *IEEE Trans. on Visualization and Computer Graphics*, 7(4), 2001.
- [3] Y. Boykov and V. Kolmogorov. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *IEEE Trans. on PAMI*, 26(9), 2004.
- [4] Y. Boykov, O. Veksler, and R. Zabih. Fast approximate energy minimization via graph cuts. *IEEE Trans. on PAMI*, 23(11), 2001.
- [5] P. J. Burt and E. H. Adelson. A multiresolution spline with application to image mosaics. *ACM Trans. on Graphics*, 2(4), 1983.
- [6] P. Debevec, C. Taylor, and J. Malik. Modeling and rendering architecture from photographs: A hybrid geometry and image-based approach. In *SIGGRAPH*, 1996.
- [7] Z. Jankó. *Photorealistic 3D models of real-world objects*. PhD thesis, Eotvos Lorand Univ., Hungary, 2007.
- [8] Z. Jankó and D. Chetverikov. Photo-consistency based registration of an uncalibrated image pair to a 3D surface model using genetic algorithm. In *3DPVT*, 2004.
- [9] V. Kolmogorov and R. Zabih. What energy functions can be minimized via graph cuts? In *ECCV*, 2002.
- [10] P. Labatut, J.-P. Pons, and R. Keriven. Efficient multi-view reconstruction of large-scale scenes using interest points, Delaunay triangulation and graph cuts. In *ICCV*, 2007.
- [11] V. S. Lempitsky and D. V. Ivanov. Seamless mosaicing of image-based texture maps. In *CVPR*, 2007.
- [12] H. Lensch, W. Heidrich, and H.-P. Seidel. A silhouette-based algorithm for texture registration and stitching. *Graphical Models*, 63(4), 2001.
- [13] B. Lévy, S. Petitjean, S. Ray, and J. Maillot. Least squares conformal maps for automatic texture atlas generation. In *SIGGRAPH*, 2002.
- [14] F. Pighin, J. Hecker, D. Lischinski, R. Szeliski, and D. Salesin. Synthesizing realistic facial expressions from photographs. In *ICCGIT*, 1998.
- [15] J.-P. Pons, R. Keriven, and O. Faugeras. Multi-view stereo reconstruction and scene flow estimation with a global image-based matching score. *IJCV*, 72, 2007.
- [16] B. Purnomo, J.-D. Cohen, and S. Kumar. Seamless texture atlases. In *SGP*, 2004.
- [17] C. Rocchini, P. Cignoni, C. Montani, and R. Scopigno. Acquiring, stitching and blending diffuse appearance attributes on 3D models. *The Visual Computer*, 18(3), 2002.
- [18] A. Sheffer and J. Hart. Seamster: inconspicuous low-distortion texture seam layout. In *IEEE Visualization*, 2002.
- [19] L. Velho and J. Sossai. Projective texture atlas construction for 3D photography. *The Visual Computer*, 23(9), 2007.
- [20] E. Zhang, K. Mischaikow, and G. Turk. Feature-based surface parameterization and texture mapping. *ACM Trans. on Graphics*, 24(1), 2005.