A SCALE SPACE FOR TEXTURE+DEPTH IMAGES BASED ON A DISCRETE LAPLACIAN OPERATOR

Anonymous ICME submission

ABSTRACT

In this paper we design a smoothing filter for texture+depth images based on anisotropic diffusion. Our proposed filter is linear and allows to generate a *scale space* of the texture image which preserves internal image structure (i.e. it does not smooth through object boundaries). Moreover, we show experimentally that our filter demonstrates better stability under camera position changes in such a way that its response is rather intrinsic to the content and less dependent on the observer position in the scene. To illustrate the practical utility of the scale space with such properties, we integrate our filter into SIFT salient visual point detector, getting a substantial improvement of the repeatability of detected keypoints under significant viewpoint position changes.

Index Terms— image smoothing; texture+depth; anisotropic diffusion; keypoint repeatability.

1. INTRODUCTION

Scale-space theory is a fundamental tool in image processing and computer vision, as it provides a framework to problems as image and shape description, denoising, enhancement, etc. For a given input image, the scale space is generally understood as a family of its low-pass filtered versions, obtained by a filter satisfying certain axioms [1]. Among the large variety of image smoothing filters, only a limited number actually yields scale spaces. Historically, the first and most employed scale space has been that engendered by the bidimensional Gaussian filter [2, 1]. The practical effectiveness of Gaussian multi-scale representation on conventional 2D images has been widely proved in a number of different computer vision applications, including image description and retrieval. Several local image descriptors, such as SIFT [3] or BRISK [4], exploit Gaussian scale space for interest points detection. Such applications typically assume a *causality axiom*, i.e. no new features should be generated when passing from a finer to a coarser scale [5]. An example of a filter violating scale space axioms (notably the causality), but yet being used as a computationally efficient scale space approximation, is 2D rectangular filter. It is involved in *Fast-Hessian* keypoint detector, a part of SURF visual features extraction algorithm [6]. Filters with a more complex behavior, like edge preserving smoothing, may also satisfy the necessary axioms to engender scale space (e.g., the classic Perona and Malik's anisotropic diffusion [5]).

The motivation of this work comes from the availability, nowadays, of newer visual formats that go beyond the traditional 2D visual representation and enrich it with additional information, such as scene geometry. An important example is the texture+depth video format, which has gained momentum in the past few years thanks to applications as immersive communication and free viewpoint television [7]. In the texture+depth format, the conventional 2D scene intensity content (texture) is complemented with a per pixel depth map, which stores the distance of each pixel from the camera plane.

The knowledge of scene geometry affects the filter behavior with respect to the traditional imaging, e.g., the filtering process should not necessarily be spatially invariant, and could rather take advantage from the local geometrical structure of the scene. Thus, it is reasonable to question about whether one can design a new filter which uses depth information and still engenders a scale space.

In this paper we address this question and offer two contributions: i) we propose an anisotropic diffusion process based on the scene depth map that generates a scale space; ii) we show experimentally that keypoints detected in the proposed scale space are more robust to viewpoint changes than those obtained through the Gaussian scale space. Our proposed filter inherits the adaptivity properties of Perona and Malik's approach [5], but provides similar stability and linearity properties as the Gaussian filter. More precisely, it is numerically stable and linear as a function of the texture map, it satisfies the *causality* and *semigroup* axioms and enables a sort of edge-preserving filtering, i.e., it smooths out texture regions belonging to the same surface, while it prevents smoothing over boundaries of objects with different depth.

The rest of the paper is organized as follows. We introduce the necessary background on scale spaces and present related work in Section 2. In Section 3 we select a scene surface parametrization and define a discrete Laplacian operator to set up the diffusion process. We show that the resulting process generates a scale space, propose a numerical scheme for the diffusion simulation and ensure its numerical stability. At the end of that section we also discuss the relation with the Laplace-Beltrami operator that generalizes diffusion processes on manifolds. In Section 4 we illustrate the theoretical results with examples of filtered images and show that the proposed filter can increase the robustness to viewpoint changes in a keypoint detection scenario. Finally, Section 5 concludes the paper.

2. BACKGROUND AND RELATED WORK

One of the most common linear image smoothing operators is the convolutional filter with uniform Gaussian kernel:

$$K_\sigma(x, y) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right). \tag{1}$$

The kernel separability allows for a faster filter response computation: the 2D convolution may be replaced by a set of 1D convolutions over image lines and columns. The importance of Gaussian filter, as well as its principal advantage to other common low-pass filters, raises from diffusion equation framework [1]. It is well-known that PDE problem:

$$\begin{cases} \dfrac{\partial f}{\partial t} = \dfrac{\partial^2 f}{\partial^2 x} + \dfrac{\partial^2 f}{\partial^2 y} \equiv \Delta f \\ f|_{t=0} = f_0 \end{cases} \tag{2}$$

where Δf is the Laplacian operator, possesses a unique solution $f(t, x, y) = \left(K_{\sqrt{2t}} * f_0\right)(x, y)$, with $*$ denoting the convolution product. With this setup, a set of properties may be established for the Gaussian smoothing, proving that a sequence of progressively smoothed images forms a *scale space*. According to the definition
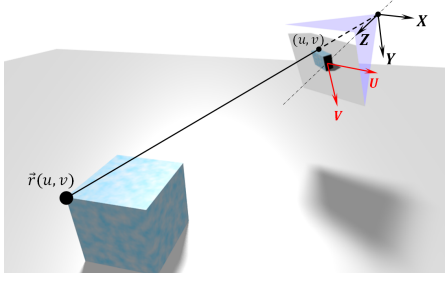
Fig. 1: Image surface parametrization in local camera coordinates.

proposed by Koenderink in [2], to be a scale space such a set of images of different scales must satisfy two hypotheses:

- **causality** (non creation of local extrema), i.e. any feature at a coarse level of resolution[1] is required to possess a (not necessarily unique) "cause" at a finer level of resolution;

- **homogeneity and isotropy**, i.e. the smoothing is spatially invariant.

Alternative definitions reject the second hypothesis in order to perform a "semantically consistent" smoothing to smooth inside the objects but not across boundaries. Examples of spatially adaptive smoothing filters that preserve image structure abound in the literature, from the classic bilateral filter [8] to the recent work in [9] on guided filter that preserves edges using an arbitrary guidance image, just to mention a few. However, no scale space properties have been proved for them so far. The first model of anisotropic scale space was first proposed by Perona and Malik [5], who formulated a non-linear PDE problem in such a way that the diffusion process is controlled by image gradient norm (no other sources of geometrical information are involved). A disadvantage of that approach is that the problem seems ill-posed, and the diffusion process simulation may be unstable [1]. Similarly to [5], our diffusion process is not isotropic; however, it is linear in the texture signal and numerically stable. Further scale space generalizations, notably the *anisotropic diffusion filtering* where the diffusivity becomes non-scalar, are discussed in [1, 10].

Recently, a scale space formulation has been proposed in [11] for lightfield content. The authors prove that their scaling is a semigroup, but they do not prove that it satisfies the causality axiom, which is the crucial property that relates scale spaces to smoothing [12].

3. PROPOSED SCALE SPACE FILTER DESIGN

3.1. PDE problem formulation

Let $\Omega = \left[-\frac{W}{2}, \frac{W}{2}\right] \times \left[-\frac{H}{2}, \frac{H}{2}\right]$ be the image support. We denote by $D : \Omega \to \mathbb{R}^+$ the depth map associated to the image I being processed. We suppose known the horizontal angle of view ω of the camera. It can be easily shown that the function $\vec{r} : \Omega \to \mathbb{R}^3$ defined below parametrizes the image surface in local camera coordinates as illustrated in Figure 1):

$$\vec{r}(u, v) = \begin{pmatrix} 2u \tan \frac{\omega}{2} \\ 2v \frac{H}{W} \tan \frac{\omega}{2} \\ 1 \end{pmatrix} D(u, v). \tag{3}$$

Let us now proceed to a discrete image support Ω_d obtained by sampling Ω with step h in both dimensions. For the same function f

---

[1]Here *resolution* means *scale* and not the image size.

still defined on the continuous support Ω, we introduce the following differential quantities:

$$\partial_u f = \frac{f(u+h, v) - f(u-h, v)}{\|\vec{r}(u+h, v) - \vec{r}(u-h, v)\|} = \frac{f(u+h, v) - f(u-h, v)}{r_u^{+-}} \tag{4}$$

$$\partial_v f = \frac{f(u, v+h) - f(u, v-h)}{\|\vec{r}(u, v+h) - \vec{r}(u, v-h)\|} = \frac{f(u, v+h) - f(u, v-h)}{r_v^{+-}} \tag{5}$$

where r_u^{+-} and r_v^{+-} are introduced in order to simplify notations. Recurrent applications of this operators give second order differential quantities, e.g. $\partial_{uu} f = \partial_u (\partial_u f)$. For a better operator kernel locality, we also introduce a definition through one-sided finite differences as follows:

$$\partial_{u+} f = \frac{f(u+h, v) - f(u, v)}{\|\vec{r}(u+h, v) - \vec{r}(u, v)\|} = \frac{f(u+h, v) - f(u, v)}{r_u^+} \tag{6}$$

$$\partial_{u-} f = \frac{f(u, v) - f(u-h, v)}{\|\vec{r}(u-h, v) - \vec{r}(u, v)\|} = \frac{f(u, v) - f(u-h, v)}{r_u^-} \tag{7}$$

$$\partial_{uu} f = \frac{\partial_{u+} f - \partial_{u-} f}{r_u^{+-}} \tag{8}$$

$$= \frac{f(u+h, v)}{r_u^+ r_u^{+-}} - \frac{f(u, v)}{r_u^+ r_u^{+-}} - \frac{f(u, v)}{r_u^- r_u^{+-}} + \frac{f(u-h, v)}{r_u^- r_u^{+-}} \tag{9}$$

and $\partial_{vv} f$ is then defined in a similar way. Thus, we may introduce a Laplacian-like second order differential operator $L \equiv \partial_{uu} + \partial_{vv}$ and then formulate the following PDE problem, linear in f_0:

$$\begin{cases} \dfrac{\partial f}{\partial t} = Lf \\ f|_{t=0} = f_0 \end{cases} \tag{10}$$

This problem is very similar to the classic diffusion problem (2). To study this similarity and set up some useful properties, let us return back to the continuous definition domain. We obtain continuous generalizations of our differential quantities by tending h to zero:

$$\begin{aligned} \mathcal{D}_u f &= f_u \|\vec{r}_u\|^{-1} \\ \mathcal{D}_{uu} f &= f_{uu} \|\vec{r}_u\|^{-2} - f_u \|\vec{r}_u\|^{-4} (\vec{r}_u, \vec{r}_{uu}) \end{aligned} \tag{11}$$

Thus we get the continuous version of the problem (10).

$$\begin{cases} \dfrac{\partial f}{\partial t} = \mathcal{D}_{uu} f + \mathcal{D}_{vv} f \\ f|_{t=0} = f_0 \end{cases} \tag{12}$$

We notice that if D is constant (i.e., we have a non-informative depth map), this problem is purely equivalent to the classic linear diffusion filtering (2), as the differential operator on the right side of the equation becomes the classic Laplacian up to a constant multiplier due to $\vec{r}_u = \vec{r}_v \equiv const$ and $\vec{r}_{uu} = \vec{r}_{vv} \equiv 0$. This allows for a "backward compatibility" of the proposed scale space to the classic Gaussian scale space in cases when the depth map is not provided. Moreover, this property is satisfied locally, i.e., at points where D is continuous and the surface normal is parallel to the camera optical axis.

3.2. Well-posedness, scale space properties and numerical solution

To establish the well-posedness of problem (10) as well as the properties of its solution, we use some of the results of [1]. We rewrite (10)

in a vector form, i.e., $f(t) \in \mathbb{R}^{W \times H}$ and the application of L to f is represented by a matrix multiplication $\mathcal{A}f$. The coefficients of matrix $\mathcal{A}$ depend only on $\vec{r}$ and are explicitly deduced from definition (9).

First we apply theorem 4 of [1]. It is straightforward to show that the operator matrix satisfies all the conditions except the symmetry, i.e., it has vanishing row sums *(S3)*, nonnegative off-diagonals *(S4)* and is irreducible *(S5)*. Lipschitz-continuity *(S1)* is satisfied unconditionally as $\mathcal{A}$ does not depend on f. The violated condition of the matrix symmetry *(S2)* is not required for well-posedness and maximum-minimum (causality) principle, as it is noticed afterwards [1, p. 76]. Thus, we prove that not only is the problem well-posed (i.e., has unique solution that depends continuously on the initial data), but that the solution f respects the extremum principle allowing to set up the causality.

Furthermore, theorem 8 of [1] proves a sufficient criterion of stability for the following explicit numerical scheme that allows to simulate the diffusion process:

$$f^{(n+1)} = f^{(n)} + \tau A f^{(n)}$$
$$f^{(0)} = f_0 \qquad (13)$$

The condition of stability consists in limiting the temporal step of simulation τ. We reinterpret theorem 8 of [1] to obtain the analytic expression:

$$\tau \leq \tau^* = \left[2 \max_{\Omega_d} \left\{ \frac{1}{r_u^+ r_u^{+-}} + \frac{1}{r_u^- r_u^{+-}} + \frac{1}{r_v^+ r_v^{+-}} + \frac{1}{r_v^- r_v^{+-}} \right\} \right]^{-1} \qquad (14)$$

Now we are able to perform the computation of the filter response for a given image $I = f_0$ and depth map D. The quantity of resulting smoothing is determined by $t^{(n)} = n\tau$.

Thus, we designed a linear smoothing filter that engenders a scale space. More complex filter behavior, such as edge preserving, mainly comes from the definition of the first order differential operators (5), where we adjust a common approximation of derivative computed on two neighboring samples by the real distance between the corresponding sample points, known from the depth map. This allows to respect the object contours by limiting the contribution of neighboring samples belonging to the background rather than to the object. For surface discontinuities in real images the denominator value of (5) is generally large enough to reduce the background contribution to zero.

3.3. Relation to the Laplace-Beltrami operator

The gradient-like quantities defined in (5) enable to smooth the image *intrinsically* to the surface, using the geometric properties conveyed by depth. Even if the proposed Laplacian $\mathcal{D}_{uu} + \mathcal{D}_{vv}$ is not strictly invariant to orthogonal coordinate changes, in case of smooth surfaces and limited high frequency variations of texture its response remains stable. Moreover, the operator vanishes on locally planar surfaces when the texture is a linear function, whereas classic 2D Laplacian, applied to the texture image only, does not. This allows for a *viewpoint-coherent* filter behavior, as observed in filtered images (see Section 4).

Filtering intrinsic to a surface is naturally formalized as a diffusion process on manifolds. These processes are classically described through the *Laplace-Beltrami operator* [13, 10]. The scene surface parametrization $\vec{r}$ may be regarded as a mesh with boundaries having a regular local topology but very irregular vertex spatial density. Thus, it is possible in theory to establish a diffusion process over such a mesh by using the Laplace-Beltrami operator. An important property
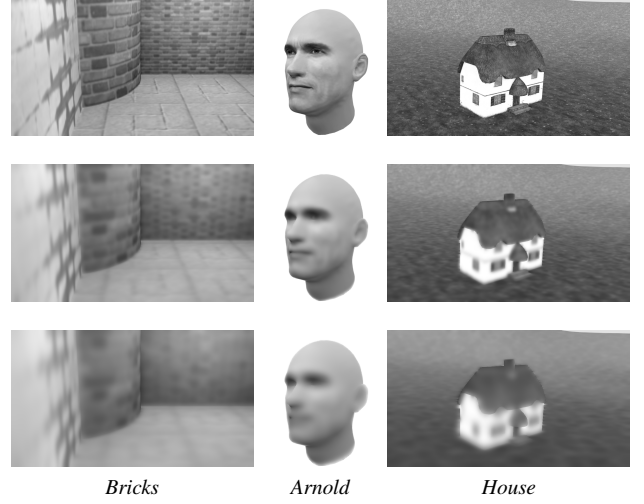

Bricks   *Arnold*   *House*

Fig. 2: Examples of filtered images. The proposed diffusion process preserves the discontinuities across different objects, while it smooths out progressively pixels that belong to the same surface.

of the surface-intrinsic characteristics of the Laplace-Beltrami operator is that the diffusion process is, in principle, "viewpoint-covariant", i.e., its action in a given point of the scene surface is completely independent of the camera position and orientation. Nevertheless, as we aim at designing a scale space, such an approach exhibits two important difficulties.

First, such a filter might not necessarily engender a scale space. One may easily verify that the proposed Laplacian $\mathcal{D}_{uu} + \mathcal{D}_{vv}$ acts exactly as the Laplace-Beltrami operator at points where the tangent plane to $\vec{r}$ is parallel to the camera plane. However, in general our Laplacian is different from Laplace-Beltrami operator. This is partially a consequence of the absence of a term containing mixed derivative f_{uv} in the analytic expression defining the operator (11). A Laplacian could be defined differently, but some coefficients of the operator matrix $\mathcal{A}$ (especially corresponding to the absent mixed derivative term) would become negative, and neither the scale space properties nor the numerical stability would be ensured by referring to [1].

Second, the Laplace-Beltrami operator is known to be hardly discretizable. Specifically, numerous discrete Laplacian operators do not converge to the Laplace-Beltrami operator, but satisfy some other desired properties and are largely used in practice [13, 14]. As for our Laplacian, it is easy to verify that, referring to the results in [14], it satisfies *locality*, *linear precision*, and *positive weights* properties.

4. EXPERIMENTS AND DISCUSSION

4.1. Qualitative assessment

We tested our filtering process on some synthesized images, for which we know exactly the corresponding depth maps. Some examples of filtered images are presented in Fig. 2. As it can be seen from the pictures, smoothing preserves surface discontinuities in each example: rounded wall boundary in *Bricks*, nose contour in *Arnold*, roof top border in *House*.

As we noticed before, our filter response is intrinsic to the surface and less dependent on the camera position with respect to the scene
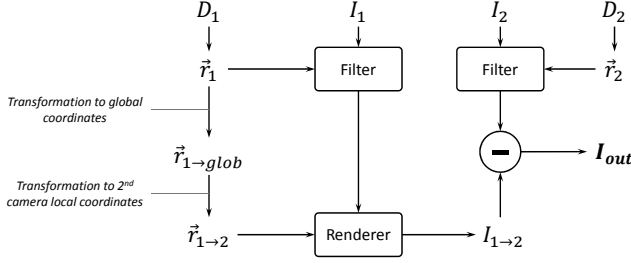
Fig. 3: Setup for simulating filter coherence to viewpoint position changes.

Fig. 5: Matching score for *Bricks* and *Graffiti* sequences. *Graffiti* texture+depth sequence is re-synthesized from the frontal view of the original *Graffiti* sequence used in [17].

content. We illustrate this showing the similarity between filtered images corresponding to different viewpoints. The simulation follows the scheme in Fig. 3. A given scene is captured from two different positions obtaining two views (I_1, D_1) and (I_2, D_2). As the camera positions and orientations are known, we may reproject the first view texture image on the second camera plane, i.e., reconstruct partially the second view from the first one. The corresponding image is referred to as $I_{1\rightarrow2}$. When no filter is applied, $I_{1\rightarrow2}$ is close to I_2 at pixels whose 3D origins are present in both images. The coherent filter behavior to viewpoint changes will manifest itself by a limited dynamics of $I_{1\rightarrow2}$.

We compare our filter with Gaussian and Perona and Malik's filters. The amount of smoothness, as well as all the other filter parameters, are set up experimentally in such a way that the filtered images are visually similar. To reduce sampling and depth quantization effects, instead of pixel-by-pixel difference, we take a set of neighboring points in $I_{1\rightarrow2}$ and compute the minimal difference with the corresponding pixel value in I_2. More precisely, we take 8 points on a circle of 1 pixel radius and its center. As $I_{1\rightarrow2}$ is sampled from a scattered point set obtained by the reprojection, we may interpolate it in the desired way. An example of the filtering process is reported in Fig. 4 for the *Bricks* content, where one can see that filtered images from different viewpoints match better with each other than those obtained through conventional scaling processes. This illustrates pictorially why the proposed scale space yields a certain viewpoint change robustness. In the following we discuss more quantitatively this property for a keypoint detection scenario.

4.2. Application to keypoint detection

In our second experiment we show how the proposed scale space can increase the repeatability in keypoint detection under substantial viewpoint changes. It is known that the classic SIFT keypoint detector [3] based on Gaussian scale space exhibits limited repeatability under significant viewpoint position changes [3, 17]. This kind of keypoint stability may be improved by using a scale space that changes in a covariant way with camera displacements. Moreover, anisotropic diffusion may have a positive influence on the keypoint stability as well [15]. To check this out, we compute scale space pyramid as described in [3] using our filter instead of the Gaussian, and inject the computed images in SIFT detector implemented in the *VLFeat* library [16]. We use this approach to detect keypoints in several different views of a given scene and then match the computed keypoint sets (the first view is matched to the other). The scenes we use are provided with ground-truth data allowing to compute exact position of 3D spherical area of each keypoint in camera-independent coordinates. Similarly to [17], the detector performance is evaluated in function
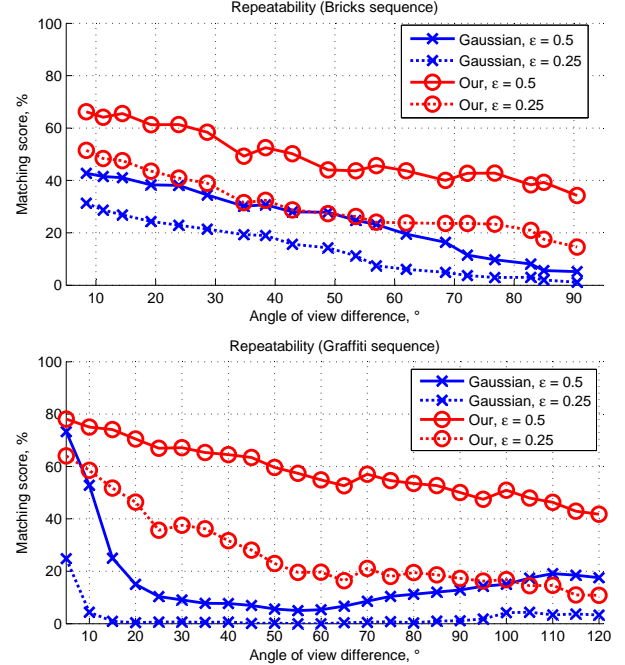
of the *matching score* representing the portion of matched keypoints to the maximum possible number of matches. Specifically, two keypoints i and j are matched if the volume of intersection of the spheres S_i and S_j is large enough, i.e., $|S_i \cap S_j| \geq (1 - \epsilon)|S_i \cup S_j|$ for $\epsilon \in (0, 1)$. This is an important difference to [17], where the intersection of flat ellipses (and not spheres) is computed, as entirely planar scenes are used for tests. To keep the comparison with SIFT fair, in our test we set the initial amount of smoothing σ_0 for each scene (for all the views) in such a way that the numbers of detected keypoints in all the images remain comparable to those of the standard SIFT detector (all the other processing parameters for the both detectors are strictly the same).

The matching score in function of view angle difference between the views is presented in Fig. 5, which clearly shows a larger robustness of the proposed scale space over the Gaussian one used in SIFT. We also show in Fig. 6 an example of images with matched keypoints. The number of keypoints in both images is approximately the same, but with the proposed scale space the number of matches is much higher.

4.3. Discussion

The proposed diffusion scheme assumes the knowledge of geometric information under the form of depth map. In this work, we assume that the depth information is noiseless, i.e., we use synthesized scenes from 3D models where the geometric information (including the camera parameters) are perfectly known. When a real data is used, an issue may be to pre-process the depth map in such a way that our proposed scale space be robust to imperfect knowledge of the geometry.

From the computational point of view, our approach has a cost equivalent to the anisotropic filtering of [5], given that it requires
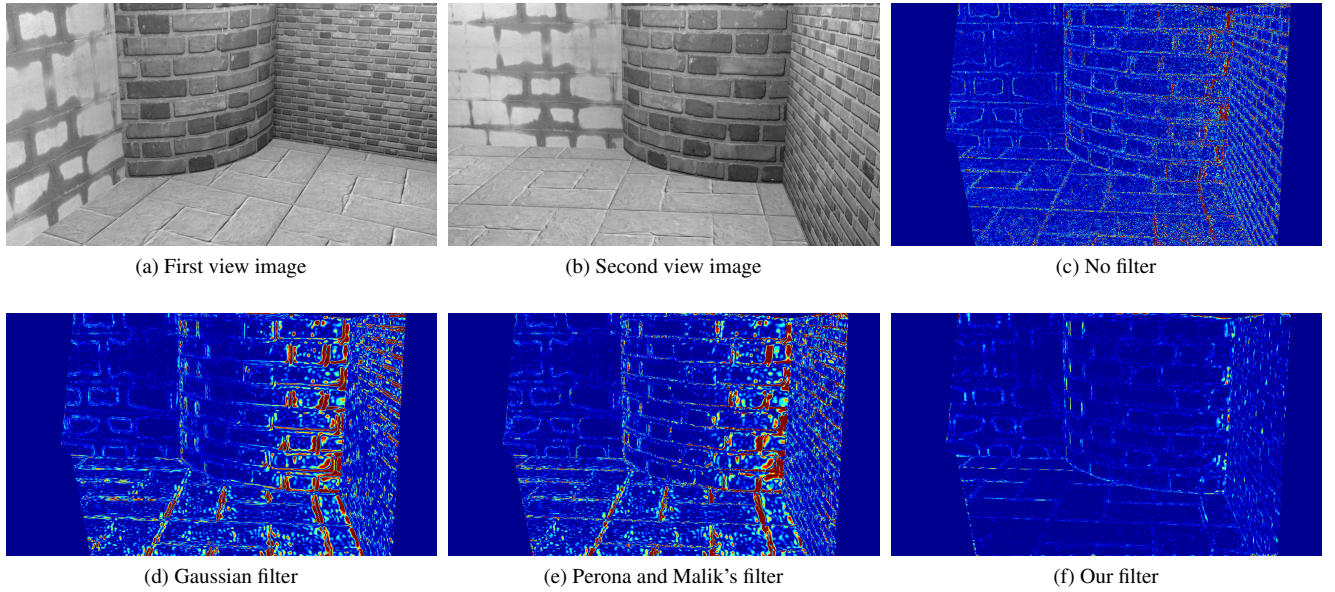
(a) First view image

(b) Second view image

(c) No filter

(d) Gaussian filter

(e) Perona and Malik's filter

(f) Our filter

Fig. 4: The first (I_1) and the last (I_2) view of the test scene and $I_{out} = |I_2 - I_{1\rightarrow 2}|$ computed without filter, with Gaussian, Perona and Malik, and our filter for an image from *Bricks* sequence. The difference images are thresholded by value of 3% of overall intensity scale for better visibility.
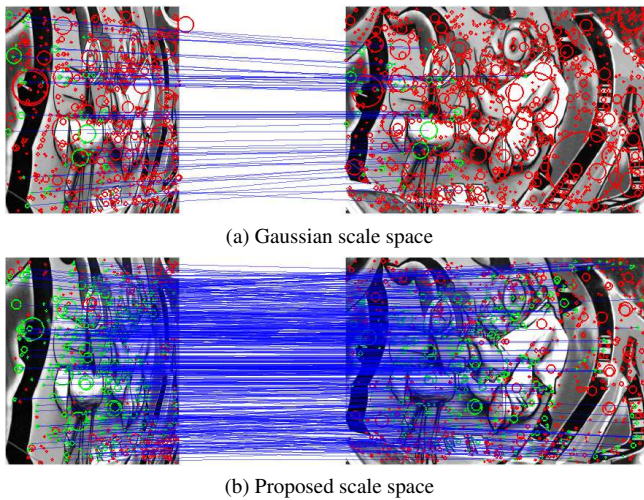

(a) Gaussian scale space

(b) Proposed scale space

Fig. 6: Matching of SIFT keypoints detected using different filters in a pair of images from *Graffiti* sequence (angle of view difference is 70°, $\epsilon = 0.5$). Green keypoints are the matched ones, i.e. occupying close spherical areas of the scene (the matching pairs are connected by blue lines). Red keypoints have no matches.

a diffusion simulation and may not be computed through a convolution. However, the scheme in (13) is trivially parallelizable, thus its complexity can be scaled linearly if implemented on a multicore architecture . Moreover, the operator matrix $\mathcal{A}$ is very sparse (up to border effects, for an image of $N \times M$ pixels it contains only $5NM$ nonzero elements) and does not change with iterations. This means that the simulation may be further accelerated by precomputing the operator matrix power $\mathcal{A}^m$ and applying it to the image instead of $\mathcal{A}$, i.e. $f^{(n+m)} = \mathcal{A}^m f^{(n)}$ – $\mathcal{A}^m$ storage cost is reasonable and its application to $f^{(n)}$ is fast thanks to the sparsity. The latter property is not valid for classical anisotropic filtering, where the matrix $\mathcal{A}$ is changing during simulation.

5. CONCLUSION

In this paper we designed an image smoothing filter for texture+depth content that makes use of the depth image in order to smooth the texture image. Our filter is linear, preserves internal scene structure avoiding the smoothing through object boundaries, and engenders a scale-space. We provided a stable numerical scheme for the filter response computation. Finally, we illustrated the viewpoint-coherent filter behavior and showed that SIFT detector with our scale space replacing the Gaussian one has better repeatability in case of significant differences between viewpoint positions.

In future work we will address the local visual features stability to significant viewpoint position changes and out-of-plane rotation with such a scale space, not only on the detector but on the descriptor side as well.

6. REFERENCES

[1] J. Weickert, *Anisotropic diffusion in image processing*. Teubner Stuttgart, 1998, vol. 1.

[2] J. J. Koenderink, "The structure of images," *Biological cybernetics*, vol. 50, no. 5, pp. 363–370, 1984.

[3] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International journal of computer vision*, vol. 60, no. 2, pp. 91–110, 2004.

[4] S. Leutenegger, M. Chli, and R. Y. Siegwart, "BRISK: Binary robust invariant scalable keypoints," in *IEEE International Conference on Computer Vision (ICCV), 2011*. IEEE, 2011, pp. 2548–2555.

[5] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 12, no. 7, pp. 629–639, 1990.

[6] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, "Speeded-up robust features (SURF)," *Computer vision and image understanding*, vol. 110, no. 3, pp. 346–359, 2008.

[7] F. Dufaux, B. Pesquet-Popescu, and M. Cagnazzo, *Emerging Technologies for 3D Video: Creation, Coding, Transmission and Rendering*. John Wiley & Sons, 2013.

[8] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *Computer Vision, 1998. Sixth International Conference on*. IEEE, 1998, pp. 839–846.

[9] K. He, J. Sun, and X. Tang, "Guided image filtering," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 35, no. 6, pp. 1397–1409, 2013.

[10] G. Sapiro, *Geometric partial differential equations and image analysis*. Cambridge university press, 2006.

[11] I. Tosic and K. Berkner, "3d keypoint detection by light field scale-depth space analysis," in *Proceedings of the International Conference on Image Processing*, Paris, France, October 2014.

[12] T. Lindeberg, "Scale-space theory: A basic tool for analyzing structures at different scales," *Journal of applied statistics*, vol. 21, no. 1-2, pp. 225–270, 1994.

[13] G. Xu, "Discrete Laplace–Beltrami operators and their convergence," *Computer Aided Geometric Design*, vol. 21, no. 8, pp. 767–784, 2004.

[14] M. Wardetzky, S. Mathur, F. Kälberer, and E. Grinspun, "Discrete Laplace operators: no free lunch," in *Symposium on Geometry processing*, 2007, pp. 33–37.

[15] M. Gobara and D. Suter, "Feature detection with an improved anisotropic filter," in *Computer Vision–ACCV 2006*. Springer, 2006, pp. 643–652.

[16] A. Vedaldi and B. Fulkerson, "Vlfeat: An open and portable library of computer vision algorithms (2008)," 2012.

[17] K. Mikolajczyk, T. Tuytelaars, C. Schmid, A. Zisserman, J. Matas, F. Schaffalitzky, T. Kadir, and L. Van Gool, "A comparison of affine region detectors," *International journal of computer vision*, vol. 65, no. 1-2, pp. 43–72, 2005.