Affine Non-local Means Image Denoising

Vadim Fedorov, and Coloma Ballester

Abstract—**This work presents an extension of the Non-Local Means denoising method, that effectively exploits the affine invariant self-similarities present in images of real scenes. Our method provides a better image denoising result by grounding on the fact that in many occasions similar patches exist in the image but have undergone a transformation. The proposal uses an affine invariant patch similarity measure that performs an appropriate patch comparison by automatically and intrinsically adapting the size and shape of the patches. As a result, more similar patches are found and appropriately used. We show that this image denoising method achieves top-tier performance in terms of PSNR, outperforming consistently the results of the regular Non-Local Means, and that it provides state-of-the-art qualitative results.**

Index Terms—**Image denoising, patch-based method, patch similarity, affine invariance.**

I. INTRODUCTION

THERE are many different ways of obtaining a digital image of a real world scene: digital cameras, scanning devices, etc. Unless an image is generated completely by a computer, it is subject to the acquisition noise. Many different kinds of denoising algorithms are widely used to suppress this noise while preserving image details, textures and colors. They are embedded in firmware of digital cameras and are also applied in the very beginning of image and video processing pipelines.

This work presents an extension of the Non-Local Means framework that effectively exploits the affine invariant self-similarities present in images of real scenes. Like the Non-Local Means strategy, our method considers 2D patches (local neighborhoods) and works by collecting the most similar patches and averaging them weighted by their similarity. Besides, our proposal grounds on the fact that in many occasions similar patches exist in the image but have undergone a transformation. Indeed, the geometry of the observed 3D scene affects the appearance of local neighborhoods, for example due to a different position with respect to the camera, and also affects patch comparison. Our proposal uses an affine invariant patch similarity measure that considers and estimates the associated distortions from local texture content and automatically adapts the size and shape of the patches to perform the appropriate comparison. As a result, more similar patches are found and used in the denoising of the image, which provides better results. We demonstrate that our method outperforms the Non-Local Means in terms of PSNR values and also produces results that visually look as good as the state-of-the-art results or, in some cases, even more pleasing.

The remainder of the paper is organized as follows. In Section II we revise previous work on image denoising.

V. Fedorov and C. Ballester are with the Department of Communication and Information Technologies, University Pompeu Fabra, Barcelona 08018, Spain (e-mail: vadim.fedorov@upf.edu; coloma.ballester@upf.edu).

Section III motivates our approach and reviews the required definitions. We explain the proposed method in Section IV, including the most important implementation details. Then in Section V we briefly discuss the choice of parameters of the method. In Section VI we present both quantitative and qualitative assessments of the proposed method. Finally, Section VII concludes the paper.

II. RELATED WORK

The problem of removing the noise of an image while preserving its structure and characteristics like geometry, texture or color has been extensively investigated over the last decades. Some early works approach denoising by local methods, for example, using anisotropic diffusion [27]. The total variation denoising model of [34] inspired many variational formulations for image denoising. In [34] it was applied for images corrupted by the Gaussian noise, and has also been used for Riccian [24], salt-and-pepper [26] and multiplicative noise [2]. Other operators have been proposed, e.g., the total generalized variation [4] or non-local regularizers as [15], [5]. Frequency domain methods have been developed and extensively studied as well, e.g., windowed discrete cosine transform (DCT) [38], [39], wavelet transform based methods [10], [32], curvelets [35], etc. More recently it was proposed [11], [23] to learn dictionaries for sparse and redundant image representation from images themselves instead of exploiting predefined ones (like wavelets or DCT). These methods attempt to model the space of image patches.

Non-local or patch-based approaches are used in most of the state-of-the-art methods for image denoising, restoration, super-resolution, inpainting and object recognition [6], [28], [33], [31], [14], [30], [17], [12]. Image denoising has gone along with the advances in patch-based techniques. Non-Local Means [6] exploits the notion of non-local self-similarity of image patches and works by collecting and averaging the most similar patches in the image. In the current work we present an extension of this seminal approach that inspired many denoising methods and has also led to the study and development of non-local regularity priors [29], [1]. One of the state-of-the-art denoising methods, the BM3D method [8], combines spatial self-similarity with thresholding in the frequency domain through the collaborative filtering of groups of similar patches. The resembling method of [40] exploits principal component analysis instead of DCT. Both methods involve two almost identical steps: the first one provides the basic estimate and the second one refines the final result using the initial estimate as a guide. A Bayesian interpretation of the method of [40], named Non-Local Bayes, was recently proposed in [18]. Non-Local Bayes is a spatial-based method that improves Non-Local Means by considering a Gaussian
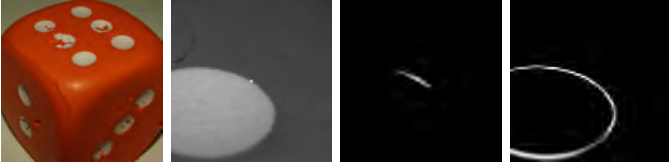
Fig. 1. From left to right: original Dice image, a zoom of a search window with its central (reference) point shown in white, similarity values computed between the central point and all other points using the similarity measure (2) and square patches, and using the proposed affine invariant patch similarity measure. White color encodes the highest similarity.

probability model for each set of similar patches. In many denoising works it is assumed that the noise type and level are known in advance. The Noise Clinic method [20] for blind denoising goes beyond this assumption and combines a noise model estimation with a multiscale modification of the Non-Local Bayes method.

Let us finally note that rotationally invariant patch comparison for image denoising was already studied in the past [41], [16], [42]. Besides that, it was noticed that denoising with regular square patches may cause noise halos around contrasted edges. To remove this kind of artifacts it was proposed in [9] to replace square patches by several types of neighborhoods with various shapes (discs, bands, half-discs, quarter-discs) and to combine their estimations using Stein's Unbiased Risk Estimate [36]. Our similarity measure originates from a multiscale patch comparison theory [3], [13] and considers a richer space of patches that includes all affine-transformed patches. However, as we will state in the following Section, for each pair of patches the transformations are uniquely determined using the image content.

III. AN AFFINE INVARIANT SIMILARITY MEASURE

In this section we recall the definition of an affine invariant patch similarity measure first given in [13]. We start by motivating the use of the proposed similarity measure for patch-based denoising. Additionally, we propose a modification that better adapts the similarity measure for denoising.

A. Motivation

Recent progress in image denoising is associated with the concept of self-similarity which is pervasive for natural images. In a broad sense, such methods as [6], [18], [8], [17] work by combining neighborhoods of points or patches of a noisy image that appear similar to each other and extracting from them a common estimate of the original noiseless image. The basic principle is that, by the variance law in probability theory, the noise standard deviation of the average of samples decreases. This approach is intuitively consistent, if one can find many samples for every image detail. In order to find patches which are most similar to a reference patch, many denoising methods use a simple patch distance measure:

$$\mathcal{D}(x,y) = \frac{1}{|P|} \sum_{i \in P} (u(x+i) - v(y+i))^2, \qquad (1)$$

where P denotes a neighborhood of the origin $0 \in \mathbb{R}^2$ forming a patch, and $|P|$ denotes the area of P (in practice, the number of pixels in P). This can be rewritten in the continuous setting as:

$$\mathcal{D}(x,y) = \frac{1}{\varrho} \int_{\mathbb{R}^2} \eta(h)(u(x+h) - v(y+h))^2 \, dh, \qquad (2)$$

where $\eta(h)$ is a characteristic function of the patch which equals to 1 when $h \in P$ and 0 otherwise. Sometimes η is a weighted characteristic function such as a Gaussian. Anyhow, ϱ is the normalizing factor.

Usually, patches are set to be squares and no transformations are allowed between two patches being compared. Since it is numerically impractical to check the whole image, similar patches are searched in a relatively small window around a reference point. The usual size of the search window ranges from 21 by 21 to 35 to 35 pixels. This allows one to reduce the overall computational expenses, however, when the level of noise increases, it is becoming less probable to encounter enough similar patches. We claim that allowing for some transformations between patches being compared can increase the amount of similar patches within the search window and thus improve the denoising result. Fig. 1 illustrates it with a simple experiment on the classical Dice image. For a given point and a search window around it we calculate similarity between that point and all other points in the window using both the similarity measure (2) and our proposal. As can be seen, in our case much more points along the edge receive high similarity values. This in turn leads to a better denoising result, as shown in Section VI.

An appropriate similarity measure was recently proposed in [13]. Like (2), it operates on intensities but now within elliptical patches whose size and shape are not fixed a priori. It is defined by taking into account the underlying local geometrical content of the image, thus is related to the shape-from-texture approach, and uses structure tensors to define shapes of patches and to extract a local transformation between any two of them. The proposed structure tensors are proved to be affine covariant and the similarity measure in turn is affine invariant.

B. Affine Covariant Structure Tensors

First of all we introduce the notion of affine covariant structure tensors and explain how they can be used to define shape-adaptive patches. Given a real-valued image u, we consider an image-dependent tensor field T_u as a function that associates a tensor (a symmetric, positive semi-definite 2×2 matrix) to each point x in the image domain (which we assume to be $\mathbb{R}^2$ to simplify). The tensor field is said to be *affine covariant* if, for any affinity A,

$$T_{u_A}(x) = A^T T_u(Ax) A, \qquad (3)$$

where $u_A(x) := u(Ax)$ denotes the affinely transformed version of u. Given a tensor $T_u(x)$ we can associate to it an elliptical region of "radius" r centered at x

$$B_u(x,r) = \{y \in \mathbb{R}^2 : \langle T_u(x)(y-x), (y-x) \rangle \leq r^2 \}. \quad (4)$$

Fig. 2. Affine covariant neighborhoods (patches) computed at corresponding points in two images taken from different viewpoints. Despite the change in appearance, patches capture the same visual information.

$$T_v(y)^{\frac{1}{2}} \qquad R_v^{-1}(y) \qquad R_u^{-1}(x) \qquad T_u(x)^{\frac{1}{2}}$$
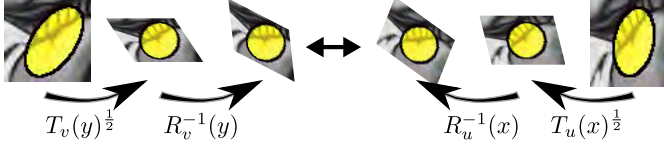
Fig. 3. An affine invariant patch comparison which is achieved by normalizing the patches to discs and aligning them with suitable rotations.

When the tensor is affine covariant, we have that $AB_{u_A}(x,r) = B_u(Ax,r)$. This means that the tensors can be used to define affine covariant regions which transform appropriately via an affinity. We will refer to these regions as *shape-adaptive* or *elliptical patches* to distinguish them from the square patches of fixed size. Fig. 2 illustrates the patches defined by the affine covariant tensors of [13], computed for a set of corresponding points in two images related by a homography. Note that even though the transformation is not an affinity, the patches still match, since a homography can be locally approximated by an affinity.

Affine covariant tensors at two corresponding locations allow us to extract the affine distortion between corresponding elliptical patches, up to a rotation, as shown in [13]. Indeed, for any affine transformation A, there exists an orthogonal matrix R such that

$$A = T_u(Ax)^{-\frac{1}{2}} R T_{u_A}(x)^{\frac{1}{2}}. \qquad (5)$$

This last equation provides an intuitive geometric relationship between the tensors, the associated elliptical patches and the affinity. Consider a point x and the corresponding affine covariant elliptical patch $B_{u_A}(x)$. Mapping $B_{u_A}(x)$ by the affinity yields $B_u(Ax)$. The application of A can be decomposed in three steps. First, applying $T_{u_A}(x)^{\frac{1}{2}}$, we transform $B_{u_A}(x)$ into a disc or radius r. We refer to the resulting patch as a *normalized patch*. Then, a rotation is applied to the normalized patch. Finally, $T_u(Ax)^{-\frac{1}{2}}$ maps the rotated normalized patch to the elliptical patch $B_u(Ax)$.

To fully determine the affinity A, one needs to find the rotation R. Any rotation would yield an affinity that maps the elliptical patch associated to T_{u_A} at x to the one associated to T_u at Ax. For a wrong value of the rotation, the image content inside both patches will not match. Therefore, the right value for the rotation can be computed by aligning the image content of both patches. For this aim, we decompose the rotation as $R = R_u(Ax)R_{u_A}^{-1}(x)$, where $R_u(Ax)$ and $R_{u_A}(x)$ are estimated from the image content inside the patches. In practice, we calculate them by aligning the dominant orientation of the normalized patches to the horizontal axis.

To compute the dominant orientation we use histograms of gradient orientations as in the SIFT descriptors [22].

C. Computation of Affine Covariant Tensors

In this work, we use the particular dense field of affine covariant tensors and associated neighborhoods proposed in [13], which are computed from the following iterative algorithm:

$$T_u^{(k)}(x) = \frac{\int_{B_u^{(k-1)}(x,r)} Du(y) \otimes Du(y)\, dy}{\mathrm{Area}(B_u^{(k-1)}(x,r))}, \qquad (6)$$

where u is the given image and $B_u^{(k)}$ is the elliptical patch related to $T_u^{(k)}$, defined by

$$B_u^{(k)}(x,r) = \{y : \langle T_u^{(k)}(x)(y-x), (y-x)\rangle \le r^2\} \qquad (7)$$

for $k \ge 1$, and

$$B_u^{(0)}(x,r) = \{y : |Du(x)^T(y-x)| \le r\} \qquad (8)$$

for $k = 0$. To simplify notation in the following, let us denote by $T_u(x)$ the affine covariant structure tensor $T_u^{(k)}(x)$ for a fixed value of k (in our experiments $k = 30$, as in [13]) and a fixed value of r (notice that $r > 0$ is a free parameter). Similarly, we denote by $B_u(x)$ the affine covariant neighborhood $B_u^{(k)}(x,r)$.

D. An Affine Invariant Patch Similarity

We are interested in comparing patches of an image. Let us generalize and consider the problem of comparing the patches around two points x, y defined in images u and v, respectively. We begin with an intuition behind the proposed similarity measure. The patches are defined by the local metric given by the tensors $T_u(x)$ and $T_v(y)$. In order to compare both patches, equation (5) suggests the following mapping between them:

$$\mathcal{C}(x,y) = T_v(y)^{-\frac{1}{2}} R_v(y) R_u^{-1}(x) T_u(x)^{\frac{1}{2}}. \qquad (9)$$

We can interpret $\mathcal{C}(x,y)$ as an affinity, mapping the elliptical patch associated to $T_u(x)$ into the one associated to $T_v(y)$. If u in the vicinity of x is an affinely transformed version of v in the vicinity of y, then $\mathcal{C}(x,y)$ recovers the true affinity. Intuitively, an affine invariant patch similarity could be computed as the distance between the elliptical patch at y and the result of applying $\mathcal{C}(x,y)$ to the patch at x. In practice, it is more suitable to transform both elliptical patches to discs of radius r (as depicted in Fig. 3) and to compare the aligned normalized patches:

$$\mathcal{D}^{\mathrm{a}}(t,x,y) = \int_{\Delta_t} g_t(h) \left(u(x + T_u^{-\frac{1}{2}} R_u(x)h) - v(y + T_v^{-\frac{1}{2}} R_v(y)h) \right)^2 dh, \qquad (10)$$

where Δ_t denotes a disc centered at the origin with radius proportional to $t > 0$ (the so-called scale) and big enough such that the weighting function g_t has effective support in Δ_t. $\mathcal{D}^{\mathrm{a}}$ is a multiscale affine invariant patch distance which intrinsically extends the set of available patches.

Let us now review the origins of the similarity measure corresponding to (10). It was derived in [13] as a computationally tractable approximation in the linear case of the multiscale analyses of similarity measures introduced in [3]. There, the authors show that all scale spaces of similarity measures $\mathcal{D}(t, x, y)$ satisfying a set of appropriate axioms are solutions of a family of degenerate partial differential equations. Images are considered in those papers as Riemannian manifolds endowed with a metric defined by a tensor field. If this tensor field is affine covariant, the resulting similarity measure is affine invariant. In this Riemannian framework, the defined affine covariant tensors can be seen as a metric in the image plane and $\mathcal{C}$ in (9) defines an isometry between the tangent spaces in the two manifolds $(\mathbb{R}^2, T_u(x))$ and $(\mathbb{R}^2, T_v(y))$. The authors refer to $\mathcal{C}$ as the *a priori connection*, since it is related to the notion of connection appearing in parallel transport (see [3] for details). WKB approximation method, named after Wentzel, Kramers and Brillouin, was used in [13] to obtain (10) as an approximate solution to a linear partial differential equation with spatially varying coefficients as a convolution with a short-time space-varying kernel. Let us finally notice that g_t in (10) can be either an approximated geodesic weighting function or a Gaussian, and represents a weighted characteristic function of both patches being compared and determines the comparison scale.

E. Patch Size Constraint

In most of the patch-based methods square patches of a fixed size are used. The size of the patches may change depending on the noise level, however, it is always the same within a single image. Commonly square patches are chosen to be small, their size ranges from 3 by 3 to 7 by 7 pixels (sometimes slightly bigger). Our method allows the patch shapes to adapt to the local image content, thus the patches may well have different sizes. We would like to preserve this shape-adaptiveness and at the same time to be able to limit the maximum patch size to capture only small pieces of visual information. To some extent the r parameter in (4) controls the size of a covariant elliptical patch, however, it also depends on the image content itself (see Fig. 4, left image). For the same value of r, elliptical patches are always significantly bigger in homogeneous regions than in textured regions or close to edges. As shown in Fig. 4, in the presence of noise elliptical patches tend to shrink due to the additional gradients associated with noise. This effect is accentuated in homogeneous regions, therefore elliptical patches become more uniform in size. On the other hand, as these additional gradients have random magnitudes and directions, the *shape* of patches does not change a lot. This phenomenon motivates the way of placing an upper boundary constraint on the size of elliptical patches.

We are interested in simulating additive Gaussian noise in our tensor computation scheme (Eq. (6)-(8)). From this perspective, additional noise can be seen as a multitude of extra gradients with random directions and expected magnitude related to the variance of noise. We approximate it by

adding a small constant value to the diagonal elements and define a new tensor to be

$$\widetilde{T}_u(x) = T_u(x) + \beta I, \qquad (11)$$

where β controls the amount of simulated noise and, therefore, the maximum allowed size of the associated elliptical patches, say it $\widetilde{B}_u(x)$. To relate β with the size constraint, let the reference point x be located in an infinitely wide region with uniform color. Since all the gradients are equal to zero, the term $T_u(x)$ is equal to zero as well and the resulting tensor is defined solely by β. If $\beta = 0$, an elliptical patch at x would have infinite size. If $\beta > 0$, the patch becomes a disc. From (4), its boundary is given by

$$\partial \widetilde{B}_u(x) = \{y : \langle \beta I(y - x), (y - x) \rangle = r^2\}$$
$$= \{y : \beta \|y - x\|^2 = r^2\}. \quad (12)$$

Let R_{max} be the radius of this disc. Then $\|y - x\|^2 = R_{max}^2$ and

$$\beta = \frac{r^2}{R_{max}^2}. \qquad (13)$$

The relation (13) allows us to parametrize the patch size constraint in terms of the maximum possible radius of a patch, shall it appear in a completely homogeneous region. Fig. 5 shows the effect of applying the size constraint with different values of R_{max}. Note that for $R_{max} = 5$, elliptical patches are almost uniform in size and still follow the image content.

Of course, this term, added to the diagonal of the structure tensors, breaks the affine covariant property of tensors. Most strongly it affects invariance of the similarity measure to scaling. This is easily verified mathematically and can be illustrated with the following experimental example (Fig. 6). Let two corresponding points be given on two images related by an affinity (a scaling in this case). Two elliptical patches calculated using the original scheme are shown in Fig. 6a and 6b. They cover the same visual information. Elliptical patches with the size constraint applied are shown in Fig. 6c and 6d. The patch shown in Fig. 6c cannot grow enough to take the edges into consideration. As a consequence, the corresponding elliptical patches do not capture the same visual information, moreover tensors do not recover the correct local affinity. For some applications this breaking of affine covariance is unacceptable, however, in the case of denoising we do not expect to encounter severe scaling within a relatively small search window.

The size constraint plays the most important role for the cases of small levels of noise when it prevents elliptical patches from growing too big in uniform regions. When there is a clearly defined structure within a patch it is less affected by the size constraint.

IV. NON-LOCAL MEANS DENOISING WITH AN AFFINE INVARIANT SIMILARITY MEASURE

In this section we describe our extension of the well known Non-Local Means denoising method.

Fig. 4. Elliptical patches on images with different levels of noise. From left to right, standard deviation of noise: 0 (original noiseless image), 2, 5, 10, 15.

Fig. 5. Elliptical patches on the original noiseless image, calculated using different size constraints (different values of R_{max} in (13)). From left to right, R_{max}: ∞ ($\beta = 0$, no size constraint), 20, 15, 10 and 5.
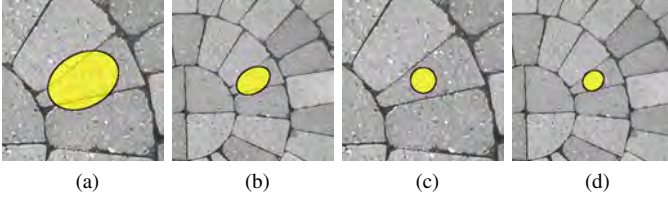

| (a) | (b) | (c) | (d) |

Fig. 6. Effect of constraining the patch size. (a) and (c) show part of an image in its original size, (b) and (d) show part of the same image scaled down by 2. (a) and (b) illustrate elliptical patches without any size constraints, (c) and (d) illustrate elliptical patches with the size constraint $R_{max} = 25$ applied.

A. Description of the Method

In this work we follow the strategy of the original Non-Local Means scheme proposed in [6]. In particular, given a noisy image we go through all of its points and, for every point, we calculate patch distances between that reference point x and other points y around it. As in the other works, we limit this set of surrounding points to a square window of a given size w, centered at the reference point x. Instead of the very common patch distance measure (2) we use the recently proposed measure (10). Consequently, instead of common square patches, we consider elliptical patches given by the affine covariant structure tensor at every point of an image. Note that in the denoising problem, v in (10) coincides with u.

To be able to denoise color images we adapt the point-wise square difference in (10) to operate on vector values:

$$\mathcal{D}^{a}(t, x, y) = \int_{\Delta_t} g_t(h) \left\| \mathbf{u}(x + T_u^{-\frac{1}{2}} R_u(x)h) - \mathbf{u}(y + T_u^{-\frac{1}{2}} R_u(y)h) \right\|^2 dh,$$

(14)

where $\mathbf{u}$ denotes a color images in RGB color space and u is a gray-scale image associated with it. Let us remark that we

use the gray-scale version u to calculate the structure tensor field and the associated shape-adaptive patches.

In order to convert distances into similarities we use the exponential function and define the similarity of the patches centered at x and y as:

$$\mathcal{S}(x, y) = e^{-\frac{\mathcal{D}^{a}(t, x, y)}{\lambda^2}},$$

(15)

where $\lambda = b\sigma$ is a bandwidth that depends on the standard deviation σ of the noise and is controlled by the parameter $b > 0$ of the method. Note that the distance $\mathcal{D}^{a}$ is already squared.

To denoise an elliptical patch at the reference point x, we average all the patches at surrounding points y using their similarity values $\mathcal{S}(x, y)$ as weights. Note that in this aggregation process elliptical patches have to be appropriately transformed to match the patch at x being denoised. We calculate this transformation from the structure tensors using (9). The denoised patch is then given by a weighted average:

$$\widehat{P}_u(x) = \frac{1}{\varrho(x)} \sum_{y \in W_w(x)} \mathcal{S}(x, y) \cdot \mathcal{C}(y, x) P_u(y),$$

$$\varrho(x) = \sum_{y \in W_w(x)} \mathcal{S}(x, y),$$

(16)

where $W_w(x)$ denotes the search window of size w and centered at x. $\mathcal{C}(y, x)$ denotes the a priori connection (9) that transforms an elliptical patch at y to an elliptical patch at x. That is, it first transforms an elliptical patch to a disc, then rotates this disc and finally transforms it into another elliptical patch. Of course, in the discrete setting, some kind of interpolation of color values has to be done after this transformation. For this we use the Nadaraya-Watson estimator [25], [37] with Gaussian kernel.

Since denoised patches may overlap each other, every pixel of the resulting image receives multiple color estimates.

Fig. 7. Left to right: original images, and homogeneous regions, in white, computed for noisy images with noise standard deviation 2, 10 and 30.
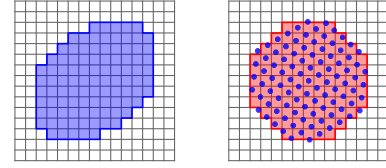

Fig. 8. Elliptical patch normalization. On the left an elliptical patch is shown in blue. On the right a set of scattered points of that patch, being normalized by $T_u(x)^{\frac{1}{2}}$ to a disc, is shown in blue over a regular grid which is shown in red. In this case, the grid resolution (number of grid nodes along the diameter) $g = 11$.

The estimates coming from a denoised patch are weighted depending on their distance to the center of that patch. The resulting color for a pixel is calculated as a weighted average of all estimates. We use here the same anisotropic intra-patch weights, given by the affine covariant tensors, as in the similarity measure (10). In this way patch aggregation complies with patch comparison.

Homogeneous region test. Our patch comparison by means of the affine invariant similarity measure can be related to the shape-from-texture techniques in the sense that the shape of elliptical patches is defined by the underlying image texture. Thus, in the absence of clear textural information (in a homogeneous region), the shape of elliptical patches becomes very dependant on the image noise. In the denoising application it leads to a high frequency noise that is suppressed but still remains in homogeneous regions. To reduce this dependency, the parameters of the method can be adjusted to allow for bigger patches, however, it usually leads to a blurry reconstruction in textured regions. To avoid contradiction in the choice of parameters for denoising homogeneous regions and fine image details (for example, textures and standalone edges) we treat them differently.

Motivated by [18], we first check whether a reference patch can be considered as located in a homogeneous region. For that we calculate the variance of all color values in a set of L patches in $W_w(x)$, most similar to the reference patch at x. L is a parameter of the method. As commented in [18], for homogeneous regions that color variance is expected to be small and close to the noise variance. We check if it is less than $H\sigma^2$, where σ is the noise variance and H is a thresholding parameter of the method. Note that we use the same values for both L and H in all our experiments (see Section V). If the homogeneous region test is passed (i.e., the patch belongs to a homogeneous region), the reference patch is denoised by filling it with a single average color $c(x)$, computed from the similar patches. For every channel separately:

$$c(x) = \frac{1}{\varrho(x)} \sum_{y \in \mathcal{P}_L(x)} \sum_{z \in B_u(y)} u(z),$$

$$\varrho(x) = \sum_{y \in \mathcal{P}_L(x)} \sum_{z \in B_u(y)} 1, \qquad (17)$$

where $\mathcal{P}_L(x)$ is a set of at most L centers of patches which are most similar to a patch centered at x. If the homogeneous region test is failed, the algorithm proceeds normally. Fig. 7 shows some examples, computed for noisy images with different levels of noise. The points satisfying the homogeneity

criterion are shown in white. With this trivial modification we can avoid using the similarity measure in the case it was not designed for.

B. Implementation Details

Patch comparison. In the proposed method elliptical affine covariant patches are used instead of the square patches of a fixed size. The scheme described in Section III-C allows one to calculate at any given point x both the structure tensor and the elliptical patch defined by it. As shown in Section III-B, an elliptical patch can be normalized to a disc of radius r using its corresponding tensor. Since in practice digital images are discrete, after normalization any patch turns into a set of scattered points (shown in blue in Fig. 8). In order to compare one normalized patch with another, we interpolate these sets of scattered points to a regular grid. For interpolation we use the Nadaraya-Watson estimator [25], [37], with Gaussian kernel which standard deviation $\bar{\sigma}$ depends on the density of the scattered points. More specifically, it depends on the radius of a normalized patch (that is determined by the parameter r of the method) and on the number of points within a patch being normalized:

$$\bar{\sigma} = \frac{r}{|B_u(x)|}. \qquad (18)$$

Since a normalized patch is a disc, the resolution of the regular grid can be specified by the number of grid nodes along its diameter. This grid resolution g is a free parameter of the method. The choice of g for different values of noise variance is described in Section V.

As commented in Section III-B, a local affine transformation can be estimated from two tensors, but only up to a rotation. To compensate for the missing rotation we calculate dominant orientations of the normalized patches after interpolating them to the regular grid. For this we use histograms of gradient orientations as in the SIFT descriptors [22]. Several dominant orientations can be extracted from a single patch and every one of them gives a possible additional rotation. Therefore, every patch can be associated with multiple candidate trans-formations $G_i(x) = R_{u,i}^{-1}(x) T_u(x)^{\frac{1}{2}}$, as many as there are dominant orientations. According to our observations, most commonly the number of dominant orientations does not exceed two. Of course, patch normalization and interpolation to the regular grid should be repeated for each of these transformations. When calculating patch distance between two patches, we compare every candidate normalization of the first
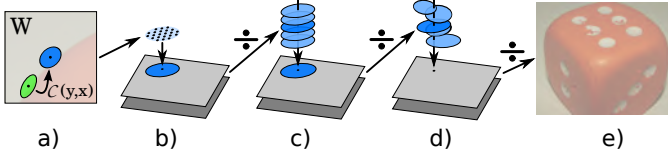
Fig. 9. Diagram illustrating the aggregation step of the proposed method: a) normalization (W denotes the search window); b) interpolation; c) patch denoising; d) color estimates aggregation; e) resulting denoised image.

with every candidate normalization of the second and keep the configuration that gives the smallest distance.

Since the tensor field is associated with the input noisy image, it can be precomputed in the very beginning. Moreover, the regular grid is constructed only once using a given resolution g, therefore, all possible normalized versions for every elliptical patch can also be computed in advance and stored in memory, together with their corresponding candidate transformations $G_i(x)$. In this way, the calculation of the proposed patch distance between two points boils down to several squared array differences as in (1).

Patch aggregation. As commented above, to compare elliptical patches in a more efficient way, we normalize them to the regular grid in the pre-processing step. In the aggregation step, similar patches have to be transformed once again to match a reference patch. For that we calculate the complete transformation $\mathcal{C}(y, x)$ using (9). Since halves of this transformation, namely $R_u^{-1}(x)T_u(x)^{\frac{1}{2}}$ and $R_u^{-1}(y)T_u(y)^{\frac{1}{2}}$, are computed in advance and stored in memory in pair with their corresponding normalizations, the $\mathcal{C}(y, x)$ computation costs only one matrix inversion and one matrix multiplication.

After transforming a similar patch to its reference patch, another interpolation is needed to aggregate color values from that similar patch. For this we use the same Nadaraya-Watson estimator as before, but for now the Gaussian standard deviation of the kernel is specified by a free parameter i of the method. This parameter allows us to control the sharpness of the denoising result. The choice of its value, same for all levels of noise, is discussed in Section V.

In the aggregation process, patches within the search window are weighted depending on their similarity (15) to the reference patch. This suggests the weight value of 1 for the reference patch itself. To avoid overweighting of the noise from the reference patch we instead take the maximum weight among the other patches in the search window.

For performance reason we use three pairs of buffers of the same size as an input image during the aggregation (see Fig. 9). In every pair, the first buffer accumulates color values and the second buffer – total weights. Normalized values are then obtained by element-wise division of the first buffer by the second. The first pair of buffers is used to interpolate every patch from within the search window after transforming it to its corresponding reference patch. The second pair of buffers is used to aggregate these transformed patches in order to denoise their reference patch. And, finally, the third pair of buffers aggregates color estimates for every pixel of the output image from different denoised patches.

TABLE I
AVERAGE PSNR VALUES OBTAINED FOR A TEST SET OF COLOR IMAGES WHILE VARYING THE r AND R_{max} WITH ALL THE REST PARAMETERS BEING FIXED. THE BEST CONFIGURATION IS IN **BOLD**

$\sigma = 2$				$\sigma = 5$		
r	25	**30**	35		15	**20**	25
R_{max}				R_{max}			
2	45.00	**45.02**	45.01	2	39.52	39.60	39.65
3	44.85	44.85	44.83	**3**	39.68	**39.69**	39.67
4	44.75	44.74	44.72	4	39.67	39.66	39.61

$\sigma = 10$				$\sigma = 20$		
r	20	**25**	30		40	**45**	50
R_{max}				R_{max}			
4	35.92	36.05	36.06	7	32.61	32.64	32.64
5	35.95	**36.07**	36.05	**8**	32.61	**32.66**	32.65
6	35.96	36.06	36.03	9	32.61	32.65	32.64

$\sigma = 30$				$\sigma = 40$		
r	60	**65**	70		85	**90**	95
R_{max}				R_{max}			
12	30.65	30.69	30.69	18	29.26	29.28	29.27
13	30.67	**30.71**	30.68	**19**	29.28	**29.31**	29.29
14	30.66	30.69	30.69	20	29.29	29.28	29.28

TABLE II
PARAMETERS OF THE METHOD CHOSEN FOR DIFFERENT NOISE LEVELS σ

σ	R_{max}	r	w	g
2	2	30	29	9
5	3	20	29	9
10	5	25	31	9
20	8	45	33	13
30	13	65	35	13
40	19	90	35	21

TABLE III
PARAMETERS OF THE METHOD THAT DO NOT DEPEND ON THE NOISE LEVEL

t	b	i	L	H
1	0.35	0.35	30	0.35

V. PARAMETERS SELECTION

The proposed method depends on the following parameters. Below, we explain the criteria to fix them.

- r – the "radius" used in the elliptical patch calculation (4) (Section III-C).
- R_{max} – the maximum size constraint for elliptical patches. It is specified in terms of the radius of a patch, shall it appear in a wide and completely homogeneous region (Section III-E).
- t – the "scale" derived from the theoretical framework. It controls the intra-patch Gaussian weights (Section III-D).
- w – the size of the search window around a reference patch (Section IV-A).
- g – the resolution of the regular grid used for interpolation during the patch distance calculation (Section IV-B).
- b – the bandwidth multiplier (Section IV-A).
- i – the interpolation coefficient used in the aggregation step (Section IV-B).
- L – the number of most similar patches to be considered in the homogeneous region test (Section IV-A).
- H – the decision threshold of the homogeneous region test (Section IV-A).
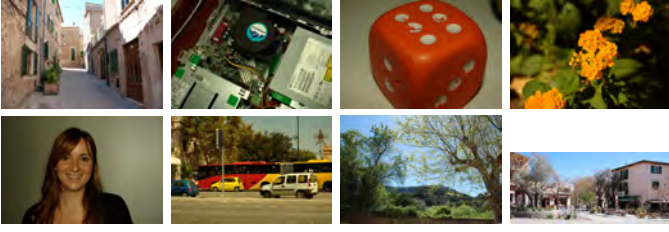
Fig. 10. Noiseless images used in the experiments: Alley, Computer, Dice, Flowers, Girl, Traffic, Trees and Valldemossa.

Fig. 11. "Method noise" benchmark for two noisy images (first row – Computer, second row – Traffic) with added Gaussian white noise of standard deviation $\sigma = 5$. From left to right: "method noise" for BM3D, Non-Local Bayes, original Non-Local Means and for our method. For the visualization purposes, the difference values were scaled from the range of $[-4\sigma, 4\sigma]$ to the range of $[0, 255]$. Consider observing this figure on screen and with zoom.

Let us first remark how the scale is specified in (10) using the t parameter. Instead of making the sigma of the Gaussian weighting function $g_t(h)$ to be equal to t directly, with a slight abuse of notation, we define it as $\sigma_t = \frac{r}{t}$. This allows us to abstract from the r in the scale specification. Therefore, t can be intuitively seen as the number of sigmas that would fit inside a patch.

Since it would be impractical to test all possible combinations of all the parameters, we split them into several groups and adjust one group at a time. We measure performance of the method on each configuration of parameters by the average PSNR value, computed for the test set of eight color images shown in Fig. 10.

The first group of parameters includes r, R_{max} and t which control the size of elliptical patches. Some of the PSNR values obtained while varying the r and R_{max} parameters are shown in Table I as an example. Let us remark that only a few values around the maximum are shown for every level of noise, however, much wider ranges were actually tested.

The second group of parameters includes w and g which as well should be picked for every noise level separately. The size of the search window w should be bigger for higher levels of noise to provide more patch candidates. The value of g should roughly follow the value of R_{max} to ensure enough resolution of the regular grid to represent normalized patches. These intuitions are confirmed while testing the ranges of values for these two parameters.

The third group of parameters includes b and i. In contrast to the original work of [6] we have not observed any significant effect of varying the b parameter for different noise levels. The value of i has no relations with the noise level, therefore a single value, was picked for all the experiments.

The fourth group of parameters includes L and H which are associated with the homogeneous region test. The same values were picked for all levels of noise.

The values of parameters selected for different levels of noise are summarized in Table II. The values of parameters that do not depend on the noise level, and thus fixed for all our experiments, are summarized in Table III.

VI. EXPERIMENTAL RESULTS AND ASSESSMENT

In this section we present assessment of the proposed method and compare it with the original Non-local Means method [6], and the state-of-the-art BM3D [8] and Non-local Bayes [18] methods. In order to assess the performance of the proposed method, we demonstrate the "method noise"
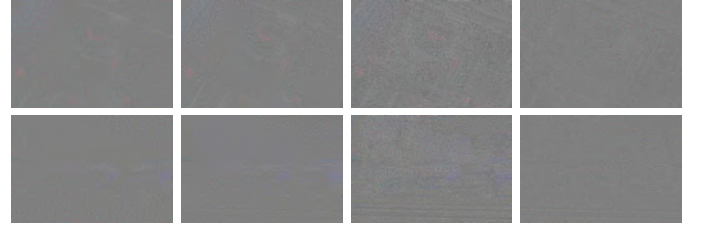
and the "noise to noise" benchmarks commonly used in the literature. For a quantitative evaluation, we show the PSNR values computed for multiple test images and different values of the standard deviation of noise. Finally, we show some of the denoised images for a qualitative assessment.

To obtain the results of Non-Local Means and Non-Local Bayes we use public implementations made by the authors [7], [21] and available online at http://www.ipol.im/. For the results of BM3D we use the implementation made by the authors and publicly available on the dedicated web page[1]. For PSNR evaluation, visual comparison and some other experiments we use the set of eight color images shown in Fig. 10, available online at http://www.ipol.im/ under a Creative Commons CC-BY license. Images Alley and Valldemossa by A. Buades, the rest six images by M. Colom.

A. "Method Noise" and "Noise to Noise" Assessments

At first, we assess the proposed method by calculating the so called "method noise" [6]. For that, we calculate the difference between a noisy image and an output of a denoising method. The "method noise" should contain as little structure from a noisy image as possible. Fig. 11 shows the comparison between BM3D, Non-Local Bayes, the original Non-Local Means and our proposed method. As can be seen, for the proposed method the "method noise" looks almost like a white noise and is very close to the one of BM3D and Non-Local Bayes methods. In contrast the original Non-Local Means method suppresses strong edges (e.g. around the cooler on the first images and along the bus on the second image) which then appear on the difference image.

According to the "noise to noise" principle [19], a denoising method should transform white noise into white noise. If in contrast a method creates some structure from that noise, it will introduce similar artifacts in the denoising result. For this benchmark a uniform image with color values $(127, 128, 129)$ is used as an input. White noise of standard deviation 30 is added separately to every channel of that image and the obtained noisy image is then processed by a denoising method. For visualization purposes the resulting denoised images are amplified: for every pixel the difference to the mean color $(128, 128, 128)$ is magnified by the factor of 5 (see [19] for
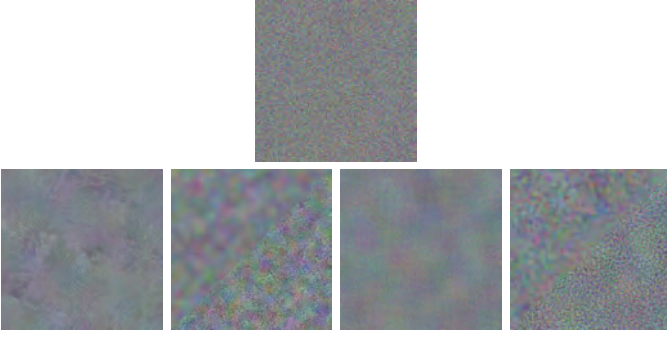
[1]http://www.cs.tut.fi/~foi/GCF-BM3D/index.html

Fig. 12. "Noise to noise" benchmark. First row: an image of uniform color $(127, 128, 129)$ with added Gaussian noise of standard deviation $\sigma = 30$. Second row, from left to right: results of applying BM3D, Non-Local Bayes, original Non-Local Means and our proposed method. Note that for the Non-Local Bayes and for our method, top-right half was processed as homogeneous region (as it should be) while bottom-left as non-homogeneous (this might happen if the homogeneity criterion is not accurate).

TABLE IV
PSNR VALUES FOR NOISE $\sigma = 2, 5$ AND 10

$\sigma = 2$				
	BM3D	NL-Bayes	NL-means	**our**
Alley	45.03	**45.28**	42.68	43.37
Computer	45.30	**45.81**	43.93	44.67
Dice	48.89	**49.17**	48.12	48.22
Flowers	47.40	**47.75**	46.31	46.89
Girl	47.35	**47.67**	46.71	46.71
Traffic	44.66	**45.17**	43.45	44.00
Trees	43.13	**43.44**	42.15	42.62
Valldemossa	44.81	**45.07**	43.26	43.67
$\sigma = 5$				
	BM3D	NL-Bayes	NL-means	**our**
Alley	39.01	**39.14**	37.25	37.31
Computer	40.10	**40.54**	38.93	39.15
Dice	45.98	**46.02**	44.93	45.22
Flowers	43.16	**43.29**	42.17	42.74
Girl	44.09	**44.18**	43.36	43.37
Traffic	38.75	**39.39**	37.59	38.01
Trees	36.11	**36.54**	34.71	35.03
Valldemossa	38.37	**38.62**	35.96	36.70
$\sigma = 10$				
	BM3D	NL-Bayes	NL-means	**our**
Alley	**34.89**	34.82	33.64	33.55
Computer	36.44	**36.68**	35.54	35.40
Dice	**43.40**	43.20	41.92	42.63
Flowers	**39.66**	39.53	38.59	39.38
Girl	**41.57**	41.43	40.40	40.85
Traffic	34.61	**35.15**	34.05	34.10
Trees	31.24	**31.70**	29.59	30.31
Valldemossa	33.81	**33.96**	32.15	32.33

TABLE V
PSNR VALUES FOR NOISE $\sigma = 20, 30$ AND 40

$\sigma = 20$				
	BM3D	NL-Bayes	NL-means	**our**
Alley	**31.23**	31.17	29.98	30.15
Computer	32.86	**32.98**	31.67	32.03
Dice	**40.37**	40.17	38.31	39.62
Flowers	36.06	**36.14**	34.53	36.08
Girl	**38.98**	38.62	36.92	38.02
Traffic	30.85	**31.24**	30.14	30.46
Trees	26.87	**27.36**	26.37	26.40
Valldemossa	29.57	**29.72**	28.44	28.51
$\sigma = 30$				
	BM3D	NL-Bayes	NL-means	**our**
Alley	**29.30**	29.15	27.85	28.37
Computer	**30.71**	30.68	29.28	30.06
Dice	**38.46**	37.95	36.92	37.54
Flowers	33.89	33.85	32.35	**34.11**
Girl	**37.43**	36.69	35.58	36.25
Traffic	28.80	**29.03**	27.74	28.54
Trees	24.56	**25.03**	23.79	24.31
Valldemossa	27.23	**27.35**	25.89	26.46
$\sigma = 40$				
	BM3D	NL-Bayes	NL-means	**our**
Alley	**27.79**	27.77	26.48	27.11
Computer	28.71	**29.04**	27.61	28.57
Dice	36.00	**36.24**	35.26	36.05
Flowers	32.06	32.13	30.58	**32.70**
Girl	**35.25**	35.06	34.17	35.00
Traffic	27.27	**27.52**	26.23	27.17
Trees	23.03	**23.50**	22.42	22.88
Valldemossa	25.65	**25.81**	24.46	25.04

the complete methodology). Fig. 12 shows the "noise to noise" benchmark for BM3D, Non-Local Bayes, the original Non-Local Means and our proposed method. For Non-Local Bayes and for our method the noisy image is treated as homogeneous and also as non-homogeneous. As can be seen, the proposed method leaves no structural artifacts in both homogeneous and non-homogeneous cases. This is also the case for the Non-Local Bayes method as long as the homogeneity threshold γ is set correctly. Meanwhile, the artifacts appear when the image is treated as non-homogeneous. Non-Local Means method leaves some occasional artifacts without any clear structure. Finally, BM3D method produces noticeable characteristic artifacts which might also be observed in the denoising results.

B. PSNR and Visual Comparison

Mean square error (MSE) and peak signal-to-noise ratio (PSNR) are commonly used in the denoising literature to quantify a relative performance of different denoising methods. Tables IV and V show PSNR values for denoising results obtained for the eight color images of Fig. 10 and standard deviation of noise $\sigma \in \{2, 5, 10, 20, 30, 40\}$. Our method almost always outperforms the original Non-Local Means method in terms of PSNR. As expected, for high levels of noise, PSNR values for our method are not very far from the ones of the state-of-the-art Non-Local Bayes and BM3D methods.

Even though comparison by PSNR provides a handy objective metric, an important evaluation is visual comparison that for now can only be done by a human being. Fig. 13 and 14 show several denoising results for high levels of noise. To better illustrate the comparison, the images are zoomed around some interesting regions. As can be seen, the original Non-Local Means suppresses small image details and thus produces "flattened" results. This is especially noticeable on the images Alley, Girl and Flowers. The Non-Local Bayes, while providing high PSNR values, may introduce abrupt jumps in colors in regions where a smooth transition should take place. Appearance of these new edges leads to a "staircasing" effect that is most visible on the images Dice, Girl and Flowers. It is also affected by the noise halo (*rare patch effect* [9]) around high contrasted edges which can be observed on the images Traffic and Computer. While providing good overall quality of the results, the BM3D method may introduce
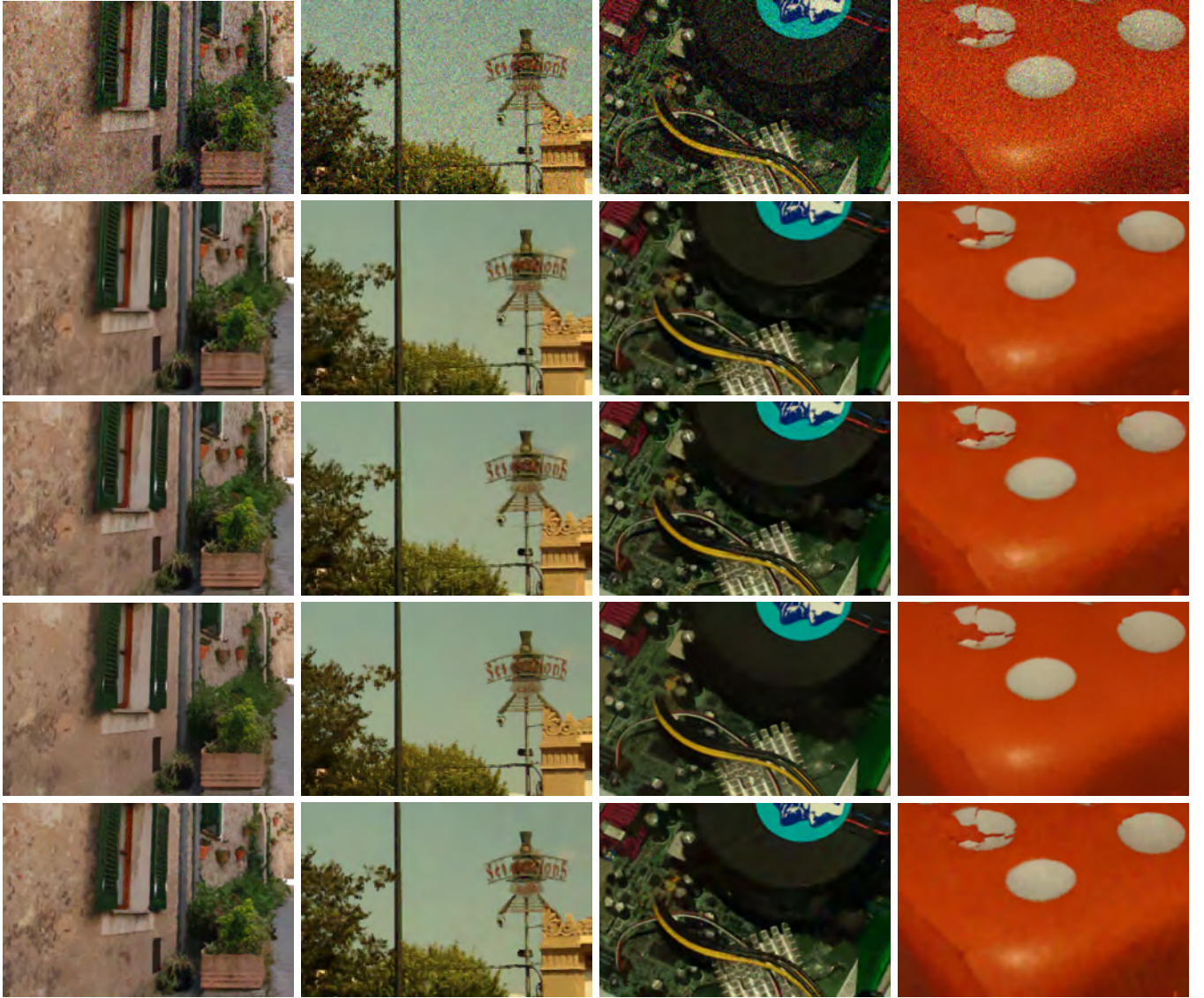
Fig. 13. Denoising results. In rows: noisy image, BM3D, NL-Bayes, NL-Means, our method. In columns: Alley (noise $\sigma = 20$), Traffic (noise $\sigma = 20$), Computer (noise $\sigma = 30$), Dice (noise $\sigma = 30$).

characteristic structural artifacts predicted by the "noise to noise" benchmark. They are most noticeable on the images Trees and Valldemossa. Our method is capable of denoising both small image details, sharp edges and smooth color transitions; however, results may suffer from oversmoothing in some particular cases.

In terms of computational time, our proposed method is on average one order of magnitude slower than the original Non-Local Means method. Partially this is explained by the fact, that there might be several dominant orientations associated with every shape-adaptive patch. For a pair of patches all combinations of orientations should be compared in order to find the smallest patch distance. In our experiments we limit the amount of orientations to 3 for every patch, which leads to at most 9 combinations for a pair of them. Additional computational cost is caused by the interpolation that has to be performed when one shape-adaptive patch is transformed

to be registered with another during image synthesis.

VII. CONCLUSIONS

In this paper we have presented an extension for the Non-Local Means denoising method that considers shape-adaptive patches instead of the very common square ones. The core of our approach is a recently proposed affine invariant patch similarity measure which allows us to compare patches related by a local affinity. By extending the space of patches being considered, this similarity measure helps to detect more similar patches. This in turn leads to better denoising results. Moreover, shape-adaptive patches are well-suited for denoising along contrasted edges and thus our method do not suffer from the halo artifacts. We have shown that our method almost always outperforms the original Non-Local Means both quantitatively and qualitatively. Furthermore, while showing slightly smaller PSNR values than BM3D and Non-Local

Fig. 14. Denoising results. In rows: noisy image, BM3D, NL-Bayes, NL-Means, our method. In columns: Trees (noise $\sigma = 30$), Valldemossa (noise $\sigma = 30$), Girl (noise $\sigma = 40$), Flowers (noise $\sigma = 40$).

Bayes, the proposed method provides the state-of-the-art level of visual quality of the results. We find these observations to be promising and consider similar extensions of Non-Local Bayes and BM3D methods as a possible direction for future research.

Further work will be devoted to the better suppression of the lower frequency noise that remains in homogeneous regions for the noise levels of $\sigma = 30$ and higher. Multiscale and two-step denoising are among the options that will be tested aiming to improve the results.

ACKNOWLEDGMENTS

REFERENCES

[1] P. Arias, G. Facciolo, V. Caselles, and G. Sapiro. A variational framework for exemplar-based image inpainting. *International Journal of Computer Vision*, 93:319–347, 2011. http://dx.doi.org/10.1007/s11263-010-0418-7.

[2] G. Aubert and J. F. Aujol. A variational approach to removing multiplicative noise. *SIAM Journal on Applied Mathematics*, 68(4):925–946, 2008.

[3] C. Ballester, F. Calderero, V. Caselles, and G. Facciolo. Multiscale analysis of similarities between images on riemannian manifolds. *SIAM Journal Multiscale Modeling and Simulation*, 12(2):616–649, 2014.

[4] K. Bredies, K. Kunisch, and T. Pock. Total generalized variation. *SIAM Journal on Imaging Sciences*, 3(3):492–526, 2010.

[5] T. Brox, O. Kleinschmidt, and D. Cremers. Efficient nonlocal means for denoising of textural patterns. *IEEE Transactions on Image Processing*, 17(7):1083–1092, July 2008.

[6] A. Buades, B. Coll, and J.-M. Morel. A non local algorithm for image denoising. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, volume 2, pages 60–65, 2005. http://dx.doi.org/10.1109/CVPR.2005.38.

[7] Antoni Buades, Bartomeu Coll, and Jean-Michel Morel. Non-Local Means Denoising. *Image Processing On Line*, 1, 2011.

[8] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian. Image denoising by sparse 3-d transform-domain collaborative filtering. *Image Processing, IEEE Transactions on*, 16(8):2080–2095, Aug 2007.

[9] C. A. Deledalle, V. Duval, and J. Salmon. Non-local methods with shape-adaptive patches (nlm-sap). *Journal of Mathematical Imaging and Vision*, 43(2):103–120, 2012.

[10] D. L. Donoho and J. M. Johnstone. Ideal spatial adaptation by wavelet shrinkage. *Biometrika*, 81(3):425–455, 1994.

[11] M. Elad and M. Aharon. Image denoising via sparse and redundant representations over learned dictionaries. *Image Processing, IEEE Transactions on*, 15(12):3736–3745, Dec 2006.

[12] V. Fedorov, P. Arias, G. Facciolo, and Ballester. Affine invariant self-similarity for exemplar-based inpainting. In *In Proc. Int. Conf. Computer Vision Theory and App. VISAPP-2016*, Rome, Italy, 2016.

[13] V. Fedorov, P. Arias, R. Sadek, G. Facciolo, and C. Ballester. Linear multiscale analysis of similarities between images on riemannian manifolds: Practical formula and affine covariant metrics. *SIAM Journal on Imaging Sciences*, 8(3):2021–2069, September 2015.

[14] A. Foi and G. Boracchi. Foveated self-similarity in nonlocal image filtering. *Human Vision and Electronic Imaging XVII*, 8291(1):829110, February 2012.

[15] G. Gilboa and S. J. Osher. Nonlocal operators with applications to image processing. *Multiscale Modeling and Simulation*, 7(3):1005–1028, 2008.

[16] Sven Grewenig, Sebastian Zimmer, and Joachim Weickert. Rotationally invariant similarity measures for nonlocal image denoising. *Journal of Visual Communication and Image Representation*, 22(2):117 – 130, 2011.

[17] A. Kheradmand and P. Milanfar. A general framework for regularized, similarity-based image restoration. *IEEE Trans. on Image Processing*, 23(12):5136–5151, 2014.

[18] M. Lebrun, A. Buades, and J. M. Morel. A nonlocal bayesian image denoising algorithm. *SIAM Journal on Imaging Sciences*, 6(3):1665–1688, 2013.

[19] M. Lebrun, M. Colom, A. Buades, and J. M. Morel. Secrets of image denoising cuisine. *Acta Numerica*, 21:475–576, 5 2012.

[20] M. Lebrun, M. Colom, and J.-M. Morel. Multiscale image blind denoising. *Image Processing, IEEE Transactions on*, 24(10):3149–3161, Oct 2015.

[21] Marc Lebrun, Antoni Buades, and Jean-Michel Morel. Implementation of the "Non-Local Bayes" (NL-Bayes) Image Denoising Algorithm. *Image Processing On Line*, 3:1–42, 2013.

[22] D. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2):91–110, 2004.

[23] J. Mairal, M. Elad, and G. Sapiro. Sparse representation for color image restoration. *Image Processing, IEEE Transactions on*, 17(1):53–69, Jan 2008.

[24] A. Martn, J. F. Garamendi, and E. Schiavi. Iterated rician denoising. In *In Proc. Int. Conf. Image Process. Comput. Vis. Pattern Recognit. IPCV-2011*, volume 11, pages 959–963, Las Vegas, USA, 2011.

[25] E. A. Nadaraya. On estimating regression. *Theory of Probability & Its Applications*, 9(1):141–142, 1964.

[26] M Nikolova. A variational approach to remove outliers and impulse noise. *Journal of Mathematical Imaging and Vision*, 20(1-2):99–120, 2004.

[27] P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 12(7):629–639, Jul 1990.

[28] G. Peyré. Manifold models for signals and images. *Computer Vision and Image Understanding*, 113(2):249–260, 2009. http://dx.doi.org/10.1016/j.cviu.2008.09.003.

[29] G. Peyré, S. Bougleux, and L. D. Cohen. Non-local regularization of inverse problems. In *Proceedings of the IEEE Conference on Computer Vision ECCV*, pages 57–68, 2008.

[30] N. Pierazzo, Rais Lebrun, M., M. E., J. M. Morel, and G. Facciolo. Non-local dual image denoising. In *In Image Processing (ICIP), 2014 IEEE International Conference on (pp. 813-817). IEEE.*, 2014.

[31] L. Pizarro, P. Mrázek, S. Didas, S. Grewenig, and J. Weickert. Generalised nonlocal image smoothing. *IJCV*, 90:62–87, 2010.

[32] J. Portilla, V. Strela, M.J. Wainwright, and E.P. Simoncelli. Image denoising using scale mixtures of gaussians in the wavelet domain. *Image Processing, IEEE Transactions on*, 12(11):1338–1351, Nov 2003.

[33] M. Protter, M. Elad, H. Takeda, and P. Milanfar. Generalizing the non-local-means to super-resolution reconstruction. *IEEE Trans. on IP*, 18(1):36–51, 2009.

[34] Leonid I. Rudin, Stanley Osher, and Emad Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, 60(14):259 – 268, 1992.

[35] J. L. Starck, E. J. Candès, and D. L. Donoho. The curvelet transform for image denoising. *IEEE Trans. on Image Processing*, 11(6):670–684, 2002.

[36] C. M. Stein. Estimation of the mean of a multivariate normal distribution. *The annals of Statistics*, 43(2):1135–1151, 1981.

[37] Geoffrey S. Watson. Smooth regression analysis. *Sankhya: The Indian Journal of Statistics, Series A (1961-2002)*, 26(4):359–372, 1964.

[38] Leonid P. Yaroslavsky. Local adaptive image restoration and enhancement with the use of dft and dct in a running window, 1996.

[39] Guoshen Yu and Guillermo Sapiro. Dct image denoising: a simple and effective image denoising algorithm. *Image Processing On Line*, 1, 2011.

[40] Lei Zhang, Weisheng Dong, David Zhang, and Guangming Shi. Two-stage image denoising by principal component analysis with local pixel grouping. *Pattern Recognition*, 43(4):1531 – 1549, 2010.

[41] S. Zimmer, S. Didas, and J. Weickert. A rotationally invariant block matching strategy improving image denoising with non-local means. In *International Workshop on Local and Non-Local Approximation in Image Processing*, Lausanne, Switzerland, August 2008.

[42] C. Zuo, L. Jovanov, H.Q. Luong, B. Goossens, W. Philips, Yu Liu, and M. Zhang. Rotation invariant similarity measure for non-local self-similarity based image denoising. In *Image Processing (ICIP), 2015 IEEE International Conference on*, pages 1618–1622, Sept 2015.