

Adaptive Image Decomposition into Cartoon and Texture Parts Optimized by the Orthogonality Criterion

D. Szolgay, *Student Member, IEEE* and T. Szirányi, *Senior Member, IEEE*

Abstract—In this paper a new decomposition method is introduced that splits the image into geometric (or cartoon) and texture parts. Following a total variation based preprocessing, the core of the proposed method is an anisotropic diffusion with an orthogonality based parameter estimation and stopping condition. The quality criterion is defined by the theoretical assumption that the cartoon and the texture components of an image should be orthogonal to each other. The presented method has been compared to other decomposition algorithms through visual and numerical evaluation to prove its superiority.

Index Terms—image decomposition, texture segmentation, Total Variation, Anisotropic Diffusion, quality criterion

I. INTRODUCTION

Image decomposition into meaningful components has a key role in many image processing applications. In this paper, we focus on decomposition into texture and non-texture (or cartoon) components. This kind of image decomposition can be useful for image compression where compressing the cartoon and the texture components separately can provide better results [1], for image denoising [2], [3] since zero mean oscillatory noise can be regarded as a fine texture, image feature selection [4], 2D and 3D computer graphics and main edge detection as illustrated in [5], etc.

Recently published algorithms for texture/cartoon decomposition [4]–[8] are mostly based on total variation (TV) minimization inspired by the work of Yves Meyer [9]. Total variation based regularization dates back to Tikhonov [10]. The most widely known form was introduced in image processing by Mumford and Shah [11] for image segmentation and later by Rudin et al. [3] for noise removal through the optimization of a cost function as follows:

$$\inf \left\{ E_{TV}(u) = \int_{\Omega} |Du| + \lambda \int_{\Omega} v^2 \right\} \quad (1)$$

where u is the cartoon component of the original image f , $v = f - u$ is the texture, $\int_{\Omega} |Du|$ denotes the total variation of u in Ω and λ is a regularization parameter. The first part produces a smooth image with bounded variation upon energy minimization for the cartoon component, while the second ensures that the result is close to the initial image.

Dániel Szolgay is with Pázmány Péter Catholic University, Budapest, Hungary (e-mail: szolgay.daniel@itk.ppke.hu)

Tamás Szirányi is with Computer and Automation Research Institute of the Hungarian Academy of Sciences, Distributed Events Analysis Research Group, H-1111 Budapest, Kende utca 13-17, Hungary (e-mail: sziranyi@sztaki.hu)

The regularization of Rudin et al. [3] (**ROF** in the following) was used as an image denoising and deblurring method, since it removes fine, oscillating, noise-like patterns, but preserves sharp edges.

In [9] Meyer proposed a different norm for the second, texture part of (1), which is better suited for oscillatory components than the standard L_2 norm:

$$\inf \int_{\Omega} |Du| + \lambda \|v\|_* \quad (2)$$

where $\|\cdot\|_*$ is defined on a suitably defined Banach G space as follows:

$$\|v\|_* = \inf_{g_1, g_2} \left\| \sqrt{g_1^2(x, y) + g_2^2(x, y)} \right\|_{L^\infty} \quad (3)$$

over all g_1 and g_2 such that $v = \text{div}(\vec{g})$ where $\vec{g} = (g_1, g_2)$.

Other variations of eq.(1) are summarized in [5].

Beside the choice of regularization, other techniques are used to enhance the quality of the decomposition: in [7], the authors propose an image decomposition and texture segmentation method via sparse representation using Principal Component Analysis (**DPCA**). In [8], an algorithm (**DOSV**) is introduced to find the optimal value of the fidelity parameter (λ in eq.(1)) based on the observation of Aujol et al. in [12] concerning the independence of cartoon and texture.

Looking at the palette of the different solutions, we can see that the decomposition into cartoon and textured partitions requires tackling the following main issues:

- Adaptive scale definition of texture and cartoon (cc. outline) details;
- Reasonable process that filters out textured parts while keeping the main outlines;
- Quality criterion for the efficiency of the decomposition: goal function of the process.

Some aspects of these issues have been addressed in [13] where preliminary results were introduced. In the following, we overview the related contributions and then we discuss our proposed solutions to the above issues.

A. Related Works

In this section, we shortly summarize published results closely related to the proposed method: non-linear filtering is introduced in [5], Anisotropic Diffusion in [14]–[16] and measures of independence in [12], [17], [18].

1) *BLMV Nonlinear Filter*: Buades et al. have recently proposed a non-linear method inspired by eq.(2) (*BLMV* filter in the following) that calculates local total variation (LTV) for each pixel on f before and after filtering the image with a σ sized low pass filter, L_σ , inspired by Y. Meyer [9]. The relative difference of the calculated LTVs shows if the observed pixel is part of the texture or the cartoon, since the oscillatory parts' LTV will change radically, while the LTV of the cartoon parts will be left virtually unchanged (although blurred). The cartoon image, u is composed based on this information: if the relative difference is high for a pixel r , then $u(r)$ will be equal to the low pass filtered $(L_\sigma * f)(r)$, otherwise $u(r) = f(r)$. The results of this simple method are impressive on the presented examples in [5]: the edges are preserved as long as σ is not too large and the texture components are blurred with L_σ .

The right choice of σ is important to get the best result, however, it is possible that there is no such σ which eliminates all the textures but keeps the non-texture components on the cartoon. The existence of a content adaptive scaling parameter can be derived from scale-space theory, as it has been introduced in the work of Lindeberg [19].

2) *Anisotropic Diffusion*: The general goal of diffusion algorithms is to remove noise from an image by using partial differential equations (PDE). Diffusion algorithms can be classified as isotropic or anisotropic. Isotropic diffusion can be described by the following equation:

$$\frac{\partial f(x, y, t)}{\partial t} = \nabla^2 \cdot f \quad (4)$$

where $f(x, y, t) : \mathbb{R}^2 \rightarrow \mathbb{R}^+$ is the image in the continuous domain, with (x, y) spatial coordinates, t an artificial time parameter and ∇f the image gradient. $f(x, y, 0)$ is the original image. This diffusion is equivalent to using a Gaussian filter on the image, which blurs not only the noise or texture components, but the main edges as well.

In [20] Gábor and later in [14] Perona and Malik proposed anisotropic diffusion (*AD*) functions that, according to scale-space theory (see works of Florack [21] or Alvarez, Lions and Morel [22]) allows diffusion along the edges or in edge-free territories, but penalizes diffusion orthogonal to the edge direction:

$$\frac{\partial f(x, y, t)}{\partial t} = \nabla \cdot (g(\|\nabla f\|) \nabla f) \quad (5)$$

where $\|\nabla f\|$ is the magnitude of the gradient and $g(\cdot)$ is the weighting function that controls diffusion along and across edges. The discretized form of their diffusion equation is as follows:

$$f(x, y, t + 1) = f(x, y, t) + \frac{\lambda}{|\eta(x, y)|} \sum_{(x', y') \in \eta(x, y)} \nabla^{(x', y')} \left(g \left(\left\| \nabla^{(x', y')} f(x, y, t) \right\| \right) \right) \nabla^{(x', y')} f(x, y, t) \quad (6)$$

where I is the processed image, (x, y) is a pixel position, t now denotes discrete time steps (iterations). The constant $\lambda \in \mathbb{R}^+$ is a scalar that determines the rate of diffusion, $\eta(x, y)$ is the spatial neighborhood of (x, y) , $|\eta(x, y)|$ is the number of

neighboring pixels. $\nabla^{(x', y')} f(x, y, t)$ is an approximation of the image gradient at a particular direction:

$$\nabla^{(x', y')} f(x, y, t) = f(x', y', t) - f(x, y, t), (x', y') \in \eta(x, y) \quad (7)$$

AD belongs to a theoretically sound scale-space class of differential processes ensuring the denoising of an image along with the enhancement of its main structure [23]. We will show that the AD as proposed by Perona and Malik is not suitable for cartoon/texture decomposition, since the texture part might contain high magnitude edges, which would inhibit the diffusion. As a solution to this problem, the authors of [1] suggest that the AD algorithm is used with modified weights: instead of using $\left\| \nabla^{(x', y')} f(x, y, t) \right\|$ as the parameter of the weighting function, they use the edges of the Gaussian filtered image, $\nabla(G_\sigma * f)$:

$$\left\| \nabla^{(x', y')} (G_\sigma * f)(x', y', 0) - (G_\sigma * f)(x, y, 0) \right\|, (x', y') \in \eta(x, y) \quad (8)$$

where G_σ is a Gaussian filter with σ bandwidth and $*$ denotes the convolution. Using a blurred image to control diffusion directions will give better results: texture edges will not hinder the diffusion, but the strong main edges will. Yet the quality of the solution relies heavily on the σ parameter: with small σ , some texture might remain on the cartoon, while with greater σ , some of the cartoon edges will disappear.

In Section II we will propose an algorithm which utilizes the smoothing property of AD, while it preserves edges based on whether they belong to a cartoon or texture and not their level of magnitude.

3) *The Use of Independence in Image Decomposition*: The independence of the cartoon part and the texture/noise part of the image was used in denoising, decomposition [12] and restoration [18] algorithms.

In [12], Aujol et al. propose the use of the correlation between the cartoon and the oscillatory (noise, texture) components of a decomposition to estimate the regularization parameter λ . The assumption made in their model is that these two components are uncorrelated, which makes intuitive sense (as stated in [8]), since every feature of an image should be considered as either a cartoon feature or a textural/noise feature, but not both.

In [18], the Angle Deviation Error (*ADE*) - introduced in [17] - is used as a measure of independence to automatically find the best stopping point for an iterative non-regularized image deconvolution method. As the deconvolution problem is ill-posed, after a certain point, further iterations will only amplify the noise on the estimated image. The heart of the method in [18] is to find the iteration, where the change of the estimated image in one time step $X(t) - X(t - 1)$, and the estimated image itself ($X(t)$) are the most independent of each other. The described *ADE* measure is somewhat similar to correlation [12], but it is based on the pure orthogonality of two image partitions (e.g. clear image and noise):

$$ADE(Q, P) = \left| \arcsin \left(\frac{\langle Q, P \rangle}{|Q| \cdot |P|} \right) \right| \quad (9)$$

where $Q, P \in \mathbb{R}^n$ and $\langle Q, P \rangle$ is their scalar product. This measure is different from the standard correlation, where zero-mean vectors are used to calculate the scalar product and the normalization is done with the vectors' standard deviation:

$$\text{corr}(Q, P) = \frac{\text{cov}(Q, P)}{\sigma_Q \cdot \sigma_P} = \frac{\sum_{i=1}^n (Q_i - \mu_Q)(P_i - \mu_P)}{n \cdot \sigma_Q \cdot \sigma_P} \quad (10)$$

where $\text{cov}(\cdot)$ is the covariance over the elements of vectors, σ_Q, μ_Q and σ_P, μ_P are the standard deviation and the expected values of the elements of Q and P respectively, and n is the size of the vectors.

Comparing the two measures, we can see that they are very similar: if both Q and P were zero mean, the two measures would actually give the same result. However, in cartoon texture decomposition only the texture part has an inherent zero mean, while the cartoon does not. This makes a small difference in the resulting decomposed images in favor of the *ADE* measure, as it will be shown in Section III; *ADE* strengthens the image partitions to being really independent (geometrical orthogonality in \mathbb{R}^n), while *corr* is for the estimation of regression.

B. The Contribution of the Paper

In the following we will show how independence can be used to separate better the texture and cartoon parts of the image by using *ADE* orthogonality measure to locally estimate the best parameter of the BLMV filter. The edge inhibitions of the AD are initialized by the filtered image. Then the *ADE* is calculated again on the diffused image to stop the diffusion at the point where the orthogonality of cartoon and texture components is maximal.

To sum up, we offer theoretically clear solutions for the main issues:

- Adaptive scale definition by using locally optimal BLMV filter tuned by *ADE* measure;
- Anisotropic Diffusion, initialized by the new adaptive BLMV to better separate texture from cartoon;
- Orthogonality criterion for the quality measure of the decomposition (stopping condition to AD).

In Section II, we overview in detail our proposed solutions for the above issues. To validate the proposed method, in Section III, we will show results on real life images, and also on artificial images where numerical evaluation is possible.

II. CARTOON/TEXTURE DECOMPOSITION USING INDEPENDENCE MEASURE

In this section, the orthogonality based cartoon/texture decomposition method is described in detail. The core algorithm is the AD, which is initialized and stopped using BLMV filter and *ADE* independence measure.

A. Locally Adaptive BLMV filter

As it has been mentioned earlier, the BLMV filter uses the same σ sized low pass filter for the whole image, while there is no guarantee for the existence of a single sigma that would

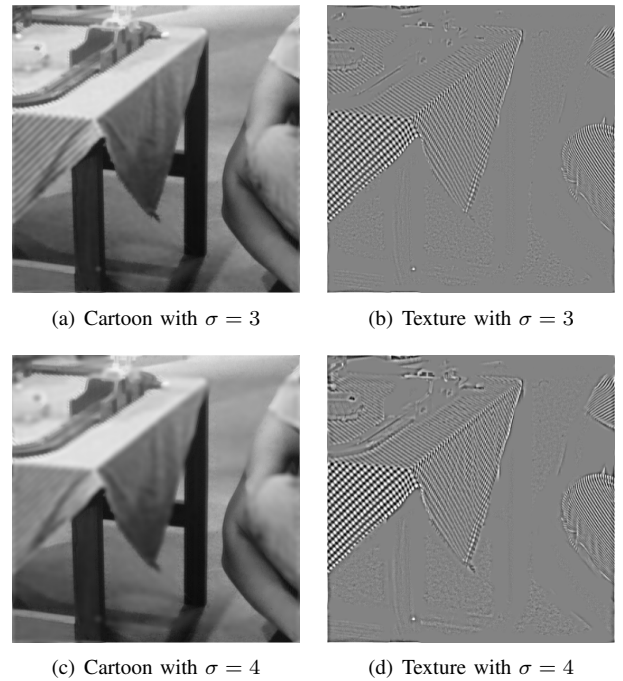


Fig. 1. The cartoon and texture component of a part of the Barbara image (see the original in Fig.6(a)) produced by the BLMV method with $\sigma = 3$ and $\sigma = 4$, respectively. Note that the texture of the tablecloth is not completely removed by the smaller sigma, while the edges of the cover are blurred if we choose a larger sigma that eliminates the texture from the cover.

remove all texture from the image without blurring the cartoon edges (see Fig.1.).

We propose the use of different σ for the different parts of the image based on the independence of the removed texture component and the remaining cartoon component. This theory is similar to the one proposed in [12], although in our case the parameter selection has to be locally adaptive. The reason for this difference lies in the purpose of the methods: while in [12] the authors' goal was noise removal, where one can assume that the parameters of the noise are the same for the whole image, here we want to remove texture components which may vary in many aspects (e.g. scale, magnitude) across the image.

To make the filter locally adaptive, BLMV filtered images were calculated for a given range of the scale parameter: $\sigma_i \in [s_1; s_2]$. Let $u_{\sigma_i}, v_{\sigma_i}$ denote the cartoon and texture components of the f input image, produced by the BLMV filter with σ_i parameter.

The image is then divided into non-overlapping small cells (5 pixel by 5 pixel in our experiments), and around each cell a larger block (21 by 21) is centered, in which the *ADE* measure is calculated:

$$\text{ADE}(u_{\sigma_i}^{(x,y)}, v_{\sigma_i}^{(x,y)}) = \left| \arcsin \left(\frac{\langle u_{\sigma_i}^b(x,y), v_{\sigma_i}^b(x,y) \rangle}{|u_{\sigma_i}^b(x,y)| \cdot |v_{\sigma_i}^b(x,y)|} \right) \right|, \quad (11)$$

where $u^b(x,y)$ and $v^b(x,y)$ denote the cartoon and texture components of the *block* which is centered around the cell containing (x,y) pixel.

It is worth noting that the texture component of an image should have zero mean, since it is the difference of the

textured area and the diffused background. To eliminate the consequences of the quantization error through the iterations, the texture component is biased to be zero mean when the ADE function is computed.

The σ with minimal ADE is chosen to be the parameter for each pixel in the cell. For the output cartoon image the value of the pixel, $u_a(x, y)$, will be the following:

$$u_a(x, y) = u_{\sigma_m^{(x,y)}}(x, y) \quad (12)$$

$$\sigma_m^{(x,y)} = \arg \min_{\sigma_i \in [s_1; s_2]} (ADE(u_{\sigma_i}^b(x, y), v_{\sigma_i}^b(x, y))) \quad (13)$$

This cell-based scheme is used to reduce the computational workload: instead of calculating the block correlation for each pixel, we calculate it for small cells. To avoid the blocking effect, a soft Gaussian smoothing was used on the parameter image of the same size as the input image and containing the corresponding σ value in each (x, y) point: $p(x, y) = \sigma_m^{(x,y)}$. For the smoothing Gaussian $\sigma = 2pixel$ was applied with $2 * (\sigma * 5) + 1$ window size. However the result is not sensitive to these numbers within a reasonable range. An example result of the described method can be seen on Fig.2. and the corresponding parameter image on Fig.3.

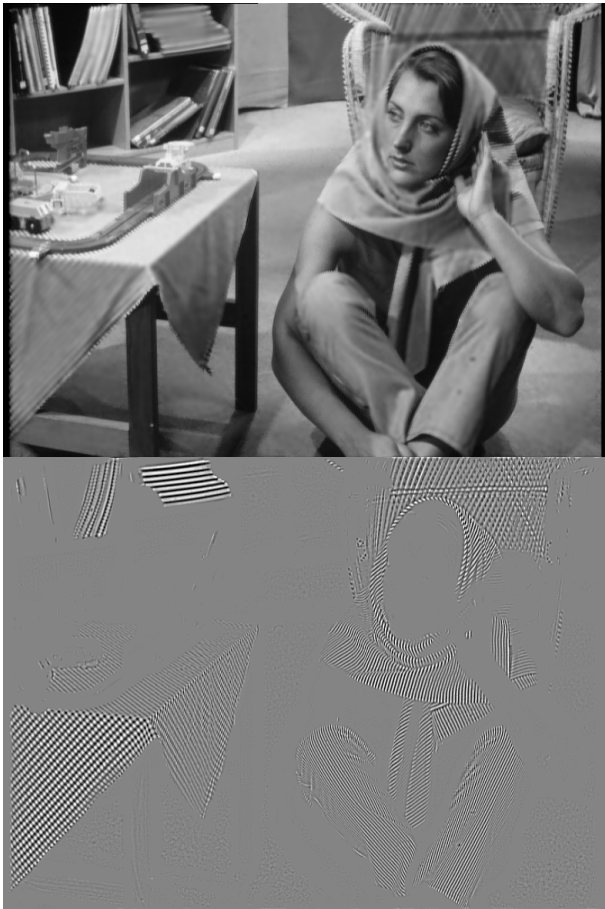


Fig. 2. Cartoon and texture components of the BLMV filtered Barbara image (Fig.6(a)) with automatic σ selection.



Fig. 3. The parameter map of Barbara image (Fig.6(a)).The brighter pixel corresponds to greater value used on that location. In this image the value of σ is between 0 and 5 and it is linearly stretched between 0 and 255 for demonstration.

B. Anisotropic Diffusion with adaptive BLMV filter and ADE stopping condition

The above described adaptive BLMV filter (aBLMV in the following) clearly performs better than the original one (see Section III), but it still faces a problem at the borders where cartoon and texture parts meet: either the cartoon edges are blurred, or the texture remains on the cartoon component close to cartoon edges. We propose to use AD initialized with a cartoon image produced by aBLMV filter and stopped by ADE measure. AD preserves high magnitude edges and blurs weaker ones, but obviously a texture can contain strong edges while a cartoon edge can be weak. As a result, AD may blur important edges of the cartoon and keep unwanted edges of the texture.

Similarly to [1], where the diffusion weight function was calculated on a Gaussian blurred version of the image, we propose to calculate the $g(\cdot)$ weight function of eq.(6) by using the aBLMV-filtered image resulting in the following diffusion equation:

$$f(x, y, t + 1) = f(x, y, t) + \frac{\lambda}{|\eta(x, y)|} \sum_{(x', y') \in \eta(x, y)} \nabla^{(x', y')} \left(g \left(\nabla^{(x', y')} u_a(x, y) \right) \right) \nabla^{(x', y')} f(x, y, t) \quad (14)$$

Note that the aBLMV filter, could be easily replaced in the algorithm by any other method, which blurs the texture but preserves cartoon edges. We tested various methods, like simple Gaussian blur or the linear filter used in [5], and we have found that the aBLMV filter performs the best. TV based methods like *TVLI* [4] and ROF were also tested, but their results were no better than the result of any of the respective methods (TVLI or ROF).

On u_a of eq.(12), the texture parts are blurred and they do not contain strong edges, while the cartoon parts are more or less preserved. Choosing a low value for the rate of diffusion (λ) means that the diffusion can preserve even the weak edges

of the cartoon part, but it blurs texture parts completely (since it is not inhibited by edges). Fig. 4 shows the cartoon and texture components produced by the method proposed above (AD with ADE).

To avoid oversmoothing of important edges, the iteration of the AD must be stopped at the right moment. For this purpose, we utilize the independence property of cartoon and texture components in the same manner as we did in Section II-A, with the difference that here we are searching for the iteration count i that minimizes $ADE(u_i, v_i)$ for each block (the size of the blocks is 21 by 21 pixels).

The cartoon component of the proposed method is produced as follows:

$$u(x, y) = f(x, y, t_{ADE}), \quad (15)$$

$$t_{ADE} = \arg \min_{i=1..I_{max}} (ADE(f^b(x, y, i), v^b(x, y, i))) \quad (16)$$

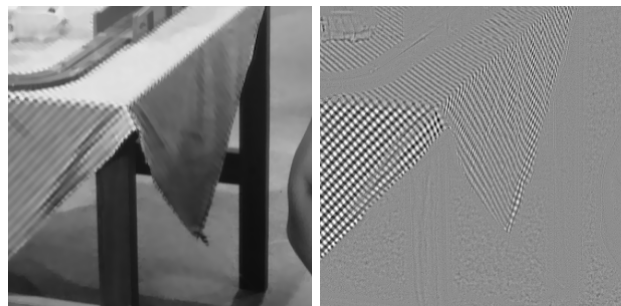
where $f(x, y, t_{ADE})$ is the (x, y) pixel of the diffused image after t_{ADE} iterations, I_{max} is the maximum number of diffusion iterations, $f^b(x, y, i)$ and $v^b(x, y, i) = f^b(x, y, 0) - f^b(x, y, i)$ are the cartoon and texture components, respectively, of the *block* that is centered around the cell containing (x, y) pixel after the i th diffusion iteration.



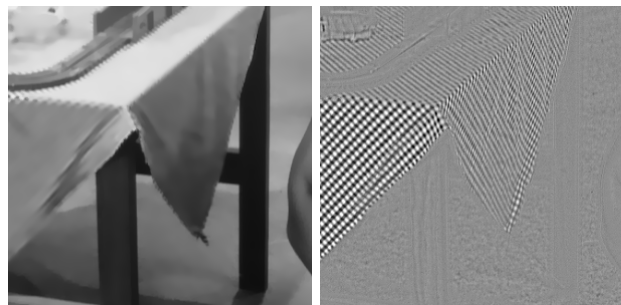
Fig. 4. The cartoon and texture components of the Barbara image produced by the proposed anisotropic diffusion model with ADE based stopping condition.

If the diffusion is not stopped automatically, but after fixed number of iterations, then the texture will not be eliminated

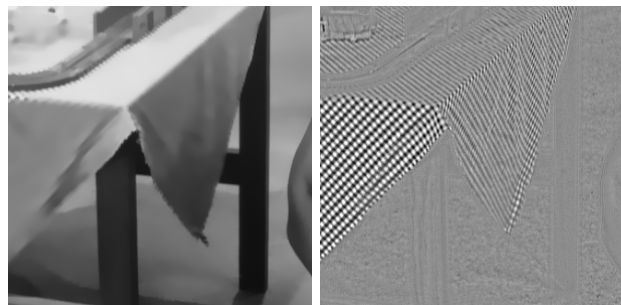
completely from the cartoon or some parts of the cartoon component will be apparent on the texture image, as it can be seen on Fig.5.



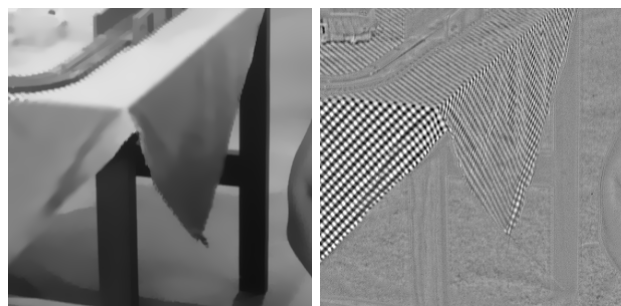
(a) Cartoon and Texture after 10 iterations



(b) Cartoon and Texture after 30 iterations



(c) Cartoon and Texture after 50 iterations



(d) Cartoon and Texture after 100 iterations

Fig. 5. The cartoon and texture components of a part of the Barbara image produced by the proposed anisotropic diffusion model after 10, 30, 50 and 100 iterations.

III. RESULTS

The evaluation of the quality of cartoon/texture decomposition is usually done on visual examples, since there is no generally accepted objective method for ground truth generation in case of real images. Sometimes it is difficult

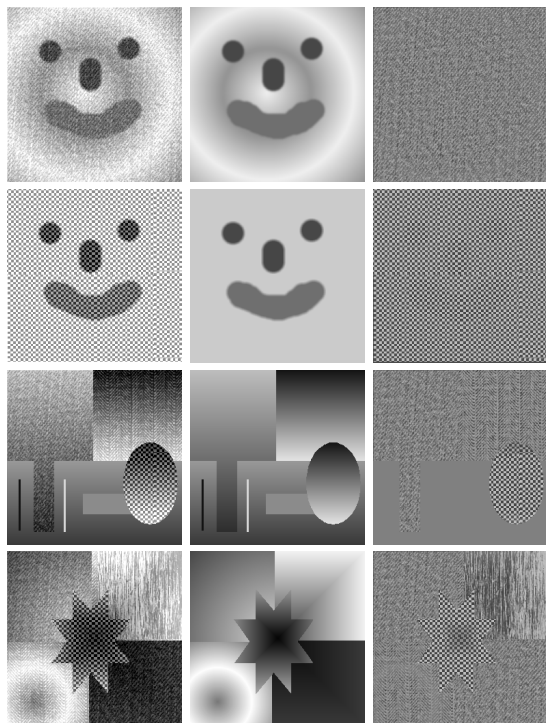


Fig. 7. The artificial images used for numerical evaluation. Left column: original image, Middle column: cartoon component, Right column: texture component.

even for a human to decide if a certain part of the image is texture or not.

Hence, to evaluate the quality of the different methods, we show the decomposition results of example images (see Fig.6), but we also evaluate numerically the competing methods on artificial images (see Fig.7) where the ground truth cartoon and texture parts are available.

We have compared the proposed method to the following decomposition methods: BLMV-filter [5], aBLMV-filter (also proposed in this paper), Anisotropic Diffusion [14], DPCA [7], DOSV [8], ROF [3], TVL1 [4]. The codes for the above methods were provided by the authors, and we used them with the best tuned parameters in each individual test case. For numerical evaluation, we used the parameters that gave the best numbers, and in case of subjective evaluation, the parameters that gave the best visual result. For the proposed method, we kept all the parameters, except one: the σ range for the aBLMV was changed to the same transparent scale parameter as it was for the original BLMV. The other parameters were set to a constant value: the maximum number of iterations for the AD was set to 100, the λ parameter of eq. (6) was set to 2. Note that $\lambda = 2$ is a low value, making the AD very sensitive to edge inhibitions, which helps the better preservation the cartoon edges. The only parameter that was not constant during the tests is the $[s_1, s_2]$ range of the σ . The usual values were $s_1 = 0.5$ and $s_2 = 7$. The only time when the values were different was for the City image, where we set $s_2 = 4$ to preserve better the cartoon details of the image.

For better visibility, the contrast of the texture images was linearly stretched on the demonstrated figures.

A. Visual Evaluation on Real Life Images

For the visual evaluation, one has to consider how strong the remaining cartoon parts on the texture image and the remaining texture part on the cartoon image are. For a part of the Barbara image, we can see on Fig.8 that 5 methods (AD, BLMV, TVL1, ROF, DOSV) cannot completely eliminate the texture from the table cover, while there are cartoon edges apparent on the texture image. DPCA can eliminate the texture from the cartoon image, but the image itself becomes less smooth, and the slow changes of gray level values are also apparent on the texture image (to observe the differences in detail please consider viewing the digital version). The BLMV with adaptive local parameter selection (aBLMV) and the proposed method eliminate the texture from the cartoon while virtually no cartoon appear on the texture image (see Fig.8).

On the Geometry image, all the methods eliminate the texture from the cartoon part, but all of them bring some cartoon edges on the texture (see Fig.9). Here one should consider how strong the cartoon edges on the texture image are, and also how precise the cartoon part is.

The third image shows city towers. This image has precise edges, which favors the TV based methods, especially ROF (see Fig.10). However, some artifacts can be seen on cartoon image of ROF, as the rectangularly shaped cloud at the top of the building on the left, or the disappearing top of the same skyscraper. Most of the methods blur some parts of the image, and almost none of them can eliminate the vertical line texture from the darkest building.

For the fourth image (Pillars), the question is how well the pillars are preserved on the cartoon image (or how strong the edges of the pillars on the texture image are) and how blurred the greenery in the background is (see Fig.11). Here we can say that BLMV, DOSV and TVL1 produce good results, but they are outperformed by aBLMV and the proposed method, while AD and ROF preforms very poorly: the texture is slightly blurred on the cartoon image and the edges of the pillar are already obviously present on the texture part. DPCA blurs the texture the best (similarly to the proposed method), but, in the meantime, it brings some strong cartoon edges to the texture part.

The last image, the Zebra is quite challenging, since the texture of the Zebra has a wide range of sizes. See results on Fig.12. AD and ROF preform poorly, since most of the texture remains on the cartoon, while non-textural parts like slow changes of gray level values and non-textural parts of the background are apparent on the texture image. As it was mentioned earlier, AD is not suited for the tasks, since the edges of the texture are stronger than some of the cartoon edges, therefore the cartoon edges are eliminated while the texture edges are kept unchanged. DOSV cannot eliminate the larger texture parts without blurring the cartoon. BLMV blurs the cartoon even more, but it eliminates most of the texture, although not as efficiently as aBLMV or the proposed method. TVL1 and DPCA perform similarly: they both eliminate most of the texture, but a lot of non-textural edges are apparent on the texture image as well.

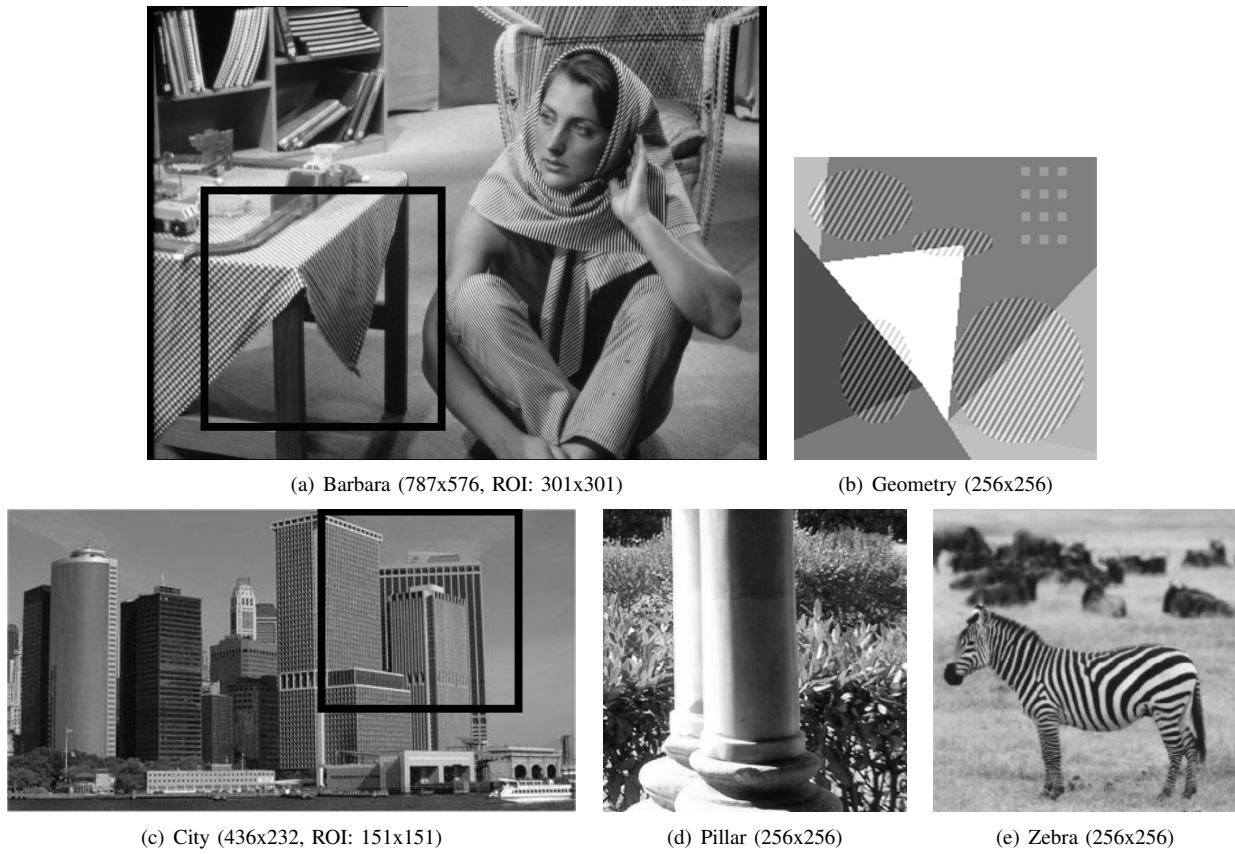


Fig. 6. Images used for visual evaluation. In parentheses after the image name there is the size of the image and the size of the Region of Interest (ROI) if the latter examples are using only a part of the image.

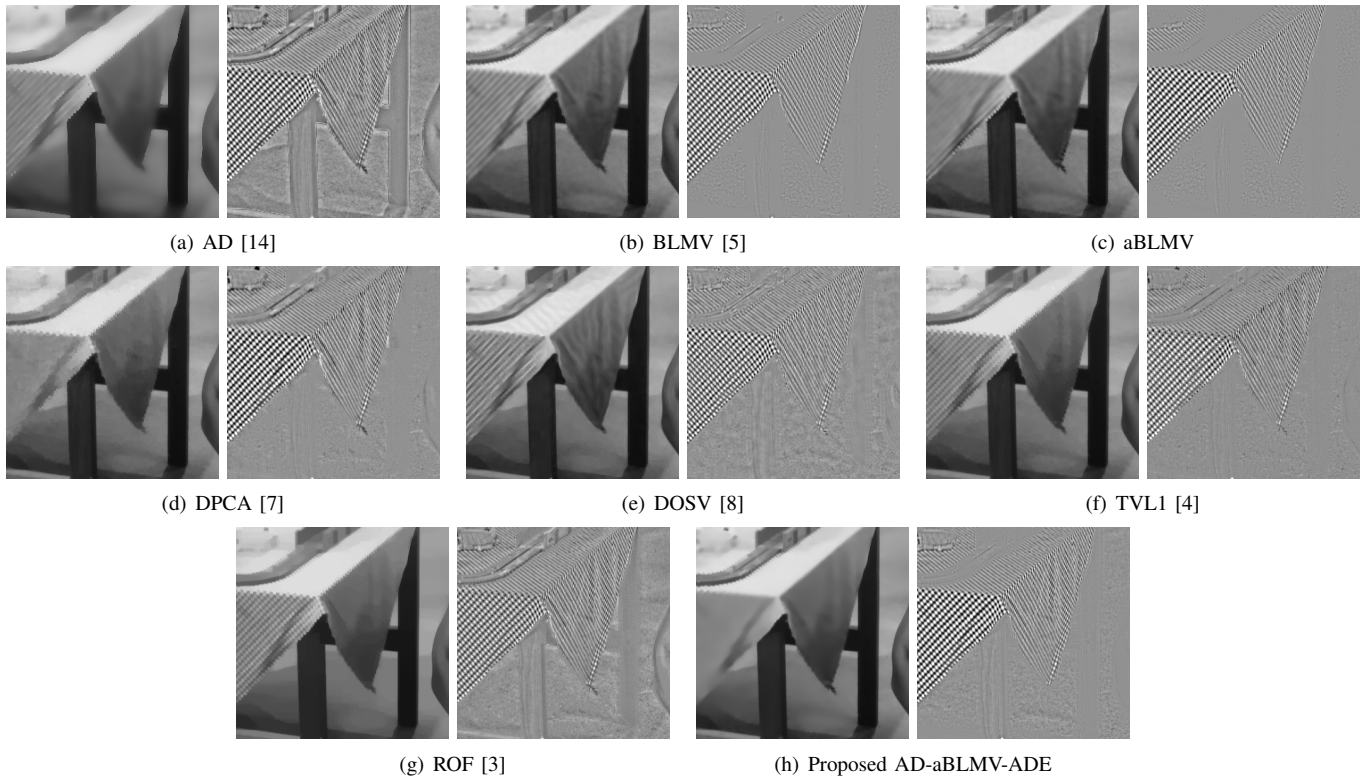


Fig. 8. Separation of cartoon and texture components (Barbara)

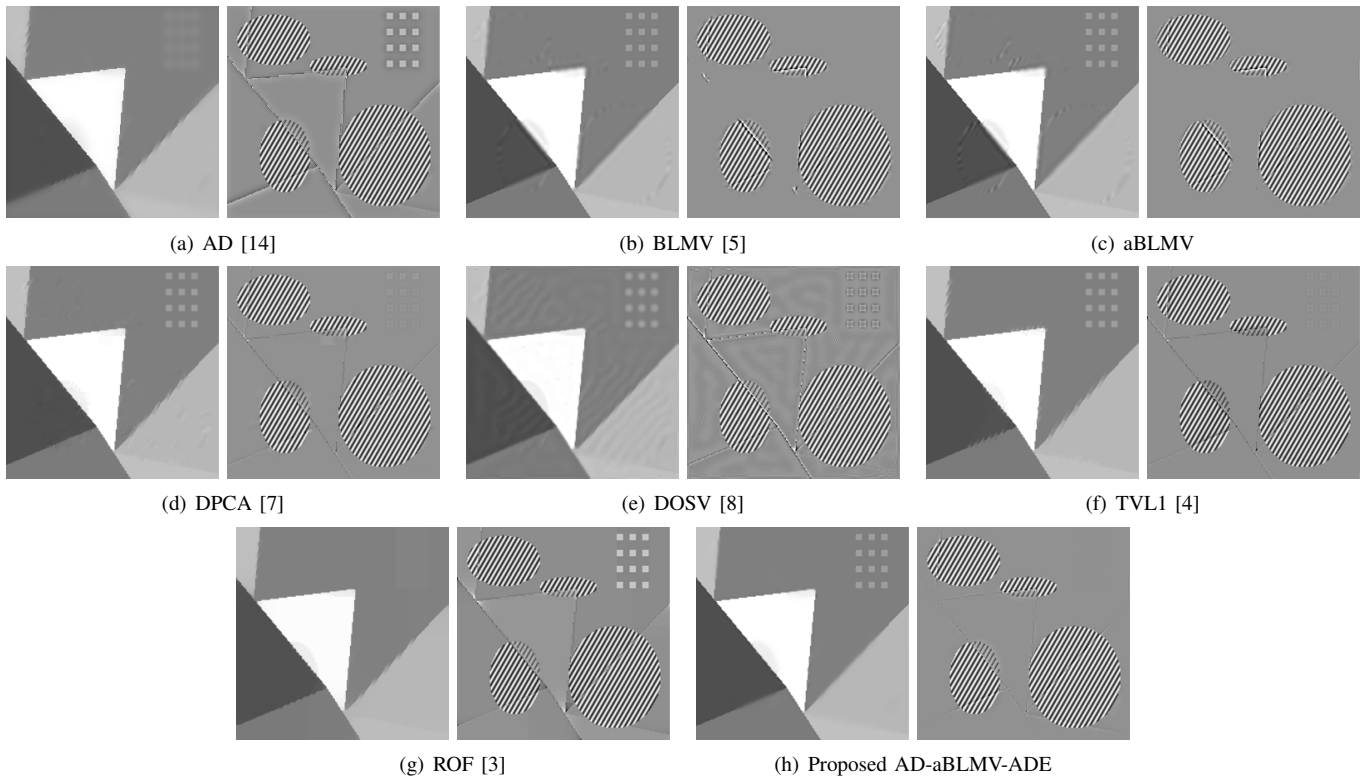


Fig. 9. Separation of cartoon and texture components (Geometry)

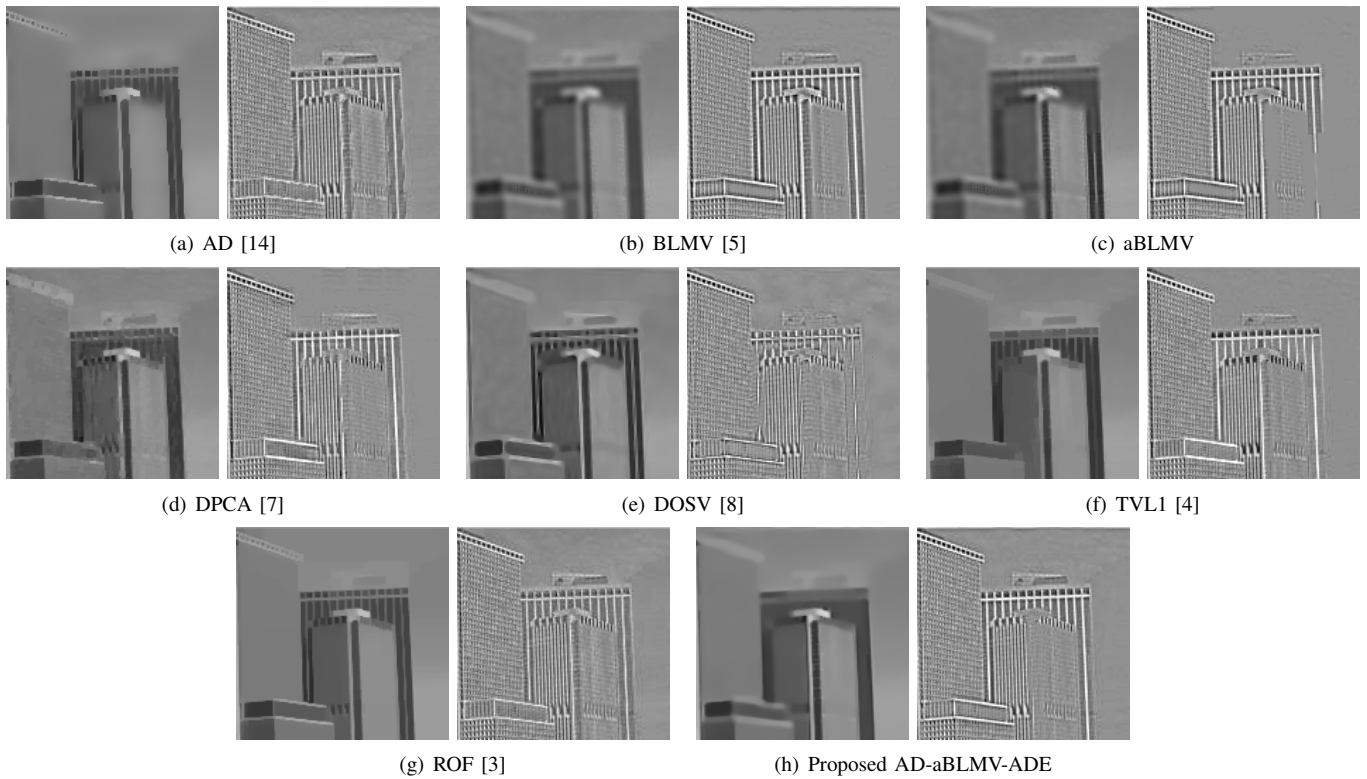


Fig. 10. Separation of cartoon and texture components (City towers)

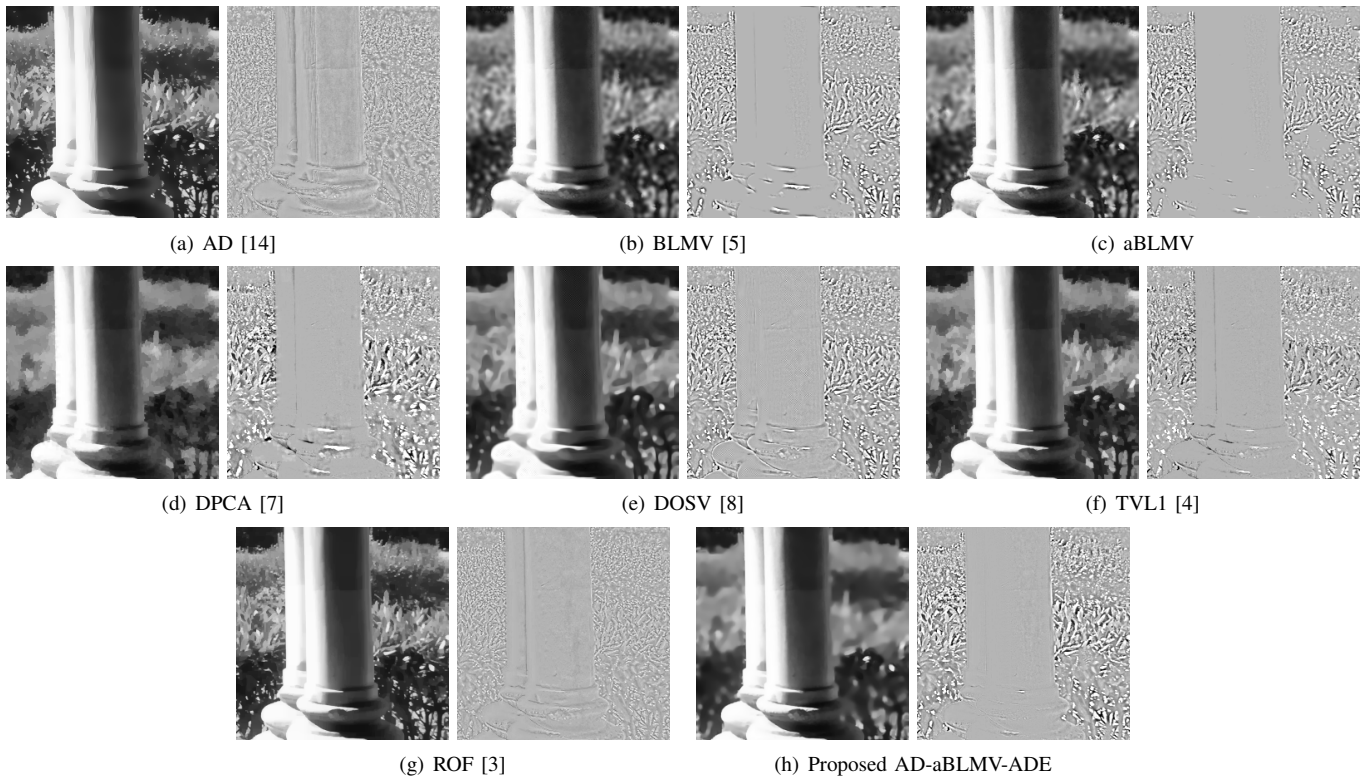


Fig. 11. Separation of cartoon and texture components (Pillar)

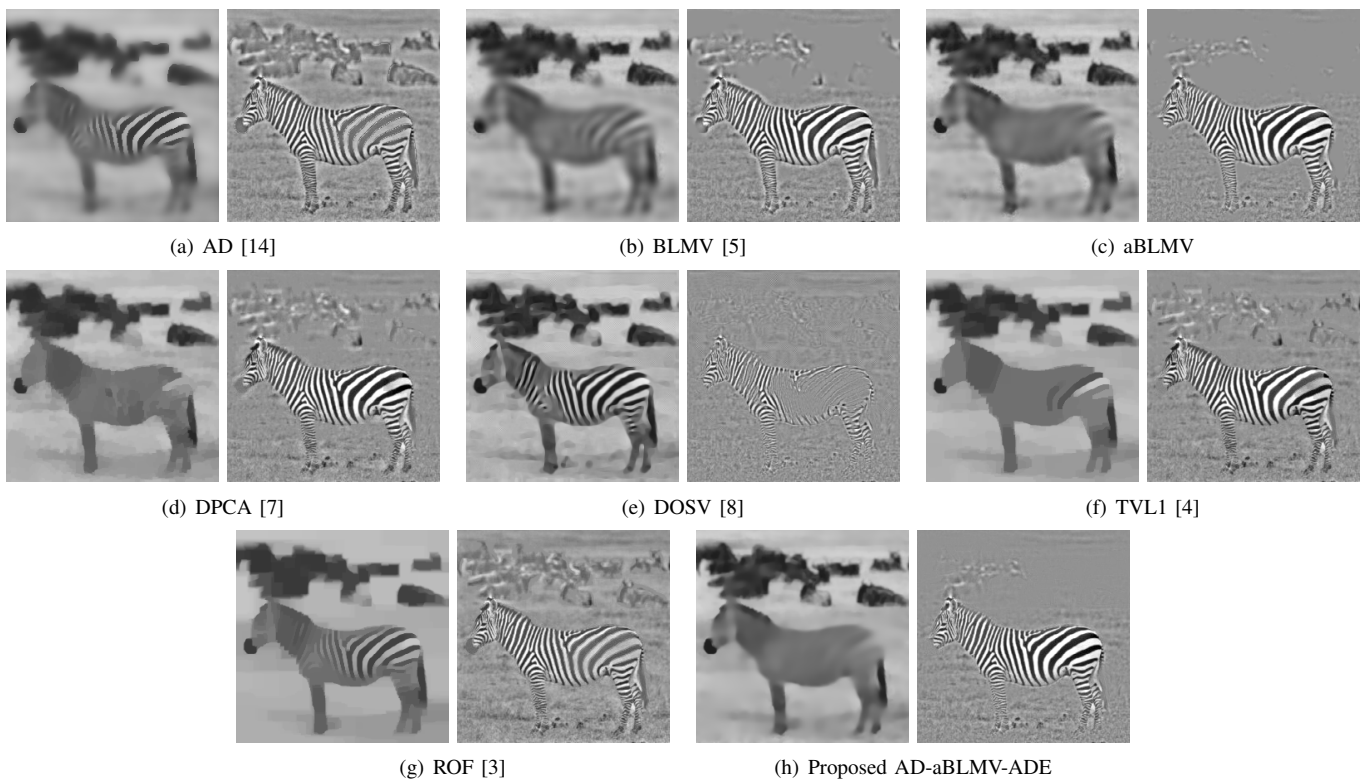


Fig. 12. Separation of cartoon and texture components (Zebra)

B. Numerical Evaluation on Artificial Images

Numerical evaluation is a difficult task for cartoon/texture decomposition since usually there is no ground truth for the images. For this reason most papers in the field lack this kind of comparison and rely only on subjective visual evaluation. We used artificial images for numerical evaluation where the ground truth is available. The following measures were calculated to compare quality: edge absolute difference of the cartoon ($ead(u)$) and texture ($ead(v)$) images, the absolute difference of the cartoon image ($ad(u)$), and the correlation coefficient of the estimated texture and the GT texture. We define these measures as follows:

$$\begin{aligned} ead(u) &= |e(u') - e(u)|, \\ ead(v) &= |e(v') - e(v)|, \\ ad(u) &= |u' - u|, \\ corr(v', v) &= \frac{cov(v', v)}{\sigma_{v'} \sigma_v}, \end{aligned} \quad (17)$$

where u and v are the ground truth cartoon and texture images, u' , v' are the cartoon and texture images produced by a decomposition method and $e(\cdot)$ is the Prewitt edge image [24]. For $ead(u)$, $ead(v)$ and $ad(u)$, the lower value means better result, while for the correlation coefficient, the higher values correspond to better results. In general the proposed AD-aBLMV-ADE method performs better than the rest (see TABLE I- IV).

	Scores for the 1st image of Fig.7			
	ead(u)	ead(v)	ad(u)	corr(v',v)
AD [14]	0.4585	0.3870	3.7430	0.9460
BLMV [5]	0.8105	0.6698	4.5584	0.9238
aBLMV	0.7668	0.6333	4.3541	0.9298
DPCA [7]	0.8322	0.6701	5.7892	0.8939
DOSV [8]	0.5556	0.5130	3.7948	0.9422
TVL1 [4]	0.9311	0.7802	7.9885	0.7893
ROF [3]	0.6683	0.5237	4.6568	0.9367
AD-aBLMV-ADE	0.5075	0.4532	3.6115	0.9479

TABLE I

NUMERICAL RESULTS FOR THE 1ST IMAGE OF FIG.7. THE BEST RESULTS ARE HIGHLIGHTED IN BOLD.

	Scores for the 2nd image of Fig.7			
	ead(u)	ead(v)	ad(u)	corr(v',v)
AD [14]	0.5631	0.5373	2.1536	0.9564
BLMV [5]	0.8393	0.7611	2.6356	0.9263
aBLMV	0.7922	0.7182	2.4525	0.9311
DPCA [7]	0.7420	0.6596	2.8667	0.9361
DOSV [8]	0.4137	0.3911	1.5184	0.9780
TVL1 [4]	0.4339	0.3288	3.8850	0.8693
ROF [3]	0.5926	0.4855	2.6941	0.9508
AD-aBLMV-ADE	0.2782	0.2408	1.0342	0.9862

TABLE II

NUMERICAL RESULTS FOR THE 2ND IMAGE OF FIG.7. THE BEST RESULTS ARE HIGHLIGHTED IN BOLD.

We have also compared the proposed method using ADE to the case when correlation is used instead as an independence measure. It shows that the pure orthogonality criterion

	Scores for the 3rd image of Fig.7			
	ead(u)	ead(v)	ad(u)	corr(v',v)
AD [14]	0.3279	0.2386	1.7500	0.9678
BLMV [5]	0.4363	0.3229	1.8498	0.9661
aBLMV	0.3703	0.3037	1.5932	0.9811
DPCA [7]	0.5535	0.4322	2.3272	0.9658
DOSV [8]	0.3217	0.2788	1.6535	0.9826
TVL1 [4]	0.5589	0.4391	2.2239	0.9681
ROF [3]	0.5105	0.4064	2.3239	0.9658
AD-aBLMV-ADE	0.2682	0.2385	1.2826	0.9876

TABLE III

NUMERICAL RESULTS FOR THE 3RD IMAGE OF FIG.7. THE BEST RESULTS ARE HIGHLIGHTED IN BOLD.

	Scores for the 4th image of Fig.7			
	ead(u)	ead(v)	ad(u)	corr(v',v)
AD [14]	0.4019	0.3268	1.6902	0.9949
BLMV [5]	0.4479	0.3147	1.4710	0.9974
aBLMV	0.3979	0.2805	1.4539	0.9974
DPCA [7]	0.6842	0.5437	9.2985	0.9118
DOSV [8]	0.3564	0.3090	1.2509	0.9984
TVL1 [4]	0.6917	0.4942	29.2620	0.7246
ROF [3]	0.2829	0.2846	2.3199	0.9926
AD-aBLMV-ADE	0.1610	0.1304	1.2245	0.9985

TABLE IV

NUMERICAL RESULTS FOR THE 4TH IMAGE OF FIG.7. THE BEST RESULTS ARE HIGHLIGHTED IN BOLD.

Comparison of independence measures			
ead(u)	ead(v)	ad(u)	corr(v',v)
97.34%	97.31%	98.73%	100.06%

TABLE V

RATIO OF THE ERROR RATES AND THE CORRELATION OF ADE BASED VS. CORRELATION BASED CALCULUS. THE RESULTS OBTAINED BY ADE ARE BETTER THAN THE ONES OBTAINED BY CORRELATION: THE ABSOLUTE DIFFERENCES HAVE DECREASED WHILE THE CORRELATION COEFFICIENT HAS SLIGHTLY INCREASED.

performs slightly better than the correlation based comparison (see TABLE V).

As a drawback of the proposed method, we have to mention that the block based computation of ADE might cause problems at the borders and might oversmooth small and weak cartoon parts, as it can be seen on the City image, where the top of the building is hardly recognizable. Also it could fail preserving cartoon edges in regions where both cartoon and texture edges are present.

The computational time is also increased compared to the fast BLMV, TVL1 and ROF methods, but it is faster than DPCA or DOSV (see TABLE VI).

Computational Time	
AD [14]	3.3
BLMV [5]	0.90
aBLMV	8.8
DPCA [7]	271.5
DOSV [8]	710.7
TVL1 [4]	0.47
ROF [3]	1.3
AD-aBLMV-ADE	21.3

TABLE VI

COMPUTATIONAL TIME (IN SECONDS) OF THE DIFFERENT METHODS FOR THE CITY IMAGE (436x232) ON A PENTIUM IV 2 GHZ NOTEBOOK WITH 3GB MEMORY.

IV. CONCLUSION

In this paper, we have introduced novel, theoretically sound solutions for the main issues of cartoon/texture decomposition, using anisotropic diffusion with ADE based iteration stopping. To initialize the diffusion inhibitions of AD, we used BLMV [5] nonlinear filter with adaptive parameter selection based on ADE calculus. Numerical results and visual comparisons show that the proposed method works with high efficiency and in quality it outperforms other algorithms introduced in recent years.

ACKNOWLEDGMENT

We would like to thank the authors of [4], [5], [7], [8] for making their algorithms available for testing. This work was partially supported by the Hungarian Research Fund #80352.

REFERENCES

- [1] N. Sprljan, M. Mrak, and E. Izquierdo, "Image compression using a cartoon-texture decomposition technique," *Proc. Int. Work. on Image Analysis for Multimedia Interactive Services (WIAMIS)*, p. 91, 2004.
- [2] S. Osher, A. Sole, and L. Vese, "Image decomposition and restoration using total variation minimization and the h^{-1} norm," *Sci. Multiscale Model. Simul.*, vol. 1, pp. 349–370, 2002.
- [3] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D*, vol. 60, pp. 259–268, November 1992.
- [4] W. Yin, D. Goldfarb, and S. Osher, "Image cartoon-texture decomposition and feature selection using the total variation regularized L1 functional," in *Variational, Geometric, and Level Set Methods in Computer Vision*. Springer, 2005, pp. 73–84.
- [5] A. Buades, T. Le, J.-M. Morel, and L. Vese, "Fast cartoon + texture image filters," *IEEE Transactions on Image Processing*, vol. 19, no. 8, pp. 1978–1986, 2010.
- [6] J.-L. Starck, M. Elad, and D. Donoho, "Image decomposition via the combination of sparse representations and a variational approach," *IEEE Transactions on Image Processing*, vol. 14, no. 10, pp. 1570–1582, 2005.
- [7] F. Zhang, X. Ye, and W. Liu, "Image decomposition and texture segmentation via sparse representation," *Signal Processing Letters, IEEE*, vol. 15, pp. 641–644, 2008.
- [8] R. Shahidi and C. Moloney, "Decorrelating the structure and texture components of a variational decomposition model," *IEEE Transactions on Image Processing*, vol. 18, no. 2, pp. 299–309, 2009.
- [9] Y. Meyer, *Oscillating Patterns in Image Processing and Nonlinear Evolution Equations: The Fifteenth Dean Jacqueline B. Lewis Memorial Lectures*. Boston, MA, USA: American Mathematical Society, 2001.
- [10] A. N. Tikhonov and V. Y. Arsenin, *Solutions of ill-posed problems*, ser. Scripta series in mathematics. Washington: Winston, 1977.
- [11] D. Mumford and J. Shah, "Optimal approximations by piecewise smooth functions and associated variational problems," *Communications on Pure and Applied Mathematics*, vol. 42, no. 5, pp. 577–685, 1989.
- [12] J.-F. Aujol and G. Gilboa, "Constrained and snr-based solutions for tv-hilbert space image denoising," *J. Math. Imaging Vis.*, vol. 26, pp. 217–237, November 2006.
- [13] T. Sziranyi and D. Szolgay, "Geometrical and textural component separation with adaptive scale selection," *MUSCLE Int. Workshop on Computational Intelligence for Multimedia Understanding*, 2011, Pisa.
- [14] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, pp. 629–639, 1990.
- [15] G. Sapiro and D. Ringach, "Anisotropic diffusion of multivalued images with applications to color filtering," *IEEE Transactions on Image Processing*, vol. 5, no. 11, pp. 1582–1586, nov 1996.
- [16] J. Weickert, *Anisotropic Diffusion in Image Processing*. Stuttgart, Germany: Teubner-Verlag, 1998.
- [17] L. Kovacs and T. Sziranyi, "Focus area extraction by blind deconvolution for defining regions of interest," *IEEE Tr. Pattern Analysis and Machine Intelligence*, vol. 29, no. 6, pp. 1080–1085, 2007.
- [18] D. Szolgay and T. Sziranyi, "Optimal stopping condition for iterative image deconvolution by new orthogonality criterion," *Electronics Letters*, vol. 47, no. 7, pp. 442–444, 31 2011.
- [19] T. Lindeberg, "Edge detection and ridge detection with automatic scale selection," *Int. Journal of Computer Vision*, vol. 30, pp. 117–154, 1998.
- [20] D. Gabor, "Information theory in electron microscopy," *Laboratory Investigation*, vol. 14/6, pp. 801–807, 1965.
- [21] L. Florack, *Image Structure*. Kluwer Academic Publishers, 1997.
- [22] L. Alvarez, P.-L. Lions, and J.-M. Morel, "Image selective smoothing and edge detection by nonlinear diffusion. ii," *SIAM J. Numer. Anal.*, vol. 29, pp. 845–866, June 1992.
- [23] I. Kopilovic and T. Sziranyi, "Artifact reduction with diffusion preprocessing for image compression," *Optical Engineering*, vol. 44, no. 2, Feb. 2005.
- [24] J. M. S. Prewitt, *Object Enhancement and Extraction*. Picture Processing and Psychopictorics, B. S. Lipkin and A. Rosenfeld, eds., Academic Press, New York, 1970.



Dániel Szolgay received his MsC degree in Electronic and Computer Engineer from the Pázmány Péter Catholic University, Budapest, Hungary in 2006. He defended his PhD in a joint program between the Pázmány Péter Catholic University and the University of Bordeaux 1, Bordeaux, France in 2011. In 2007 he gained the Égide - Eiffel Excellence Scholarships for International PhD Students and a year later the Scholarship of the French government for PhD students in co-tutoring. He has worked on video event analysis for stationary and wearable

camera recordings and also on image deconvolution and image segmentation problems.



Tamás Szirányi received his Ph.D. and D.Sci. degrees in 1991 and 2001, by the Hungarian Academy of Sciences, Budapest. He was appointed to a Full Professor position in 2001 at Pannon University, Veszprém, Hungary, and, in 2004, at the Péter Pázmány Catholic University, Budapest. He leads the Distributed Events Analysis Research Group at the Computer and Automation Research Institute, Hungarian Academy of Sciences. His research activities include machine perception, pattern recognition, texture and motion segmentation, Markov

Random Fields and stochastic optimization, remote sensing, surveillance systems for panoramic and multiple camera systems, intelligent networked sensor systems, image quality, graph based clustering, digital film restoration, image rendering and coding. Dr. Szirányi was the founder and past president (1997 to 2002) of the Hungarian Image Processing and Pattern Recognition Society. He was an Associate Editor of IEEE T. Image Processing (2003-2009). He was honored by the Master Professor award in 2001, and by the Széchenyi professorship. He is a senior member of IEEE and he is the member of the board of Trustees of Bolyai Research Fellowship. He is a Fellow both of the IAPR and the Hungarian Academy of Engineering from 2008. He has more than 200 publications including 40 in major scientific journals.