Image and Video Abstraction by Anisotropic Kuwahara Filtering

Jan Eric Kyprianidis[1*]　　　Henry Kang[2]　　　Jürgen Döllner[1]

[1] Hasso-Plattner-Institut, Germany
[2] University of Missouri, St. Louis

Abstract

We present a non-photorealistic rendering technique to transform color images and videos into painterly abstractions. It is based on a generalization of the Kuwahara filter that is adapted to the local shape of features, derived from the smoothed structure tensor. Contrary to conventional edge-preserving filters, our filter generates a painting-like flattening effect along the local feature directions while preserving shape boundaries. As opposed to conventional painting algorithms, it produces temporally coherent video abstraction without extra processing. The GPU implementation of our method processes video in real-time. The results have the clearness of cartoon illustrations but also exhibit directional information as found in oil paintings.

Categories and Subject Descriptors (according to ACM CCS): I.3.3 [Computer Graphics]: Picture/Image Generation—Display algorithms

Figure 1: *Examples of abstracted images created using our technique.*

1. Introduction

Photorealistic depictions often contain more information than necessary to communicate intended information. Artists therefore typically remove detail and use abstraction for effective visual communication. A typical approach to automatically create stylized abstractions from images or videos is the use of an edge-preserving filter. Popular examples of edge-preserving filters used for image abstraction are the bilateral filter [TM98] and mean shift [CM02]. Both smooth low-contrast regions while preserving high-contrast edges and therefore may fail for high-contrast images where either no abstraction is performed or relevant information is removed because of the thresholds used (Figure 9). They also often

---

*　e-mail: kyprianidis@hpi.uni-potsdam.de

(a) Original image   (b) Kuwahara [KHEK76]   (c) Gaussian Kuwahara [vdB02]

(d) Weighted Kuwahara   (e) Papari et al. [PPC07]   (f) Proposed method

Figure 2: *Comparison of different variants of the Kuwahara filter.*

fail for low-contrast images where typically too much information is removed (Figure 10). An edge-preserving filter that overcomes this limitation is the Kuwahara filter [KHEK76]. Based on local area flattening, the Kuwahara filter properly removes detail in high-contrast regions while also protecting shape boundaries in low-contrast regions. Therefore, it helps to maintain a roughly uniform level of abstraction across the image while providing an overall painting-style look. Unfortunately, the Kuwahara filter is unstable in the presence of noise and suffers from block artifacts. Several extensions and modifications have been proposed to improve the original Kuwahara filter. The most recent work by Papari et al. [PPC07] introduces new weighting windows and a new combination rule. Even though this improves the output quality significantly, clustering artifacts are still noticeable.

In this work, we present a generalization of the Kuwahara filter that removes clustering artifacts by adapting shape, scale and orientation of the filter to the local structure of the input (Figure 3). Due to this adaption of the filter to the local structure, directional image features are better preserved and emphasized. This results in overall sharper edges and a more feature-abiding painterly effect. Our approach is therefore an interesting alternative to typical painterly rendering algorithms [Lit97, Her03, HE04]. These also remove detail, but often fail to preserve sharp edges and surface integrity. In addition, their processing speed is typically slow due to the need for handling individual brush strokes as graphical primitives. Also, when processing video, they require expensive motion estimation to maintain temporal coherence of

the strokes. By contrast, our approach does not require any sophisticated mechanism for handling strokes, and shows excellent temporal coherence without video motion estimation. The GPU implementation of our approach processes video in real-time.

2. Related Work

Nonlinear filters that preserve edges have been an active research topic for several decades. Classical filters in this field are the well-known Kuwahara filter [KHEK76] and its variants [TT77, NM79]. The general idea behind these filters is to replace the input pixel with the mean of the subregion with minimum variance. The variants differ in the way the subregions are defined. The Kuwahara and Tomita-Tsuji filter use rectangular subregions, while the Nagao-Matsuyama filter uses a rotated elongated bar mask. The use of rectangular subregions results in visible artifacts in the output. In addition, the subregion selection process is unstable if noise is present or subregions have the same variance. This results in randomly chosen subregions and corresponding artifacts (Figure 2(b)). A detailed discussion of limitations of the Kuwahara filter can be found in [PPC07]. Van den Boomgaard [vdB02] replaced the box filter by a Gaussian filter, but this leads only to minor improvements (Figure 2(c)).

Recently, Papari et al. [PPC07] have presented a generalized variant of the Kuwahara filter. They define a new criterion to overcome the limitations of the minimum standard deviation criterion. Instead of selecting a single subregion,
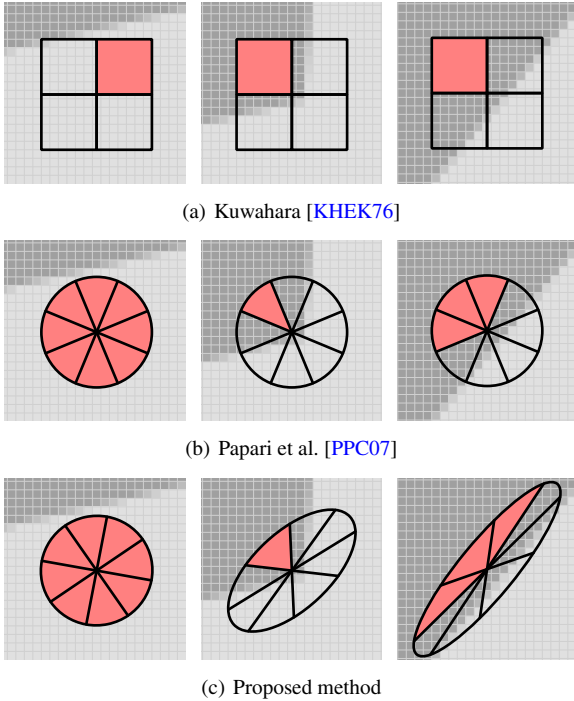
(a) Kuwahara [KHEK76]

(b) Papari et al. [PPC07]

(c) Proposed method

Figure 3: *Comparison of our method to isotropic approaches: (a) Rectangular subregions; a single subregion is selected. (b) Sectors of a disc as subregions; multiple sectors are chosen. (c) Anisotropic filter shape derived from the local structure and divided into subregions; multiple filter responses are chosen. Note that the subregions in (a), (b) and (c) are defined to slightly overlap.*
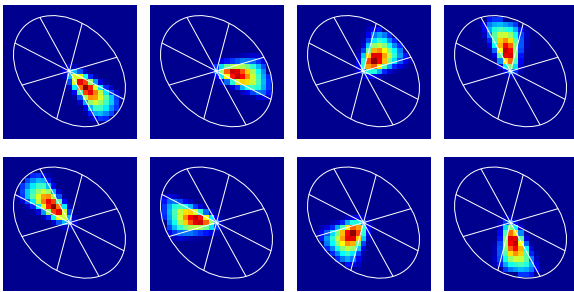

Figure 4: *Filter kernel used by our method ($N = 8$). The shape and orientation of the filter kernel is adjusted to the local structure of the input.*

they define the result as weighted sum of the means of the subregions. The weights are defined based on the variances of the subregions. As can be seen in Figure 2(d), this results in smoother region boundaries and less artifacts. To improve this further, Papari et al. define smooth weighting functions over subregions that are sectors of a disc. This significantly improves the quality of the output (Figure 2(e)), but fails to

capture directional features and results in clustering artifacts. Our approach addresses these limitations by adapting the filter to the local structure of the input. As can be seen in Figure 2(f), this avoids clustering and moreover creates a painterly look for directional image features.

Several classical methods have been successfully extended to take the local structure of images into account. The most popular example is maybe anisotropic diffusion by Perona and Malik [PM90]. Another example is the work of Yang et al. [YBFU96] where an anisotropic Gaussian kernel is used for edge-preserving smoothing. Greenberg and Kogan [GK06] replace the Gaussian kernel in the algorithm of Yang et al. by a median filter and suggest using an elliptic shape for the filter kernel. Wang et al. [WTXC04] replace the radially symmetric kernel, that is used in mean shift to estimate local density, with an anisotropic kernel of elliptic shape.

A popular filter used for image abstraction is the bilateral filter [TM98]. Winnemöller et al. [WOG06] use a separable approximation of the bilateral filter [PvV05] in combination with color quantization to abstract video in real-time. In [KD08] a variant of the bilateral filter is presented that is aligned to the local structure of the image. By this alignment, block artifacts are avoided and smooth boundaries are created between color regions. Kang et al. [KLC09] presented another variant of the bilateral filter that is guided by a feature flow field, which outperforms the full-kernel bilateral filter in terms of feature preservation, noise reduction, and stylization.

DeCarlo and Santella [DS02, SD04] use eye-tracking data to guide image abstraction based on mean shift segmentation [CM02] at different scales. Anisotropic mean shift segmentation [WTXC04] is used in [WXSC04] to transform videos into cartoon animations. Collomosse et al. [CRH05] use mean shift segmentation to extend static non-photorealistic rendering styles to video, including painterly, sketchy, and cartoon shading. An interactive system to create highly abstracted color sketches based on mean shift is presented in [WLL*06]. Apart from failing for high-contrast images like Figure 2(a), a limitation of the mean shift algorithm is that it does not produce smooth boundaries (Figure 9(d)). Therefore, the resulting region boundaries usually require post-processing in order to attain a stylistic look.

Other recently presented image abstraction approaches include a segmentation-based image vectorization system called ARDECO [LL06], structure-preserving manipulation of photographs based on gradient domain image processing [OBBT07], and shape-simplifying image abstraction based on constrained mean curvature flow [KL08]. An edge-preserving smoothing operator, based on the weighted least squares framework, and its application to progressive image abstraction is presented in [FFLS08]. These techniques require some extensive processing and have not been tested for video applications.
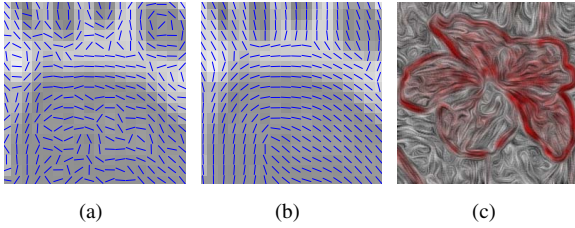
     (a)   (b)   (c)

Figure 5: *Eigenvectors of the structure before (a) and after (b) smoothing with a Gaussian filter. (c) Resulting vector field visualized using line integral convolution. Color indicates the quantity of the maximum eigenvalue $\sqrt{\lambda_1}$.*

3. Method

We begin with calculating the structure tensor and smooth it with a Gaussian filter. Local orientation and a measure for the anisotropy are then derived from the eigenvalues and eigenvectors of the smoothed structure tensor. Finally, structure-aware smoothing is performed using a nonlinear filter. This nonlinear filter uses weighting functions defined over an ellipse, whose shape is based on the local orientation and anisotropy (Figure 4). The filter response is defined as a weighed sum of the local averages, where more weight is given to those averages with low standard deviation.

3.1. Orientation and Anisotropy Estimation

The local orientation and anisotropy estimation is based on the eigenvalues and eigenvectors of the structure tensor [BWBM06]. We calculate the structure tensor directly from the RGB values of the input [KD08]. Let $f \colon \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ be the input image and let $G_{\sigma,x}$ and $G_{\sigma,y}$ be the spatial derivatives in x- and y-direction of a Gaussian with standard deviation σ:

$$G_\sigma(x) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{\|x\|^2}{2\sigma^2}\right)$$

Then, approximations of the partial derivatives of f can be calculated by

$$f_x = G_{\sigma_g,x} \star f \quad \text{and} \quad f_y = G_{\sigma_g,y} \star f \, ,$$

where $\star$ denotes convolution. The structure tensor of f is then defined as:

$$(g_{ij}) = \begin{pmatrix} f_x \cdot f_x & f_x \cdot f_y \\ f_x \cdot f_y & f_y \cdot f_y \end{pmatrix} =: \begin{pmatrix} E & F \\ F & G \end{pmatrix}$$

Here, $\cdot$ denotes the scalar product. The eigenvalues of the structure tensor correspond to the squared minimum and maximum rate of change. The eigenvectors correspond to the respective directions. Selecting the eigenvector corresponding to the minimum rate of change gives a vector field. As shown in Figure 5(a) this vector field has discontinuities. In order to smooth the vector field, smoothing of the structure tensor is performed. The result of applying a Gaussian filter is shown

in Figure 5(b). Smoothing the structure tensor is a linear operation on the tensor, but the effect on the eigenvectors is highly nonlinear and corresponds geometrically to principal component analysis. In our examples, we use a Gaussian filter with standard deviation $\sigma = 2.0$. Note that we do not normalize the tensor, as for example in [WB05]. Therefore, structure tensors corresponding to edges with large gradient magnitude get more weight during smoothing. Hence, orientation information of edges is distributed into the neighborhood of the edges (Figure 5(c)).

The eigenvalues of the structure tensor are non-negative real numbers and are given by:

$$\lambda_{1,2} = \frac{E + G \pm \sqrt{(E-G)^2 + 4F^2}}{2}$$

The eigenvector oriented in direction of the minimum rate of change is given by:

$$t = \begin{pmatrix} \lambda_1 - E \\ -F \end{pmatrix}$$

We define local orientation to be $\varphi = \arg t$. To measure the amount of anisotropy, we use the approach proposed in [YBFU96]:

$$A = \frac{\lambda_1 - \lambda_2}{\lambda_1 + \lambda_2}$$

The anisotropy A ranges from 0 to 1, where 0 corresponds to isotropic and 1 corresponds to entirely anisotropic regions.

3.2. Anisotropic Kuwahara Filtering

Let $f \colon \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ denote our input image and let $x_0 \in \mathbb{R}^2$ be a point. Let φ be the local orientation and A be the anisotropy at x_0. We use the method proposed in [Pha06] to define an elliptical filter shape. To adjust the eccentricity depending on the amount of anisotropy we set:

$$S = \begin{pmatrix} \frac{\alpha}{\alpha+A} & 0 \\ 0 & \frac{\alpha+A}{\alpha} \end{pmatrix}$$

The parameter $\alpha > 0$ is a tuning parameter. For $\alpha \to \infty$ the matrix S converges to the identity matrix. We use $\alpha = 1$ in all examples, which results in a maximum eccentricity of 4. Let R_φ be the matrix defining a rotation by φ and let $h = \lceil 2\sigma_r \rceil$. Then the set

$$\Omega = \left\{ x \in \mathbb{R}^2 : \|SR_{-\varphi} x\| \le h \right\}$$

defines an ellipse, whose major axis is aligned to the local image orientation. This ellipse has high eccentricity in anisotropic regions and becomes a circle in isotropic regions.

The mapping $SR_{-\varphi}$ defines a linear coordinate transform that maps Ω to a disc of radius h. To define weighting functions over Ω, we define corresponding weighting functions over the disc and then pull these back to Ω. We divide the

disc into N equal sectors by defining characteristic functions which are 1 over the sector and 0 otherwise:

$$\chi_i(x) = \begin{cases} 1 & \frac{(2i-1)\pi}{N} < \arg x \le \frac{(2i+1)\pi}{N} \\ 0 & \text{otherwise} \end{cases}, \ i = 0,\dots,N-1$$

To define smooth weighting functions, the characteristic functions of the different sectors χ_i are first convolved and then multiplied with a Gaussian:

$$K_i = (\chi_i \star G_{\sigma_s}) \cdot G_{\sigma_r} \tag{1}$$

The convolution smooths the characteristic functions such that they slightly overlap. The multiplication achieves a decay with increasing radius. Since $\chi_i = \chi_0 \circ R_{-2\pi i/N}$, and since Gaussian functions are rotational invariant, we have:

$$\begin{aligned} K_i &= \left((\chi_0 \star G_{\sigma_s}) \cdot G_{\sigma_r}\right) \circ R_{-2\pi i/N} \\ &= K_0 \circ R_{-2\pi i/N} \end{aligned}$$

Here, $\circ$ denotes composition of functions. Moreover, since $\sum_i K_i(x) = G_{\sigma_r}(x)$ for $x \in \mathbb{R}^2$ the sum of the K_i is equivalent to a Gaussian filter. By pulling back K_i we get weighting functions w_i defined over Ω:

$$\begin{aligned} w_i &= (SR_{-\varphi})^{-1} K_i = K_i \circ SR_{-\varphi} \\ &= K_0 \circ R_{-2\pi i/N} SR_{-\varphi} \end{aligned} \tag{2}$$

Now let

$$m_i = \frac{1}{k} \int f(x)\, w_i(x - x_0)\, \mathrm{d}x \tag{3}$$

be the weighted local averages and

$$s_i^2 = \frac{1}{k} \int f^2(x)\, w_i(x - x_0)\, \mathrm{d}x - m_i^2 \tag{4}$$

be the squared standard deviations, where $k = \int w_i(x - x_0)\, \mathrm{d}x$ denotes the normalization factor. We set

$$\alpha_i = \frac{1}{1 + \|s_i\|^q}$$

and define the output of the filter by:

$$\frac{\sum_i \alpha_i m_i}{\sum_i \alpha_i}$$

Our definition of the weighting factors α_i ensures that more weight is given to sectors with low standard deviation, i.e. those that are more homogeneous. This is similar to the approach of [PPC07], where $\|s_i\|^{-q}$ is used for the weights, but avoids the indetermination, when some of the s_i are zero. The parameter q controls the sharpness of the output. We use $q = 8$ in our examples. The definition of the weights α_i assumes that the values of the input image are given as integer values, e.g. in the range $0\dots255$ for 8-bit images.

4. Implementation

We have implemented the proposed algorithm using C++ and the OpenGL Shading Language (GLSL). In our implementation equation (3) and (4) are approximated by a discrete
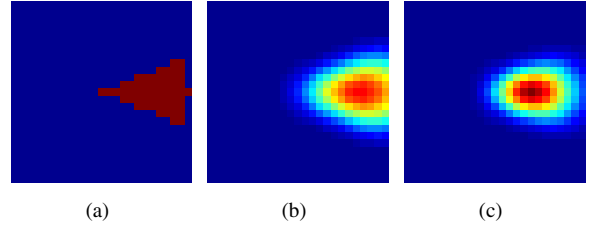

Figure 6: *Approximation of the weighting function K_0 for $N = 8$: (a) Characteristic function χ_0. (b) Characteristic function χ_0 convolved with Gaussian function G_{σ_s}. (c) Finally, multiplication with Gaussian function G_{σ_r} to obtain a decay effect with increasing radius.*

sum. Equation (2) is used to calculate the weighting functions w_i. The weighting function K_0 is approximated by sampling equation (1) into a texture map of size $(4h+2)^2$. Figure 6(c) illustrates the construction of the texture map. Sampling of this texture map is later performed using bilinear interpolation. To reduce the number of texture lookups, we use a 4-channel RGBA texture map, where K_0, K_1, K_2 and K_3 are stored in the red, green, blue and alpha channel. This allows to fetch four weights with a single texture lookup. Moreover, we make use of the fact that the ellipse Ω is symmetric to reduce the number of texture lookups further. Processing time depends on the standard deviation σ_r (i.e. the size of the filter kernel), the number of sectors N and the size of the input. Images at a resolution of 512x512 and video in DVD quality can be processed at 12 fps on a nVidia Geforce GTX 280, with our implementation using $\sigma_r = 3.0$, $\sigma_s = 1.0$ and $N = 8$. A possibility to increase the processing speed is to reduce the number of sectors to $N = 4$. This will result in slightly less sharp boundaries between color regions. We perform all calculations in RGB color space. We have also tested CIE-Lab and CIE-LUV color spaces but found no significant differences.

5. Results

Figure 7 shows a comparison with the generalized Kuwahara filter from [PPC07]. Since this approach uses an isotropic static set of filter kernels, it fails capturing small directional image features. Moreover, shape boundaries are less preserved, because the filter kernels have predefined primary orientations. Due to the fixed isotropic filter kernels, Papari et al.'s approach is prone to group pixels into clusters of circular shape. This becomes more problematic when the filter is applied iteratively to achieve further abstraction (Figure 8). By contrast, our approach avoids this and creates feature-preserving color regions even after multiple applications.

Figure 9 demonstrates the ability of our approach to abstract high-contrast images. The bilateral filter is typically not effective in abstracting high-contrast details, and thus leaves many small and noisy regions untouched (see the ground). In case of mean shift segmentation, its density estimation in
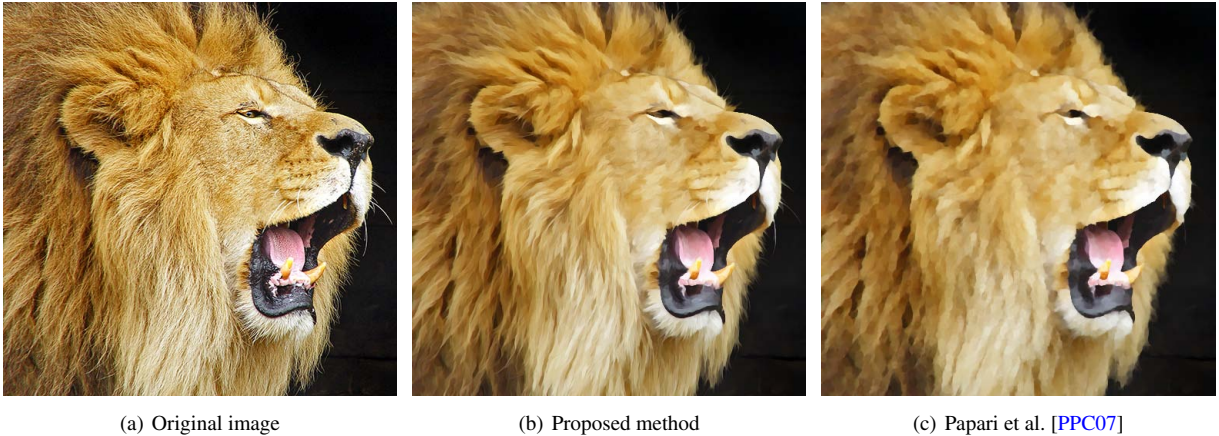
(a) Original image   (b) Proposed method   (c) Papari et al. [PPC07]

Figure 7: *A painterly look is achieved by preserving and enhancing directional image features.*

(a) Original image   (b) Proposed method (1 iteration)   (c) Proposed method (3 iterations)   (d) Papari et al. (3 iterations)
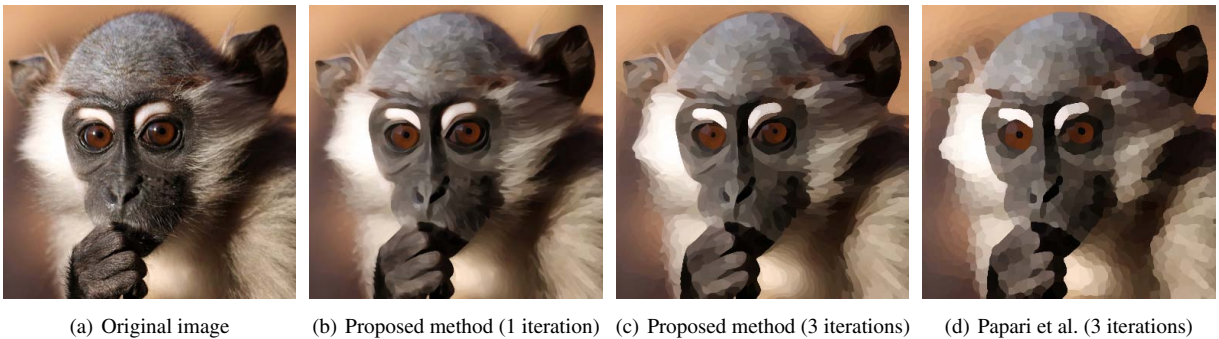
Figure 8: *Our approach creates consistent color regions without showing clustering artifacts when applied iteratively.*

the multidimensional space results in rough region boundaries. On the other hand, our method produces a clean and consistent look of abstraction across the image, without noisy pixels or rough boundaries. Figure 10 shows an example, where the bilateral filter smooths away most of the interesting low-contrast information (the cat's fur) and leaves only high-contrast details (even the black spots around mouth and nose), whereas our approach results in a consistent, feature-enhancing abstraction all over.

Figure 11 shows the vulnerability of the bilateral filter to high-contrast noise even more clearly. The input image here is artificially corrupted with Gaussian and impulse noise. While our approach successfully restores the image and creates a visually pleasing output, the flow-based bilateral filter [KLC09, Pha06], which is known to perform better on noisy input than the full-kernel bilateral filter, clearly fails to remove high-contrast impulse noise.

We also have applied our method to video abstraction and find that per-frame filtering alone provides excellent temporal coherence. Thus, no extra processing, such as motion estimation [Lit97, HE04] or adaptive color quantization [WOG06] is required. The GPU implementation of our approach processes video in real-time. Providing a painterly look, it is an

interesting alternative to video abstraction frameworks based on the bilateral filter such as [WOG06, KD08].

A limitation of our technique is that it is not suitable for creating a "rough and brushy" look as found in some oil paintings (e.g., "Starry Night" by Vincent van Gogh). To create such strong, textured brush effects, a background paper texture or a directional, cumulative alpha map [Her02] could be incorporated into our approach.

6. Conclusion

We have presented an automatic technique for image and video abstraction based on anisotropic generalized Kuwahara filtering. Guided by the smoothed structure tensor, our filter generates a feature-preserving, direction-enhancing painterly look, without having to deal with individual brush strokes. Unlike existing nonlinear smoothing filters, our technique is robust against high-contrast noise and avoids overblurring in low-contrast areas, providing a consistent level of abstraction across the image. It also ensures outstanding temporal coherence in video abstraction, even with per-frame filtering. The GPU-based implementation processes video at real-time rates.
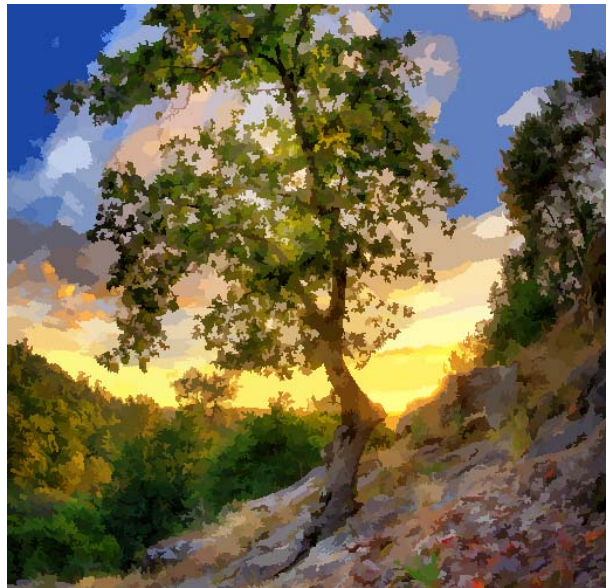
(a) Original image

(b) Proposed method
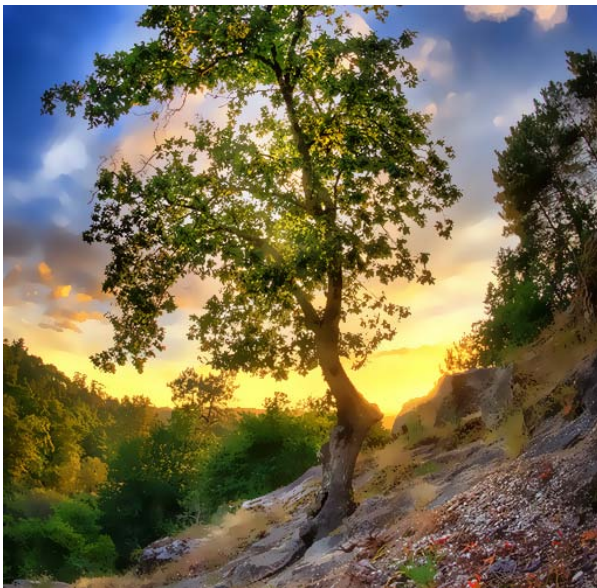
(c) Bilateral filter [TM98]

(d) Mean shift segmentation [CM02]

Figure 9: *Comparison with bilateral filter and mean shift segmentation.*

References

[BWBM06] BROX T., WEICKERT J., BURGETH B., MRÁZEK P.: Nonlinear structure tensors. *Image and Vision Computing 24*, 1 (2006), 41 – 55.

[CM02] COMANICIU D., MEER P.: Mean shift: A robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence 24*, 5 (2002), 603–619.

[CRH05] COLLOMOSSE J. P., ROWNTREE D., HALL P. M.: Stroke surfaces: Temporally coherent artistic animations from video. *IEEE Transactions on Visualization and Computer Graphics 11*, 5 (2005), 540–549.

[DS02] DECARLO D., SANTELLA A.: Stylization and abstraction of photographs. In *SIGGRAPH '02: Proceedings of the 29th annual conference on Computer graphics and interactive techniques* (2002), pp. 769–776.

(a) Original image   (b) Proposed method   (c) Bilateral filter

Figure 10: *In constrast to the bilateral filter our approach avoids overblurring in low-contrast regions.*

(a) Original image   (b) Proposed method   (c) Flow-based bilateral filter [KLC09]

Figure 11: *Application to an image with 2% Gaussian noise and 5% impulse noise.*

[FFLS08] FARBMAN Z., FATTAL R., LISCHINSKI D., SZELISKI R.: Edge-preserving decompositions for multi-scale tone and detail manipulation. In *SIGGRAPH '08: ACM SIGGRAPH 2008 papers* (2008), pp. 1–10.

[GK06] GREENBERG S., KOGAN D.: Improved structure-adaptive anisotropic filter. *Pattern Recognition Letters 27*, 1 (2006), 59–65.

[HE04] HAYS J., ESSA I.: Image and video based painterly animation. In *NPAR '04: Proceedings of the 3rd international symposium on Non-photorealistic animation and rendering* (2004), pp. 113–120.

[Her02] HERTZMANN A.: Fast paint texture. In *NPAR '02: Proceedings of the 2nd international symposium on Non-photorealistic animation and rendering* (2002), pp. 91–ff.

[Her03] HERTZMANN A.: Tutorial: A survey of stroke-based rendering. *IEEE Computer Graphics and Applications 23*, 4 (2003), 70–81.

[KD08] KYPRIANIDIS J. E., DÖLLNER J.: Image abstraction by structure adaptive filtering. In *Proc. EG UK Theory and Practice of Computer Graphics* (2008), Eurographics Association, pp. 51–58.

[KHEK76] KUWAHARA M., HACHIMURA K., EIHO S., KINOSHITA M.: *Digital processing of biomedical images*. Plenum Press, 1976, pp. 187–203.

[KL08] KANG H., LEE S.: Shape-simplifying image abstraction. *Computer Graphics Forum 27*, 7 (2008), 1773–1780. Special issue on the Pacific Graphics 2008.

[KLC09] KANG H., LEE S., CHUI C. K.: Flow-based image abstraction. *IEEE Transactions on Visualization and Computer Graphics 15*, 1 (2009), 62–76.

[Lit97] LITWINOWICZ P.: Processing images and video for an impressionist effect. In *SIGGRAPH '97: Proceedings of the 24th annual conference on Computer graphics and interactive techniques* (1997), pp. 407–414.

[LL06] LECOT G., LÉVY B.: Ardeco: Automatic region detection and conversion. In *Eurographics Symposium on Rendering* (2006).

[NM79] NAGAO M., MATSUYAMA T.: Edge preserving smoothing. *Computer Graphics and Image Processing 9* (1979), 394–407.

[OBBT07] ORZAN A., BOUSSEAU A., BARLA P., THOLLOT J.: Structure-preserving manipulation of photographs. In *NPAR '07: Proceedings of the 5th international symposium on Non-photorealistic animation and rendering* (2007), pp. 103–110.

[Pha06] PHAM T.: *Spatiotonal adaptivity in Super-Resolution of Undersampled Image Sequences*. PhD thesis, Quantitative Imaging Group, Delft University of Technology, 2006.

[PM90] PERONA P., MALIK J.: Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence 12*, 7 (1990), 629–639.

[PPC07] PAPARI G., PETKOV N., CAMPISI P.: Artistic edge and corner enhancing smoothing. *IEEE Transactions on Image Processing 16*, 10 (2007), 2449–2462.
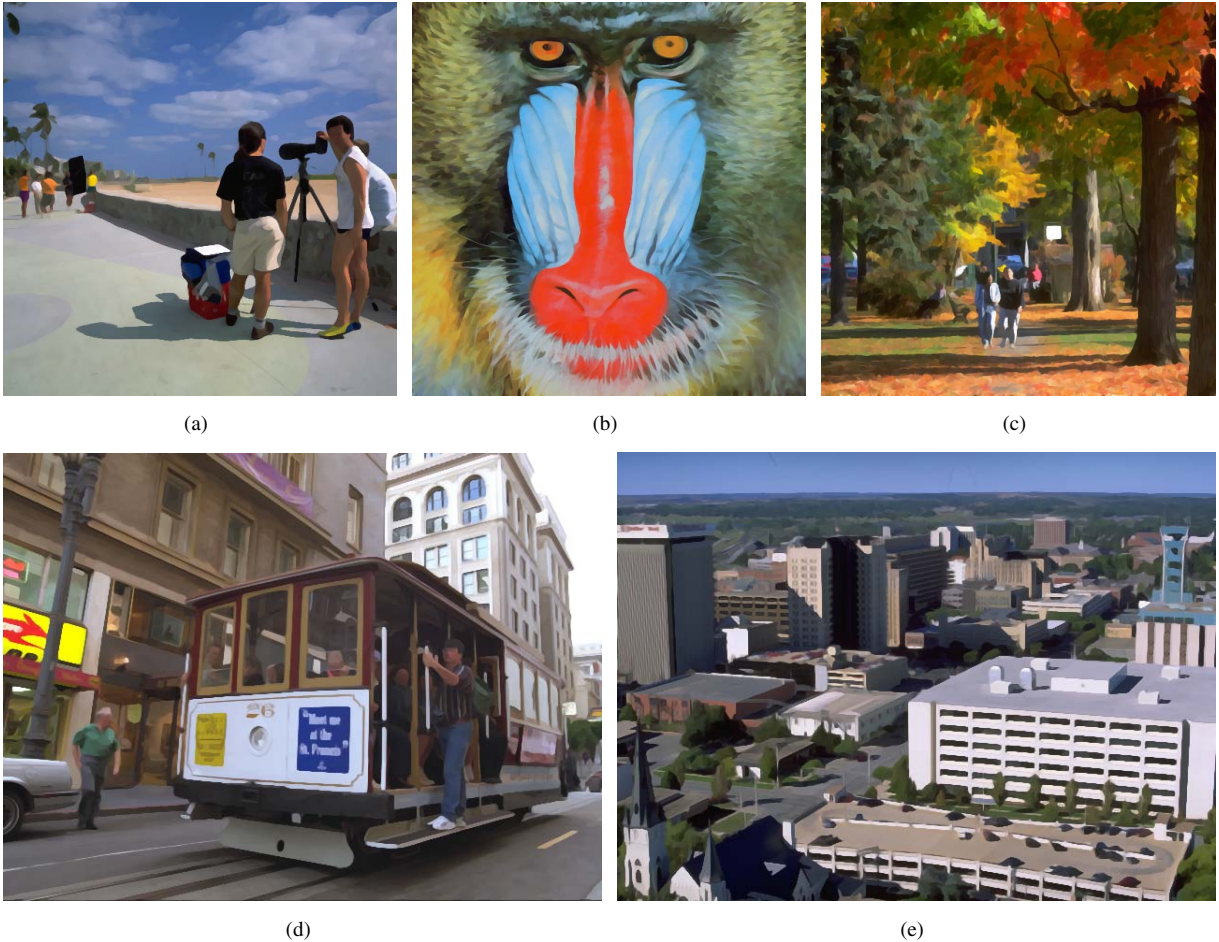
(a)

(b)

(c)

(d)

(e)

Figure 12: *Examples: $\sigma_r = 3.0$, $\sigma_s = 1.0$, $N = 8$, $q = 8$.*

[PvV05] PHAM T., VAN VLIET L.: Separable bilateral filtering for fast video preprocessing. In *IEEE International Conference on Multimedia and Expo* (2005).

[SD04] SANTELLA A., DECARLO D.: Visual interest and npr: an evaluation and manifesto. In *NPAR '04: Proceedings of the 3rd international symposium on Non-photorealistic animation and rendering* (2004), pp. 71–150.

[TM98] TOMASI C., MANDUCHI R.: Bilateral filtering for gray and color images. In *Proceedings International Conference on Computer Vision (ICCV)* (1998), pp. 839–846.

[TT77] TOMITA F., TSUJI S.: Extraction of multiple regions by smoothing in selected neighborhoods. *IEEE Transactions on Systems, Man and Cybernetics 7* (1977), 107–109.

[vdB02] VAN DEN BOOMGAARD R.: Decomposition of the Kuwahara-Nagao operator in terms of linear smoothing and morphological sharpening. In *ISMM 2002* (2002).

[WB05] WEIJER J. V., BOOMGAARD R. V.: Least squares and robust estimation of local image structure. *Int. J. Comput. Vision 64*, 2-3 (2005), 143–155.

[WLL*06] WEN F., LUAN Q., LIANG L., XU Y.-Q., SHUM H.-Y.: Color sketch generation. In *NPAR '06: Proceedings of the 4th international symposium on Non-photorealistic animation and rendering* (2006), pp. 47–54.

[WOG06] WINNEMÖLLER H., OLSEN S. C., GOOCH B.: Real-time video abstraction. In *SIGGRAPH '06: ACM SIGGRAPH 2006 Papers* (2006), pp. 1221–1226.

[WTXC04] WANG J., THIESSON B., XU Y., COHEN M.: Image and video segmentation by anisotropic kernel mean shift. In *Proceedings European Conference on Computer Vision (ECCV)* (2004), Springer, pp. 238–249.

[WXSC04] WANG J., XU Y., SHUM H.-Y., COHEN M. F.: Video tooning. In *SIGGRAPH '04: ACM SIGGRAPH 2004 Papers* (2004), pp. 574–583.

[YBFU96] YANG G. Z., BURGER P., FIRMIN D. N., UNDERWOOD S. R.: Structure adaptive anisotropic image filtering. *Image and Vision Computing 14*, 2 (1996), 135 – 145.