Image Abstraction Using Anisotropic Diffusion Symmetric Nearest Neighbor Filter

Conference Paper · December 2014

DOI: 10.1007/978-3-319-13168-9_38

3 authors, including:

John See
Heriot-Watt University Malaysia

Alfian Abdul Halin
Universiti Putra Malaysia

Image Abstraction Using Anisotropic Diffusion Symmetric Nearest Neighbor Filter

Zoya Shahcheraghi[1,2], John See[1], and Alfian Abdul Halin[2]

[1] Faculty of Computing and Informatics, Multimedia University, Cyberjaya, Malaysia
[2] Faculty of Computer Science & Information Technology, Universiti Putra Malaysia, Serdang, Malaysia

Abstract. Image abstraction is an increasingly important task in various multimedia applications. It involves the artificial transformation of photorealistic images into cartoon-like images. To simplify image content, the bilateral and Kuwahara filters remain popular choices to date. However, these methods often produce undesirable over-blurring effects and are highly susceptible to the presence of noise. In this paper, we propose an image abstraction technique that balances region smoothing and edge preservation. The coupling of a classic Symmetric Nearest Neighbor (SNN) filter with anisotropic diffusion within our abstraction framework enables effective suppression of local patch artifacts. Our qualitative and quantitative evaluation demonstrate the significant appeal and advantages of our technique in comparison to standard filters in literature.

Keywords: Image abstraction, artistic stylization, edge-preserving filters, anisotropic diffusion.

1 Introduction

Image and video abstraction, through Image-based Artistic Rendering (IB-AR) [1] have become increasingly popular in contemporary films, computer generated animation, computer games and visual communication. IB-AR involves artificially incorporating cartoon-like effects to photorealistic images. Although various image processing filters have been used, few have produced artistically appealing and interesting results. This is probably because the filters are mainly applied for photorealistic enhancement, restoration and recovery. In contrast, IB-AR generally aims towards simplification of image content. Basically, image abstraction enables details in low-contrast regions to be removed without filtering across discontinuities, leaving the overall image structure unaffected. Two prominent techniques are the bilateral and Kuwahara filters. The bilateral filter excels in smoothing low-contrast regions while preserving high-contrast edges. It however fails in high-contrast images where either no abstraction is produced or salient visual features are totally removed. Furthermore, iterative filtering using bilateral filter over-blurs edges causing an undesirable washed-out appearance. On the other hand, the Kuwahara filter is able to cope better with high-contrast

W.T. Ooi et al. (Eds.): PCM 2014, LNCS 8879, pp. 343–352, 2014.

images, but it is rather unstable in the presence of noise and suffers from blocky artifacts. Another classical edge-preserving filter called the Symmetric Nearest Neighbor (SNN) filter has relations to the mean and median filters with desirable edge-preserving properties. Unfortunately, the SNN is also susceptible to local patch-like artifacts. Surprisingly, the SNN filter has yet to be explored in literature for image abstraction. In this paper we propose an Anisotropic Diffusion Symmetric Nearest Neighbour Filter (AD-SNN) which balances region smoothing and edge preservation, while enforcing artifact suppression.

2 Related Work

In this section, we provide an overview of relevant image/video abstraction works in literature that utilize edge-preserving filters.

The bilateral filter [2] is by far the most popular edge-preserving filter and it has been widely used for image smoothing, noise reduction and segmentation. The landmark work by Winnemller et al. [3] established an effective framework for abstraction, with the use of an isotropic bilateral filter with Difference-of-Gaussians (DoG) edge extraction to produce highly stylized video frames. Kyprianidis and Döllner [4] came up with a two-pass separated implementations of the bilateral filter along the gradient and tangent directions, which estimates the local orientation based on structure tensors. Elsewhere, Kang et al. [5] also improved the original technique by introducing flow-based image abstraction, where the shapes of the bilateral and DoG filters were deformed to follow a vector field derived from salient image features. An anisotropic bilateral filter was used as opposed to an isotropic bilateral filter.

In another application, Cong et al. [6] introduced a mixed-reality selective image abstraction tool, where a new diffusion model was proposed by combining nonlinear diffusion and bilateral filter. More recently, Zang et al. [7] presented a novel image smoothing approach by using a 1-D space-filling curve to perform EMD-based filtering. Gaussian and bilateral filtering were used to consolidate the edge structure in their space filling curve based image smoothing. Generally, the bilateral filter can sometimes produce inconsistent abstraction in high-contrast images whereby any of the two extremities could occur: no abstraction performed, or over-removal of fine details.

The Kuwahara filter [8] on the other hand is a classical edge-preserving image smoothing filter that performs well in high contrast regions. The generalized Kuwahara filter was used for image and video abstraction by Papari et al. [9] whereby the original Kuwahara filter was improved by replacing the rectangular sub-regions with smooth weighting functions constructed over sectors of a circular area. Kyprianidis et al. [10] introduced an anisotropic Kuwahara filter where the weighting functions are alternatively defined over sectors of ellipses. This modification improved its resistance towards noise and artifacts. More recently, the same authors further introduced another variant that performs at multiple scales [11]. Generally, the Kuwahara filter causes unusual blocky artifacts, and is highly susceptible towards image noise. Many recent methods attempt to resolve these problems but its method of filtering with overlapping windows could be the main cause of these problems.

3 Proposed Method

The proposed method resides within an image abstraction framework that is inspired by the works of Winnemöller et al. [3] and Kang et al. [5]. The architecture employed by these well-known works can be succinctly summarized into two distinct components or processes: (i) *region smoothing*, and (ii) *edge contouring*, in which the contributions of this work involves the former. State-of-the-art methods for edge contouring [5,4] are incorporated into our framework. An overview of the processes in our framework is shown in Figure 1.

In the first step of abstraction, we perform edge-preserving region smoothing on the perceptually uniform CIELab colorspace. The edge contours are then extracted after a certain number of smoothing iterations, while color quantization is applied on the luminance channel as described in [3]. Finally, the edge contours are combined with the color-quantized output in the original RGB colorspace to produce the abstracted output. In this work, we proposed to utilize a classic Symmetric Nearest Neighbor (SNN) filter that is both edge-preserving and simplistic in computation, while adding further artifact suppression from anisotropic diffusion.

3.1 Symmetric Nearest Neighbor (SNN) Filter

Symmetric Nearest Neighbor (SNN) filter [12,13] is an edge-preserving smoothing filter that is arguably one of the most all-rounded filters in terms of both noise immunity (smoothing effect) and edge preservation. It balances both aspects equally as compared to filters such as the median filter (more smoothing), and Kuwahara filter (more edge-preserving).

The SNN filter uses a neighborhood selection technique whereby it compares the 8-connected neighboring pixels in a symmetric fashion. By example of a 3×3 filter, the symmetric pairs of neighbor pixels around the center pixel, *i.e.* N-S, W-E, NW-SE, NE-SW, are inspected by selecting the intensity values of those closest (intensity-wise, not distance-wise) to the center pixel (see Figure 2) from each pair. In other words, the pixel in each symmetric pair that is closest to the value of the center pixel is selected. Finally, the value of the center pixel is computed by the mean value of the four selected neighborhood pixels. Concisely,

$$
\begin{aligned}
\hat{P}_{N,S} &= min\{p_{x,y}, \{p_{x,y+h}, p_{x,y-h}\}\ \}, \quad \text{for N-S pair}\\
\hat{P}_{W,E} &= min\{p_{x,y}, \{p_{x+h,y}, p_{x-h,y}\}\ \}, \quad \text{for W-E pair}\\
\hat{P}_{NW,SE} &= min\{p_{x,y}, \{p_{x+h,y+h}, p_{x-h,y-h}\}\ \}, \quad \text{for NW-SE pair}\\
\hat{P}_{SW,NE} &= min\{p_{x,y}, \{p_{x+h,y-h}, p_{x-h,y+h}\}\ \}, \quad \text{for SW-NE pair}
\end{aligned} \tag{1}
$$

where $h = 1$ for the case of a 3×3 SNN filter, while $min\{\cdot\}$ selects a neighborhood pixel (from each symmetric pairing) that has an intensity value that is closest

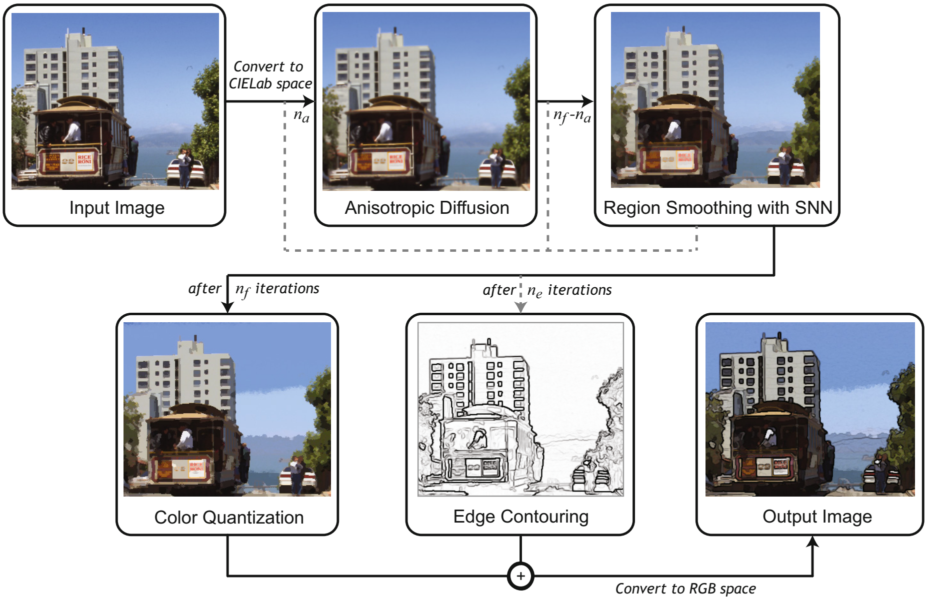

Fig. 1. Schematic overview of the proposed image stylization framework

to that of the center pixel $p_{x,y}$. In this case, the output pixel value is simply the mean of the four values obtained from Eq. 1, *i.e.*

$$f_{x,y} = \{\hat{P}_{N,S} + \hat{P}_{W,E} + \hat{P}_{NW,SE} + \hat{P}_{SW,NE}\}/4 \tag{2}$$

To further generalize this concept for a $w \times w$ SNN filter (where w should be an odd number and $w > 1$), we consider $\eta_w = \{\sum_{i=1}^{(w-1)/2} 4i\}$ number of symmetric pairs within the square filter neighborhood of size w. The example in Figure 2 illustrates a few possible symmetric pairs within a 5×5 neighborhood. Concisely, the SNN-filtered pixel value at location (x, y) can be determined as,

$$f_{x,y} = \frac{1}{\eta_w} \sum_{j=1}^{\eta_w} min\{p_{x,y}, \hat{S}_j\} \tag{3}$$

where $\hat{S}_j$ is the j-th symmetrically opposing pair of pixels. To retain the original image size, zero-padding is applied before performing filtering.

Geometrically, this concept is edge-preserving as the filter "votes" for the neighboring pixels that most closely resemble the center pixel before taking the average value of the selected neighbors. This compensates the smoothing effect (averaging) with retention of intensity gradients (closest value voting).

3.2 Anisotropic Diffusion

Anisotropic diffusion is another classical technique proposed by Perona and Malik [14] to selectively enhance contrast in parts of an image by using a modified

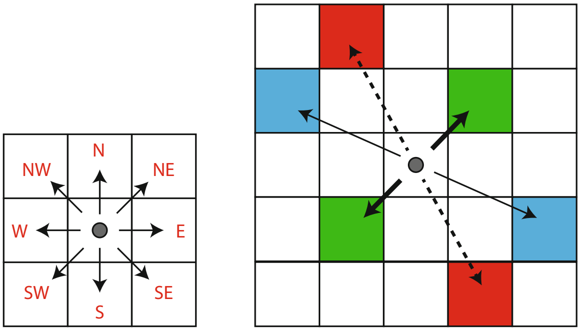

Fig. 2. Symmetric pairs of pixels in SNN filter. The 3×3 *(left)* filter has four pairs; The color-coded pixels in the 5×5 *(right)* window indicate some possible pairs.

heat diffusion equation. Anisotropic diffusion filters have the desirable property of blurring small discontinuities and sharpening edges, which is most useful for image abstraction (the iterative bilateral filter is an example of a nonlinear diffusion filter [3] where the Gaussian blur is weighted nonlinearly). The function of the SNN filter closely resembles a typical rank order mean filter, which will result in the appearance of visual artifacts at corner edges. Hence, we propose to apply the nonlinear diffusion effect on the original image features before performing filtering with SNN, as this maintains the simplicity of the SNN computation without incurring computationally expensive nonlinear filtering.

The 2-D discrete implementation of anisotropic diffusion is summarized as,

$$\frac{\partial}{\partial t} I(\bar{x}, t) = \mathrm{div}\ [c(\bar{x}, t) * \nabla I(\bar{x}, t)] \tag{4}$$

where $\bar{x}$ represents the spatial coordinates (x, y), div$[\cdot]$ denotes the divergence operator and ∇ is the gradient operator *w.r.t.* the spatial coordinates.

The intensity values are then updated locally by the sum of the flow contributions along all four directions (north, south, east, west)

$$I(\bar{x}, t + \Delta t) \approx I(\bar{x}, t) + \Delta t * \frac{\partial}{\partial t} I(\bar{x}, t) \tag{5}$$

For better isotropy, signal flow can also be computed between diagonally neighboring pixels (8-connectedness), with consideration for longer distances ($\sqrt{2}$) between diagonal neighbors. The conduction (diffusion) coefficient function,

$$c(\bar{x}, t) = e^{(-(\|\nabla I\|/K)^2)} \tag{6}$$

which prioritizes high-contrast edges over low-contrast edges, is used in our work. Good parameter values were determined empirically; we fix the threshold constant that controls the level of conduction to $K = 30$, while $\Delta t = 0.143$ (or 1/7) for numerical stability. A subsequent work by Gerig et al. [15] suggested for the diffusion process to be repeated for better convergence towards stable smoothened regions. Further details on its mathematical principles and formulation can be found in the original paper [14] and [15]. The coupling of the anisotropic diffusion and the SNN filter in our framework gives rise to the proposed Anisotropic Diffusion Symmetric Nearest Neighbor (AD-SNN) technique.

3.3 Edge Contouring

The second component of the framework involves edge contouring. Generally, this is an essential abstraction element that creates cartoon-like effects and increases the visual distinctiveness of edges at important locations such as object boundaries and background textures. In our work, we do not propose any new edge contouring methods since the scope of our contribution is in the first component. As such, two prominent methods from literature are applied to our framework, both of which are used in our experiments:

- **Multi-image gradient (MiG).** Di Zenzo [16] proposed an approach that computes the gradient image of a multi-image (or multi-spectral image) by calculating the maximum rate of change in the intensity values along the spectral dimensions. An approximation to the directional derivatives are first determined using Sobel filter. Then, the structure tensor for each point (x, y) is set in quadratic form (with g being the tensor coefficients), measuring the maximum squared rate of change of a vector $n = (n_x, n_y)$:

$$E_n = g_{xx}n_x^2 + 2g_{xy}n_x n_y + g_{yy}n_y^2 \tag{7}$$

- **Flow-based Difference-of-Gaussians (FDoG).** Kang et al. [17] introduced a method of extracting spatially coherent lines, guided by Edge Tangent Flow (ETF). The Difference-of-Gaussians (DoG) is given by,

$$E_{\sigma_e} = G_{\sigma_e} - \tau \cdot G_{\sigma_r} \tag{8}$$

  where G indicate the two Gaussian blurring functions with their respective values; σ_r is set to $1.6 \cdot \sigma_e$ to approximate the Laplacian-of-Gaussian, σ_e controls the spatial scale of edges. The threshold τ determines the sensitivity of the edge detector. We use $\sigma_e = 1.0$ and $\tau = 0.99$ in our experiments.

The final thresholding step that uses a smoothed step function similar to that described in [3],

$$T(E_{\sigma_e}) = \begin{cases} 1 & \text{if } E_{\sigma_e} > 0, \\ 1 + tanh(\varphi_e \cdot E_{\sigma_e}) & \text{otherwise} \end{cases} \tag{9}$$

where φ_e tunes the sharpness of the edges E_{σ_e}

3.4 Iterative Processing

In our framework, a total of n_f iterations are performed for region smoothing. This is a common practice in most works [3,4,5] as it provides sufficient region contrasting and a more desirable level of abstraction. Anisotropic diffusion is applied for a lesser number of iterations to soften the effect of artifact suppression— the first n_a iterations where $n_a < n_f$. Similarly, the edges are extracted after n_e iterations (where $n_e < n_f$) to reduce the presence of noise. In our experiments, we tested and used the following values: $n_f \in \{3, 4\}$, $n_a = 1$, $n_e = 2$. More precisely, both the Kuwahara and AD-SNN filters use $n_f = 3$, while the bilateral filter uses $n_f = 4$ (similar to the setting from [4]).

4 Results and Analysis

Experimental studies remain one of the most difficult problems in the area of non-photorealistic rendering (NPR) when evaluating the artistic stylization of images or videos, as acknowledged by Hertzmann [18]. It is common to perform visual assessment of the aesthetic qualities of stylized images through subjective evaluation by human observers, as employed by many previous works [3,19,20].

The proposed approach was evaluated using a mix of different quantitative and qualitative evaluation techniques. Firstly, a qualitative inspection compares our stylized images with that of other competing approaches: bilateral filter (BF) and Kuwahara filter (KF). Second, a systematic visual assessment was conducted on human observers to produce a quantitative score on the aesthetic appeal of the approaches.

4.1 Qualitative Visual Inspection

By visually examining the effects of region smoothing on the evaluated approaches, the proposed AD-SNN method appears to provide a middle ground between the BF and the KF. This is most noticeable from the sample image reviewed in Figure 3, where the strong preservation of gradients on the BF image renders it too closely to the original photograph. Meanwhile, the blurring sensation in the KF image causes obvious distortions to the shape of objects, which is a rather undesirable characteristic. Thus, the AD-SNN filter strikes a balanced compromise between the two extremes, while offering elegant and natural patch shadings. On a closer look (Figure 4), we can observe irregularities at a finer level. The color-quantized regions extracted after performing BF tends to produce incomplete edge contouring results after re-combination. Also, the blocky effect of the KF immensely deteriorates the quality of abstraction.

4.2 Quantitative Assessment

We conduct a systematic experiment using 20 popular images used in various works in literature [4,10,21]; 19 of which are taken from Philip Greenspun's photograph database[1] while one large size photograph of the Grand Canal of Venice was downloaded from the Internet. We handpicked the images based on their usage as examples in previous works, and also to provide a balanced

[1] `http://philip.greenspun.com/photography/`

Fig. 3. Edge-preserving region smoothing using bilateral filter *(top)*, Kuwahara filter *(middle)* and the proposed AD-SNN filter *(bottom)* on a photograph of the Grand Canal in Venice.

Fig. 4. Observations: Less than complete edge contouring on the color-quantized regions found by bilateral filter *(left pair)*, and obvious loss of details using Kuwahara filter *(right pair)*. The right-side image of each pair uses the proposed method.

Table 1. Mean and standard deviation of the relative amount of votes obtained each evaluated method through visual assessment

Method	Mean	Standard deviation
Bilateral	0.720	0.230
Kuwahara	0.408	0.162
AD-SNN	**0.876**	**0.098**

coverage of different types of photographs. There are 4 animal images, 6 single-person or group photographs, 4 landscape photos and 6 building/object photos. A total of 25 observers (10 of which have reasonable expertise in media arts and graphics, while the rest are novices) were given 10 sets of images; each set contains three stylized images (corresponding to the bilateral, Kuwahara and AD-SNN approaches) of a randomly selected photograph in random order. To alleviate possible bias from the edge style, we also randomly select between the MiG and FDoG (see Section 3.3) for the generated sets.

In contrast to our earlier work [20] which is primarily based on scoring or point-based feedback, we utilize a different methodology here, similar to that recently proposed in [19]. For each set of images, the observers were asked to cast two votes for the two most aesthetically pleasing stylizations. The relative amount of votes received by an approach for each participant is computed by $V_{rel} = V_{cast}/V_{total}$, where V_{cast} is the number of votes cast for the approach and V_{total} is the total number of image sets shown to the observer. The average voting score across all observers, $\overline{V_{rel}}$ is then computed to give the overall score.

Table 1 shows the overall result of our visual assessment experiment for the three evaluated filters. The AD-SNN filter provides the best artistic stylization, while its low standard deviation score highlights the consistency of opinion by the observers in favor of this method. It is unsurprising that the Kuwahara filter is visually unappealing due to the presence of artifacts discussed earlier. The bilateral filter remains competitive, but its high standard deviation offers some doubt over its attractiveness when applied to a highly variable set of 20 images and 2 different edge contours.

5 Conclusion

Concretely, we have presented a new technique called Anisotropic Diffusion Symmetric Nearest Neighbor (AD-SNN) filter for region smoothing within an image abstraction framework. With appropriate color quantization and a good choice of edge contouring, our method is able to produce aesthetically appealing images through a balanced measure of region smoothing and edge preservation. More importantly, blocky artifacts can be effectively mitigated by the effect of anisotropic diffusion. An extensive evaluation conducted shows greater appeal over other evaluated methods. In future work, the SNN filter can be further extended to an isotropic (circular) or anisotropic (elliptical) filter shape which will encourage smoother boundaries and corners in object regions.

References

1. Kyprianidis, J.E., Collomosse, J., Wang, T., Isenberg, T.: State of the 'Art': A taxonomy of artistic stylization techniques for images and video. IEEE Trans. on Visualization and Computer Graphics 19(5), 866–885 (2013)
2. Tomasi, C., Manduchi, R.: Bilateral filtering for gray and color images. In: Sixth International Conference on Computer Vision 1998, pp. 839–846. IEEE (1998)
3. Winnemöller, H., Olsen, S.C., Gooch, B.: Real-time video abstraction. ACM Transactions on Graphics (TOG) 25(3), 1221–1226 (2006)
4. Kyprianidis, J.E., Döllner, J.: Image abstraction by structure adaptive filtering. In: Theory and Practice of Computer Graphics, pp. 51–58 (2008)
5. Kang, H., Lee, S., Chui, C.K.: Flow-based image abstraction. IEEE Transactions on Visualization and Computer Graphics 15(1), 62–76 (2009)
6. Cong, L., Tong, R., Dong, J.: Selective image abstraction. Vis. Comput. 27(3), 187–198 (2011)
7. Zang, Y., Huang, H., Zhang, L.: Efficient structure-aware image smoothing by local extrema on space-filling curve. IEEE Transactions on Visualization and Computer Graphics (in press, 2014)
8. Kuwahara, M., Hachimura, K., Eiho, S., Kinoshita, M.: Processing of riangiocardiographic images. Digital Processing of Biomedical Images, 187–202 (1976)
9. Papari, G., Petkov, N., Campisi, P.: Artistic edge and corner enhancing smoothing. IEEE Transactions on Image Processing 16(10), 2449–2462 (2007)
10. Kyprianidis, J.E., Kang, H., Döllner, J.: Image and video abstraction by anisotropic kuwahara filtering. Computer Graphics Forum 28, 1955–1963 (2009)
11. Kyprianidis, J.E.: Image and video abstraction by multi-scale anisotropic kuwahara filtering. In: Proceedings of the ACM SIGGRAPH/Eurographics Symposium on Non-Photorealistic Animation and Rendering, pp. 55–64. ACM (2011)
12. Harwood, D., Subbarao, M., Hakalahti, H., Davis, L.S.: A new class of edge-preserving smoothing filters. Pattern Recognition Letters 6(3), 155–162 (1987)
13. Pietikäinen, M., Harwood, D.: Segmentation of color images using edge-preserving filters. In: Advances in Image Processing and Pattern Recognition, pp. 94–99 (1986)
14. Perona, P., Malik, J.: Scale-space and edge detection using anisotropic diffusion. IEEE Transactions on Patt. Anal. and Mach. Intell. 12(7), 629–639 (1990)
15. Gerig, G., Kubler, O., Kikinis, R., Jolesz, F.A.: Nonlinear anisotropic filtering of mri data. IEEE Transactions on Medical Imaging 11(2), 221–232 (1992)
16. Di Zenzo, S.: A note on the gradient of a multi-image. Computer Vision, Graphics, and Image Processing 33(1), 116–125 (1986)
17. Kang, H., Lee, S., Chui, C.K.: Coherent line drawing. In: Proc. of 5th Int. Symposium on Non-photorealistic Animation and Rendering, pp. 43–50 (2007)
18. Hertzmann, A.: Non-photorealistic rendering and the science of art. In: Proceedings of the 8th International Symposium on Non-Photorealistic Animation and Rendering, pp. 147–157. ACM (2010)
19. Alencar Júnior, J.D.O., de Queiroz, J., Gomes, H.: An approach for non-photorealistic rendering that is appealing to human viewers. In: 26th Conference on Graphics, Patterns and Images (SIBGRAPI), pp. 242–249. IEEE (2013)
20. Shahcheraghi, Z., See, J.: On the effects of pre-and post-processing in video cartoonization with bilateral filters. In: 2013 IEEE International Conference on Signal and Image Processing Applications (ICSIPA), pp. 37–42. IEEE (2013)
21. DeCarlo, D., Santella, A.: Stylization and abstraction of photographs. ACM Transactions on Graphics (TOG) 21, 769–776 (2002)