An Approach for Non-Photorealistic Rendering that is Appealing to Human Viewers

J. B. de O. Alencar Júnior, J. E. R. de Queiroz, H. M. Gomes
Centro de Engenharia Elétrica e Informática (CEEI)
Universidade Federal de Campina Grande (UFCG)
Campina Grande, PB, Brazil

***Abstract*—Non-photorealistic rendering is common to both Computer Graphics and Digital Image Processing, and encompasses a broad class of techniques for creating artistic images from scenes, geometric models or photographic images. In this context, this paper presents a framework designed to generate cartoon-like abstracted images, which combines techniques for region smoothing and segmentation, edge detection, adjustment of brightness/saturation and color harmonization. A subjective evaluation involving the output of related work was performed. Additionally, an experimental study involving human evaluators was conducted, in which the proposed approach was compared with five existing tools for image cartoonization. Statistical analysis performed on experimental data from one-way (1-F) ANOVA and Tukey-Kramer tests revealed significant differences in preference between results created by the proposed approach and those created by the competing approaches.**

***Keywords*-region smoothing; edge detection; non-photorealistic rendering; experimental evaluation.**

I. Introduction

Some authors argue that since the 1990s an extensive research has been produced concerning non-photorealistic rendering (NPR), in opposition to traditional forms of photorealistic representation [1]. According to Gooch [2], photorealistic rendering approaches are in general based on physical effects and properties of surfaces and materials, whereas NPR involves stylization and communication, by combining science and art. In photorealistic rendering, the effectiveness of the process is measured by the similarity between the resultant image and the original scene. On the other hand, NPR imagery can provide information about objects that may not be readily apparent in photographs or real life. Due to the subjective nature of the NPR results, strategies for accurately measuring their effectiveness do not exist [2], [3]. In addition, psychophysical evaluation and validation experiments are not trivial to devise, are time consuming to conduct, and require careful analysis [4].

Image abstraction is a NPR paradigm to provide stylized outputs that are easier to understand and more pleasant to the viewer's eyes [5]. Such a process involves two main subproblems: (i) *line extraction*, by which "significant" height discontinuities are captured and displayed; and (ii) *region smoothing*, by which "insignificant" height discontinuities are all removed. The combination of (i) and (ii) often results in a simplified, cartoon-like image [6]. The goal of *line extraction* is to find the most significant edges of the image. It results in edge maps which depict the image content on a higher abstraction level, so adding visually distinct edges to regions of high contrast further increases the visual distinctiveness of these scene locations [4]. On the other hand, the goal of *region smoothing* is to suppress unimportant information from the region interiors while preserving important shapes, i.e., region boundaries. By consequence, an image with smoothed regions is obtained, which implies a reduction of the color range of the original image. According to [5], region smoothing can be seen as the main problem in image abstraction.

Besides the aforementioned subproblems, some recent approaches (e.g., [7], [8]) include image color adjustment by using semi-automatic techniques for changing hue and brightness, with the aim to highlight some regions of interest in the image. The scope of this work involves the subproblems of region smoothing and color adjustment, being based on smoothing filters, segmentation algorithms and histogram adjustment as well as on color transformation studies. The goal of the proposed approach is to provide image abstraction from digital photographs. In this context, the subproblem of line extraction is not covered in detail in this work.

Despite the growing interest in digital image abstraction processes, new NPR approaches are still being proposed, and existing solutions still lack improvements toward producing a high level of quality in the abstractions, while requiring minimum human intervention. This fact motivated the present research, which contributes with an original approach for generating/creating cartoon-like abstracted images. Additionally, a novel extension of the bilateral filter was developed, which preserves sharpness over several iterations. More importantly, an inferential statistical analysis of the quality of the results as perceived by human subjects is provided.

The remainder of this paper is organized as follows. In next section, some related work within the NPR area is discussed, which served as source of inspiration for this research. In Section 3, the main steps and technical details of the proposed approach is presented. In Section 4, a two step evaluation of the proposed approach is presented, which includes: (i) a subjective visual inspection, including output images from related work; and (ii) an inferential statistical analysis comparing the outputs of the proposed approach with the results of a set of existing tools for image abstraction. Finally, in Section 5, some concluding remarks and proposals for further research are provided.

II. Related work

The goal of most robust image abstraction algorithms is to simplify the image visual content while preserving or even emphasizing most of the perceptually important information, making the message that it conveys more explicit and easily recognizable. Different automatic approaches for image abstraction have been proposed in the literature. Most of them often rely on image segmentation, and involve: (i) the reduction or removal of image texture, illumination and shading details (e.g., [9]); (ii) the simplification of image shapes and colors (e.g., [5]); (iii) the contrast decrease in low-contrast image regions and the contrast increase in higher ones (e.g., [10], [11]); and (iv) the use of salience maps for removing non-salient elements (e.g., [12], [13]). NPR algorithms also aim to produce artistic results that can be considered visually pleasant to the human eye [5], [7].

In spite of being mentioned as one of the main subproblems in image abstraction [6], line extraction has not been used as part of a number of approaches (e.g., [5], [14]). In those cases, the image is abstracted by means of region stylization, so this is the main subproblem for such approaches [5].

In traditional methods for color quantization the number of colors in the final image is typically reduced by error minimization [15]. In general, as those methods pick the most frequent colors, unimportant details can be preserved, such as large shaded regions, while relevant information can be removed, e.g., the contrast between small regions.

DeCarlo and Santella [12] developed a semi-automatic approach for converting images into line drawings with bold edges and large regions of constant color. They used Canny detector for edge extraction, and eye-tracking data to guide image abstraction based on mean shift segmentation [16].

Olmos and Kingdom [9] presented a rendering algorithm for image stylization that removes the soft shading from the image and gives objects extra definition through black outlines, by basing their approach on a model of the architecture of the human color vision system.

Wen et al. [7] conceived a semi-automatic NPR rendering approach in which a loose segmentation algorithm is applied to the input images. Drawing on the canvas is achieved by consulting a database of colors and applying a set of drawing rules defined by experts. Finally a form of line art, with colored strokes and filled areas, is achieved.

Medhi et al. [8] presented a tool which allows users to easily create cartoon-like images by using Gaussian blur to smooth the image and a connected-component analysis to identify the image regions. Since this method is not yet mature, the authors recommend performing the merge of the semantically related regions after the segmentation.

Winnemöller et al. [10] developed an automatic, real-time video and image abstraction framework. In their approach, the contrast reduction in low-contrast regions uses the bilateral filter [17] and artificially increased contrast in higher contrast regions with a difference-of-Gaussian (DoG) edges [18]. The authors used a separable approximation to the bilateral filter with a small kernel size and iterated the approximation to obtain a sufficiently large spatial support, achieving real-time frame rates. Color quantization is suggested as an additional step to convert the result in a cartoon-like image. The authors also suggest the use of a postprocessing technique to correct minor edge failures, and to highlight the final result as well.

Papari et al. [19] presented a nonlinear local operator that generalizes both the Kuwahara filter and the more general class of filters known in the literature as ”criterion and value filter structure”. Painting-like artistic effects were achieved from photographic images by using the operator, which is very similar to the bilateral filter used by Winnemöller et al. [10]. A similar approach was presented by Redmond and Dingliana [20], in which a technique is employed to generate multi-level abstractions of 3D textured urban scenes with the goal of highlighting the most important content in the image.

Kang and Lee [5] stated that the use of an image segmentation algorithm is a natural choice for image abstraction as it aggressively reduces the complexity of the scene while preserves important elements. Nonetheless, the region boundaries may need postprocessing, because a crude segmentation may result in a rather incomplete abstraction as the segmented regions still retain rough and complex boundaries. As a solution to that, the authors proposed to used a diretional flow-based diretional field together with an edge sharpness filter, so that the region colors and borders are simultaneously abstracted.

Kyprianidis and Döllner [11] and Kang, Lee and Chui [6] proposed extensions to the work of Winnemöller, Olsen and Gooch [10], more specifically to the region stylization using the Bilateral filter and to the contour detection by using the Difference of Gaussians (DoG) operator. In their experimental results, there was a substantial improvement in the quality of the generated contours.

Zhao et al. [13] added to the pipeline proposed by Kang, Lee and Chui [6] a technique for obtaining a saliency map, which is used to identify the most relevant image regions. This way, the level of abstraction can be controlled so that the most salient regions are emphasized and the less salient ones have their details smoothed. Also considering information on the shape of the image objects, Kyprianidis, Kang and Döllner [14] proposed an extension to the Kuwahara filter [19], which adapts this filter to the field of orientation vectors proposed by Kyprianidis and Döllner [11].

Kyprianidis [21] presents an extensive review of several automatic techniques for the creation of simplified and stylized images, including different variations of the Bilateral, Kuwahara and DoG filters. The author also suggests GPU programming for real-time processing. Along the above line, Zhao, Jin and Mao [22] present a method to automatically produce abstract photographs as well as real-time stylization of videos based on GPU. Their approach uses a direction field inspired by the Line Integral Convolution (LIC) [23] and a variant of the bilateral filter to perform region smoothing and shape abstraction simultaneously.

It is possible to characterize the reviewed work according to the types of region abstraction, color reduction and contour

extraction employed. For region abstraction, the most popular technique is bilateral filtering and its variants, as in the works of Winnemöller, Olsen and Gooch [10], Kyprianidis and Döllner [11], Kang, Lee and Chui [6], Zhao et al. [13], and Zhao, Jin and Mao [22]. An alternative to abstraction by smoothing is by region segmentation, as in the work of DeCarlo and Santella [12] that used the Mean Shift algorithm. With regards to color reduction, quantization is the most frequent technique. Finally, countour extraction can use traditional techniques, such as Canny and Sobel, but in more elaborated methods, such as the one presented by Zhao et al. [13], a combination of techniques is usually considered. The approach proposed in this paper adopted the following alternative techniques for region abstraction: (i) a novel extension of the Bilateral Filter for region smoothing; or (ii) a segmentation approach based on Mean Shift. For contour extraction, we propose DoG with ETF, although the evaluation of this aspect is out of the scope of this paper.

III. Proposed approach

The architecture of the proposed approach has been inspired by the work of DeCarlo and Santella [12], Winnemöller, Olsen and Gooch [10], Medhi, Patabandhige and Toyama [8], and Kang, Lee and Chui [6]. The architecture has two main phases: (i) image abstraction, which includes sub-phases of simplification and contour extraction; and (ii) color enhancement, which involves sub-phases of color harmonization and adjustment of brightness and contrast. The diagram presented in Figure 1 gives an overview on the architecture of the proposed approach, which is further explained in the following subsections.

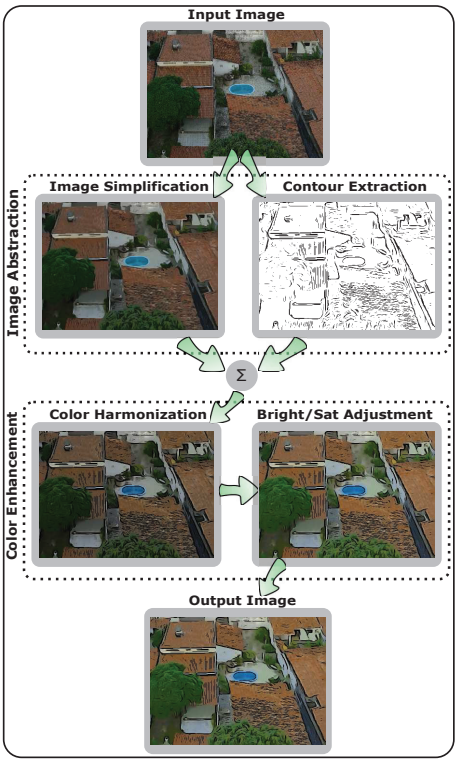


Fig. 1. Overall architecture of the proposed approach.

A. Image abstraction

The main goal of the first phase of the proposed approach is to generate a cartoon-like image. Its corresponding sub-phases are discussed in the following paragraphs.

1) Simplification by smoothing: The Bilateral filter was chosen for this sub-phase due to its edge-stopping properties. Because of the quality of its output and ease of use, the bilateral filter has caught the attention of the NPR comunity in the recent years [6]. A variant of this filter was originally employed in the work of Winnemöller, Olsen e Gooch [10], being also adopted in other more recent work (e.g. [6], [13], [11]). That variant, named Flow-based Bilateral (FBL) filter uses the information of edge orientation to strengthen the stylization and preserve the edges in the abstracted image. An original extension of the FBL filter - named AFBL - is detailed next.

Adaptive Flow-based Bilateral Filter (AFBL)

AFBL borrows the adaptive characteristic of the radial filtering proposed by Zhang and Allebach [24] in order to keep sharp edges after sucessive iterations of the filter. In the approach of Zhang and Allebach [24], the sigma and mean parameters of the Gaussian function in the radial domain are adapted to each pixel, in accordance to the LoG (Laplacian of Gaussian) operator [18]. This adaptation aims at reducing the influence of pixels that are in the transition defined by a border, thus avoiding the smoothing of these transitions, and promoting a contrast enhancement in those transition regions.

In order to add this feature to the FBL filter, the LoG calculation has been incorporated as well as the adaptive range Gaussian filter, as proposed by Zhang and Allebach [24], whose output $r(x, x_0)$ to a neighbor pixel x at position x_0 is given by Equation 1.

$$r(x, x_0) = \exp\left(-\frac{(f(x) - f(x_0) - \zeta(x_0))^2}{2\sigma_r^2(x_0)}\right) \quad (1)$$

The extension on the FBL filter has been performed only in the filtering that occurs in the direction of the edge gradient. Such extension also helped prevent the inhibition of stylization effects at the edges, which require high values for the sigma parameter. Finally, the functions ζ and σ_r that were obtained via training by Zhang and Allebach [24] have been replaced in our work by similar functions of easier adjustment, so that controlling the filter settings became easier, by allowing to increase or decrease the sigma value, as well as to regulate the displacement of the Gaussian function center (mean), which has a direct impact on the edge sharpness. The novel functions were defined according to the Equations 2 and 3.

$$\zeta(p) = m\ LoG(p) \quad (2)$$

$$\sigma_r(p) = \exp\left(-\frac{LoG^2(p)}{2\sigma_g^2}\right)(\sigma_{max} - \sigma_{min}) + \sigma_{min} \quad (3)$$

where the parameters m, σ_g, σ_{min}, σ_{max} as well as the sigma parameter used in the LoG operator are used to adjust the adaptive influence.

Pseudo-Quantization of the Illumination Channel

After obtaining the smoothed image, the cartoon-like abstracted image is generated by a color reducing process. The employed technique consists of a process of pseudo-quantization of the illumination channel, as adopted in previous studies (e.g. [6], [10], [11], [13]).

A quantization process is applied to the L* channel of the CIE L*a*b* (CIELAB) color space (see Equation 4), in which L* corresponds to the illumination and a* and b* are two cromatic components. This process is often called *pseudo-quantization* because the quantization occurs in the illumination channel, not in the crhominance channels.

$$Q(x) = q_{next} + \frac{\Delta q}{2}\tanh(\varphi.(f(x) - q_{next})) \quad (4)$$

This function partitions the L* channel into n equidistant levels (distance equals to Δq), $f(x)$ is the intensity of channel L* for pixel x, q_{next} is the level closest to $f(x)$, and φ controls the smoothness of the transitions between those levels.

2) Simplification through segmentation: The main approaches for image abstraction by using segmentation adopt the Mean Shift algorithm. It is relevant to note that, when abstracting/simplifying the image through segmentation, there is an emphasis on the edge regions, highlighting too much details of the original image. In order to correct this problem, DeCarlo and Santella [12] proposed to run a smoothing process along the image edges. Inspired on that work, the approach proposed in this paper uses the AFBL filter, previously presented, but using high values for the sigma parameter when applied in the tangential direction of the edges.

3) Contour extraction: In order to obtain image contours and to reduce the amount of high frequency details in this step, the input image is smoothed. Then, the FDoG technique was employed, which favours the generation of well defined contours from the output of an Edge Tangent Flow (ETF) filtering. ETF is a technique commonly adopted in recent work (e.g., [6], [11], [13]).

In the FDoG edge detection, contours are obtained by using the difference of two Gaussian filters in the gradient direction. Then, these contours are smoothed by using another Gaussian filter in the direction of the edge tangent. After that, the contours are merged with the abstracted image by a simple pixel multiplication operation. The next step is to perform color enhancement, which is an optional post-processing module, as discussed next.

B. Color enhancement

In this phase, hue, brightness and saturation parameters in HSV space are adjusted. The sub-phases of color harmonization and brightness/saturation adjustment are explained next.

1) Color harmonization: The goal of this sub-phase is to render the resulting image using a color set that is considered harmonic. Color harmony is described in terms of the position that the colors of an image occupy in a given color space, which provides a pleasing visual perception for humans [25].

In the proposed approach, the color harmonization technique presented by Cohen-Or et al. [25] is employed, which uses templates for spacial relations between color hues. This process may help reduce the amount of image colors since it groups hue values in sets of intervals previously defined by the harmonic templates. A minimization process is run to find the best harmonic template for a particular image. It should be noted that while making transformations in the image hues, guided by the predefined templates, there is the possibility that some objects may appear unnaturally colored (e.g., skin and sky tones may be depicted with unrealistic colors). Thus, human intervention may be needed to correct the hue angles of the selected harmonic template.

2) Brightness and saturation adjustment: The brightness and saturation adjustment aims to mimic the typical histograms of these channels in cartoon images, as observed by Ianeva [26]. According to Ianeva [26], the mean saturation and brightness of cartoon images are usually higher than those exhibited by photorealistic images. In this paper, gamma transformations the S and V channels have been applied. These were discovered by experimentation, as discussed in the next section. The result produced by this sub-phase is the final output of the proposed approach.

C. Some implementation details

The software implementing the proposed approach has been written in Microsoft Visual Studio 2005 C++. Three libraries were developed: (i) process control and file input/output; (ii) image abstraction; and (iii) color enhancement. The source code of the EDISON system [1] was used for the implementation of the Mean Shift algorithm. All experiments were run on a 2 GHz PC with 2 GB of RAM, under Microsoft Windows Vista.

IV. Evaluation

The proposed approach was evaluated by a two-step method. First, a visual confrontation and discussion of results involving an important part of the proposed approach (the AFBL filter, which is a smoothing-based approach for region abstraction), and prominent works found in the literature that contain a smoothing step. Second, a formal inferential statistical analysis (employing ANOVA and Tukey-Kramer methods) was conducted to test the appeal of resulting final images (with regards the opinion of human evaluators) produced by the proposed approach and by a selection of existing tools. This time, region abstraction step in the proposed approach was switched to segmentation based on Mean Shift in order to make a fair comparison with the evaluated tools, which were all based on segmentation. For practical reasons (unavailability of implementations for the techniques reviewed in Section II),

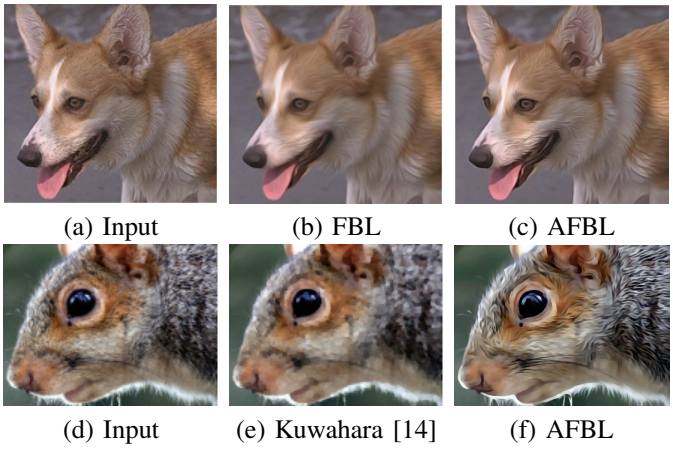
[1] EDISON stands for Edge Detection and Image SegmentatiON and is available for download at: http://coewww.rutgers.edu/riul/research/code/EDISON

the second experiment considered a set of commercial tools for cartoon generation, which had easy configuration and were readily available for use.

Hertzmann [27] stated that conducting experimental studies is one of the most difficult problems in the NPR area. Nonetheless, some methods designed to facilitate and standardize the validation of image abstraction processes can be found in the literature, such as psychological experiments related to: (i) ability of image memorization (e.g., [4]); (ii) capacity of image contents recognition (e.g., [4], [28]); (iii) visual attention (e.g., [29]); and (iv) subjective preference or appeal of resulting images (e.g. [28]). The methodology adopted in this work, which is discussed next, is aligned with the experiment categories (ii) and (iv) above mentioned.

A. Visual confrontation

In Figures 2(b)-(c), the results of the FBL and AFBL filters are shown, respectively. Besides the stylized effect produced by the FBL filter, the AFBL filter preserves the sharpness of representative high frequency details of the image (e.g., nose, eyes and tong of the dog). This is also observed when comparing the AFBL filter with a recently improved Kuwahara filter, as shown in Figure 2(e)-(f) (note the details in the neck and eyes regions of the animal, better preserved by the AFBL filter).



(a) Input (b) FBL (c) AFBL

(d) Input (e) Kuwahara [14] (f) AFBL

Fig. 2. Smoothing results. (a)-(c) comparison between FBL filter and proposed AFBL filter; (d)-(f) comparison between Kuwahara filter [14] and the proposed AFBL filter.

With regard to abstraction by segmentation, since since it was adopted the Mean Shift algorithm, which was also used in several other works as discussed in Section II, a visual confrontation was not included. However, it is relevant to mention the work of DeCarlo e Santella [12], which, besides the Mean Shift process, also employed a visual attention model trained from recordings of human eye movements to generate a strong abstraction of the background and less salient regions while retaining more details in the foreground and the more salient image regions.

Figure 3 allows a comparison between the results of the proposed method and the ones obtained by a sample of the reviewed techniques. The mandrill image in the last column had a local region magnified for better detail perception. By the results obtained using the proposed approach, it is possible to perceive a good level of region abstraction (e.g., color uniformity on large regions surrounded by darker and sharp contours), the main subject (objects, people, materials, etc) did not present noticeable imperfections, the semantic content of the original image is recognizable, and the overall cartoon-like effect is pleasing to the eye. Additional results are presented in the supplementary material.

B. Systematic evaluation

The goal of the evaluation method discussed in this section is to demonstrate that the proposed approach is adequate to its main purpose, i.e., the generation of cartoon-like images from photographs. For that purpose, an experiment has been designed to evaluate the visual appeal (as perceived by a set of human observers) of the approach results as well as the results of other available similar tools.

It should be pointed out that the selected tools did not provide detailed or specific information on the underlying techniques employed. However, these tools are well known and have sets of documented parameters for controlling the level of abstraction of the results, which is also true for the tool implementing the proposed approach. The selected tools were the following: (i) MSU Cartoonizer[2] v3.0, (ii) BeFunky Premium Cartoonizer[3], (iii) KusoCartoon General Cartoonizer[4], (iv) CaricatureSoft[5] v4.0.3557 and (v) LiangZhu[6] v4.08.

The following subsections report on the parameter calibration process for all involved tools and discusses the results of the tool comparison.

1) Parameter calibration: An initial parameter calibration process has been defined in order to identify the best parameter configurations for each of the compared software tools. The result of this process served as input to the remainder of the comparative experiment. Twenty volunteers ranked the output of the evaluated tools, when using different sets of parameters (each user ranked 576 image combinations). The scope of the calibration was limited to four different image categories: (i) images of animals; (ii) single person photographs; (iii) group photographs; and (iv) landscape/building/object photos.

Each category had only 4 representative images - 3 photographic images and 1 synthetic image, comprising 16 images. The syntetic image was designed as having low background information and simple objects with the purpose of facilitating the identification of best sets of parameters by humans. The photographic images came from the Philip Greenspun image database, which has been used in many other image abstraction work (e.g. [5], [11], [12]). The varied parameters considered adjustments in color transformation and saturation as well as

[2]http://www.compression.ru/video/cartoonizer/index_en.html
[3]http://www.befunky.com
[4]http://www.kusocartoon.com
[5]http://www.caricaturesoft.com/products/photo-to-cartoon.html
[6]http://www.liangzhuchina.com

(a) Input images

(b) Competing approaches

(c) Proposed approach

Fig. 3. Subjective evaluation of rendering results. The first line (a) contains the input images. The second line (b) contains the results of competing approaches (from left to right): Wen et al. (2006) [7]; Papari, Petkov and Campisi (2007) [19]; Kyprianidis and Döllner (2008) [11]; DeCarlo and Santella (2002) [12]; and Zhao, Jin and Mao (2013) [22]. The last line (c) contains the outputs of the proposed approach.

the amount of contours. Table I contains a list of the varied parameters per evaluated tool.

Tool	Parameter	Effect
BeFunky	Sketch Detail	Amount of Contours
	Color Detail	Amount of Colors
CaricatureSoft	Smooth	Image Smoothness
	Color Detail	Amount of Colors
	Sketch Detail	Amount of Countours
LiangZhu	Scalar	Amount of Colors
	Sleek	Image Smoothness
	Pencil Pressure	Amount of Contours
MSU	N. Edges	Amount of Contours
	N. Segments	Amount of Colors
	Saturation	Saturation Levels
KusoCartoon	Detail	Amount of Contours
	Color	Amount of Colors
	Saturation	Saturation Levels
Proposed Segmentation	Mean Shift (Radial σ)	Amount of Colors
	FDoG(Sigma 1)	Amount of Contours
	Saturation(Gamma)	Saturation Levels

TABLE I
CONTROL PARAMETERS OF THE EVALUATED TOOLS.

The amount of obtained parameter configurations was specific to each application, and aimed at prevent the use of configurations that generated too similar results, which would trouble the calibration process. The maximum number of configurations per tool was limited to 18. Apart from the BeFunky software, which had only 2 parameters to change, 3 parameters have been chosen in each software in other to produce the resulting variations.

For each of the 16 calibration images, 6 randomly selected parameter configurations were used to generate 6 resulting cartoon-like images to the human evaluator, who was asked to select from none to 2 images, when considering the following criteria (which were based on the concepts presented by Ianeva [26], Medhi, Patabandhige and Toyama [8], and Wallraven et al. [28]): (1) the cartoon has to be formed by uniformly colored areas highlighted by darker contour edges; (2) the main subject (objects, people, materials etc) should not present severe imperfections, such as missing parts as a result of the abstraction; (3) the semantic contents of the original image should be easily recognized in the cartoon-like result; and (4) the cartoon has to have a pleasing visual aspect.

After the data acquisition, a data verification process was conducted to check for inconsistencies in the results obtained from each calibrated software application. The data obtained was subsequently summarized and grouped for each participant to produce a collective score for each parameter setting.

The relative amount of votes received by a particular

parameter configuration c for each participant is defined by: $Q_{rel} = \frac{Q_{votes}}{Q_{total}}$, where Q_{votes} is the the amount of images generated by configuration c and selected by the participant, whereas Q_{total} is the total amount of images generated by configuration c and shown to the participant.

Before comparing the average of the Q_{rel} measure for the different parameter configurations of each tool, in order to choose the best ones, a null hypothesis of inexistence of significative statistical difference between those measures was formulated and verified. A 1-F ANOVA test with a 95% confidence level, commomly recommended in several statistical applications [30], [31], indicated significative difference between the averages. Then a Tukey-Kramer test allowed the selection of the groups of configuration parameters that were significantly different (higher) from the remaining, which were in turn employed in the experimental tool comparison.

Adopting a confidence level higher than 95% in this step made it difficult to see the separation between good parameter configurations and bad ones. On the other hand, in the tool comparison step, discussed next, adopting a higher confidence level (99.99%) has strenghened the credibility of the results ranking the different tools and the proposed method.

2) Tool comparison: In this step, sixty photographic images, obtained from the Phillip Greenspun's database, have been used (15 images for each of the four categories discussed in the previous subsection (parameter calibration). Only one cartoon-like image was generated by each of the compared tools, but using the best parameter configuration obtained above. For cases where more than one parameter configuration was considered best, a random selection was performed.

A distinct group of 31 test users participated in the comparison reported here. The image evaluation criteria presented to the participants of this experiment were the same of the parameter calibration step. Similarly, for each input image, six cartoon-like abstractions were presented and the participants were asked to select up to two results that were best aligned with the evaluation criteria.

It is important to note that in this step, differently from the first one, each of the cartoon-like abstraction was produced by a different tool. Moreover, in both steps of the tool comparison experiment, the strategy employed for presenting the image abstractions did not allow the identification of the tool used to produce a particular resulting image. Moreover the position of the abstractions in each evaluation screen were randomly varied throughout the experiment.

After data acquisition, the average relative votes received by each tool was calculated similarly to the parameter calibration step. The total number of images produced by the tools corresponded exactly to the number of fotographic images adopted (60 images). The average values and standard deviations of the Q_{rel} votes obtained by each tool are presented on Table II, which shows a very high average for the proposed approach when compared to the other tools.

Similarly to the previous step, a 1-F ANOVA test was performed in order to verify the presence of statistically significant differences in the average values obtained. The result of this test, when considering a confidence level of 99,99% revealed that indeed such differences are significant.

Tool	Average	Standard deviation
BeFunky	0.4387	0.2099
Caricaturesoft	0.0306	0.0365
LiangZhu	0.0866	0.0915
MSU Cartoonizer	0.0645	0.0904
KusoCartoon	0.2871	0.1752
Proposed	0.7086	0.1487

TABLE II
MEAN AND STANDARD DEVIATION OF THE RELATIVE AMOUNT OF VOTES OBTAINED BY EACH EVALUATED TOOL.

A Tukey test (99.99% confidence level) allowed to classify the average value of the proposed approach in a distinctive group, according to the average relative number of votes. From the above analyses, it is possible to conclude that the proposed approach was considered better than all the other evaluated tools, which highlights its adequacy to produce cartoon-like abstractions that are appealing to human observers, when considering the evaluation criteria.

A detailed set numerical results for the 1-F ANOVA and Tukey tests performed in the parameter calibration and tool comparison analyses are available in the online supplementary material. Snapshots of the the parameter calibration and tool comparison interfaces, as well as additional resulting images of the proposed method are also included in that material.

V. FINAL CONSIDERATIONS

The main contributions of this work were threefold: (i) the proposition of a new method for image abstraction; (ii) the development of an original extension of the bilateral filter with the purpose of maintaining edge sharpness after sucessive iterations; and (iii) the presentation of a statistical systematic evaluation methodology for comparing different approaches to the problem, which highlighted a preference for the results produced by the proposed approach.

While in the photorealism the effectiveness of automatic processes can be measured by verifying the similarity between images produced by different computational methods and ground truth images, nonphotorealism lacks an unified and objective methodology to evaluate the quality of underlying processes according to a set of criteria. In many scientific work in this area, the validation is limited to visual comparison of the obtained results with the ones of competing techniques [27], which was confirmed in the present research when reviewing the related work (e.g., [6], [14], among others).

More specifically, it was possible to conclude from the conducted experiments and from the statistical processing of the acquired data, that the images generated by the proposed approach presented a high visual appeal to human observers when compared to the results of a set of commercial tools, under the same experimental conditions, which includes a thorough parameter calibration process.

Performing user evaluation experiments has shown to be a prohibitively lengthy task for the users if using a large

set of parameters to calibrate or a large number of images to evaluate. Therefore, only a representative set of parameter configurations per tool was considered. Moreover, comparison with competing academic techniques was only possible in a subjective way. A systematic evaluation with a parameter calibration step was not feasible with those works since this would require implementations of the methods, but until now we could not find publicly available code or sufficient technical details for a faithful implementation.

The current implementations of the AFBL filter and the Mean Shift based abstraction investigated in this paper are unsuited for real-time processing. They take an average of 5 and 10s, respectively, to process a 512x512 image, when using typical parameter settings. A GPU-based parallel implementation might drastically reduce those times. Furthermore, the AFBL filter might fail under certain parameter settings. Therefore, studying the effects of different ζ and σ_r functions is left as future work.

We plan to analyze the effect of incorporating an FDoG filter to the AFBL processing flow for additional improvement. Moreover, we see that visual attention or the computation of saliency maps (e.g. [12], [13]) may help to automatically adjust the abstraction level in terms of the relevance of the image regions. Alongside this line, learning to generate NPR renderings by taking into account human reactions to the renderings is an interesting research path. The optional color enhancement modules, not included in the objective experimental evaluation will be object of further investigation.

ACKNOWLEDGMENT

Research supported by HP and CAPES. Images in Figures 2(a), 4(a3-4) and input images 1, 4, and 7 in the supplementary material are courtesy of Philip Greenspun. Thanks to Jan Eric Kyprianidis, Fang Wen, Giuseppe Papari, Doug DeCarlo, Xiaoyang Mao, Adriana Olmos, John Dingliana and Holger Winnemöller for providing result images of their methods.

REFERENCES

[1] J. Thoma, "Non-photorealistic rendering techniques for real-time character animation," Master's thesis, RWTH Aachen University, 2002. [Online]. Available: http://www.hpi.uni-potsdam.de/fileadmin/hpi/Forschung/Publikationen/Dissertationen/Nienhaus/Diss_Nienhaus.pdf

[2] B. Gooch, "Artistic vision: Automatic digital painting using computer vision algorithms," Master's thesis, The University of Utah, USA, 2001. [Online]. Available: http://www.researchgate.net/publication/33797599_Artistic_vision__automatic_digital_painting_using_computer_vision_algorithms

[3] T. Strothotte and S. Schlechtweg, *Non-Photorealistic Computer Graphics: Modeling, Rendering and Animation*. San Francisco: Morgan Kaufmann, 2002.

[4] H. Winnemöller, "Perceptually-motivated non-photorealistic graphics," Ph.D. dissertation, Northwestern University - Illinois, 2006. [Online]. Available: http://holgerweb.net/PhD/Research/papers/diss_winnemoeller06.pdf

[5] H. Kang and S. Lee, "Shape-simplifying image abstraction," *Computer Graphics Forum*, vol. 27, no. 7, pp. 1773–1780, 2008.

[6] H. Kang, S. Lee, and C. Chui, "Flow-based image abstraction," *Visualization and Computer Graphics*, vol. 15, no. 1, pp. 62–76, 2009.

[7] F. Wen, Q. Luan, L. Liang, Y.-Q. Xu, and H.-Y. Shum, "Color sketch generation," in *International Symposium on Non-photorealistic Animation and Rendering*, 2006, pp. 47–54.

[8] I. Medhi, C. R. Patabandhige, and K. Toyama, "Cartoon generation for text-free user interfaces," in *Int. Workshop on Computer Vision Applications for Developing Regions*, 2007.

[9] A. Olmos and F. A. A. Kngdom, "Automatic non-photorealistic rendering through soft-shading removal: a colour-vision approach," in *International Conference on Vision, Video and Graphics*. Eurographics Association, 2005, pp. 203–208.

[10] H. Winnemöller, S. Olsen, and B. Gooch, "Real-time video abstraction," *ACM Transactions on Graphics*, vol. 25, pp. 1221–1226, 2006.

[11] J. E. Kyprianidis and J. Döllner, "Image abstraction by structure adaptive filtering," in *Proc. EG UK Theory and Practice of Computer Graphics*. Eurographics Association, 2008, pp. 51–58.

[12] D. Decarlo and A. Santella, "Stylization and abstraction of photographs," in *Annual Conference on Computer Graphics and Interactive Techniques*. San Antonio: New York: ACM, 2002, pp. 769–776.

[13] H. Zhao, X. Mao, X. Jin, J. Shen, F. Wei, and J. Feng, "Real-time saliency-aware video abstraction," *The Visual Computer*, vol. 25, pp. 973–984, 2009.

[14] J. E. Kyprianidis, H. Kang, and J. Döllner, "Image and video abstraction by anisotropic Kuwahara filtering," *Computer Graphics Forum*, vol. 28, no. 7, pp. 1955–1963, 2009.

[15] V. Hautamäki, "Efficient color quantization by hierarchical clustering algorithms," Master's thesis, University of Joensuu, 2005. [Online]. Available: ftp://cs.joensuu.fi/franti/papers/2005_MSc_Hautamaki_Ville.pdf

[16] D. Comaniciu and M. P., "Mean shift: a robust approach toward feature space analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 5, pp. 603–619, 2002.

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

[18] D. Marr and E. Hildreth, "Theory of edge detection," *Proceedings of the Royal Society of London - Series B, Biological Sciences*, vol. 207, no. 1167, pp. 187–217, 1980.

[19] G. Papari, N. Petkov, and P. Campisi, "Artistic edge and corner enhancing smoothing," *IEEE Transactions on Image Processing*, vol. 16, no. 10, pp. 2449–2462, 2007.

[20] N. Redmond and J. Dingliana, "Adaptive abstraction of 3d scenes in real-time," in *Eurographics Conference*, Prague, 2007, pp. 77–80.

[21] J. E. Kyprianidis, "Adaptive stylization of images and video," Ph.D. dissertation, University of Potsdam, 2013. [Online]. Available: http://opus.kobv.de/ubp/volltexte/2013/6410/pdf/kyprianidis_diss.pdf

[22] H. Zhao, X. Jin, and X. Mao, "Real-time directional stylization of images and videos," *Multimedia Tools and Applications*, vol. 63, no. 3, pp. 647–661, 2013.

[23] B. Cabral and L. C. Leedom, "Imaging vector fields using line integral convolution," in *Annual Conference on Computer Graphics and Interactive Techniques*, 1993, pp. 263–270.

[24] B. Zhang and J. P. Allebach, "Adaptive bilateral filter for sharpness enhancement and noise removal," *IEEE Transactions on Image Processing*, vol. 17, pp. 664–678, 2008.

[25] D. Cohen-Or, O. Sorkine, R. Gal, T. Leyvand, and Y.-Q. Xu, "Color harmonization," in *Annual Conference on Computer Graphics and Interactive Techniques*. Boston: New York: ACM, 2006, pp. 624–630.

[26] T. Ianeva, A. de Vries, and H. Rohrig, "Detecting cartoons: a case study in automatic video-genre classification," in *International Conference on Multimedia and Expo*, vol. 1. Baltimore: IEEE, 2003, pp. 449–452.

[27] A. Hertzmann, "Non-photorealistic rendering and the science of art," in *International Symposium on Non-Photorealistic Animation and Rendering*. Annecy: New York: ACM, 2010, pp. 147–157.

[28] C. Wallraven, H. H. Bülthoff, D. W. Cunningham, J. Fischer, and D. Bartz, "Evaluation of real-world and computer-generated stylized facial expressions," *ACM Transactions on Applied Perception*, vol. 4, no. 3, pp. 16:1–16:24, 2007.

[29] A. Santella, "The art of seeing: Visual perception in design and evaluation of non-photorealistic rendering," Ph.D. dissertation, Rutgers University, 2005. [Online]. Available: http://www.research.rutgers.edu/~asantell/thesis.pdf

[30] S. Kimothi, *The Uncertainty of Measurements: Physical and Chemical Metrology : Impact and Analysis*. Asq Press, 2002.

[31] A. J. Wheeler and A. R. Ganji, *Introduction to Engineering Experimentation*. Pearson Education, Limited, 2010.