

Systematic Evaluation of Image-Based NPR Applied to Video

Alberto F.F. de Barros, Jose E.R. de Queiroz and Herman M. Gomes

Department of Systems and Computing, Federal University of Campina Grande, Campina Grande, PB, Brazil

Article history

Received: 16-12-2019

Revised: 24-03-2020

Accepted: 28-04-2020

Corresponding Authors:

Alberto F.F. de Barros
Department of Systems and
Computing, Federal University
of Campina Grande, Campina
Grande, PB, Brazil
Email: albertofagner.cav@gmail.com

Abstract: In this article, it was studied the effect of state-of-the-art approaches for Non-Photorealistic Rendering (NPR), originally developed for static images, when applied to videos. Six criteria were objectively evaluated (using a 5-point Likert scales): Simplicity, content preservation, resemblance to the original content, sharpness, temporal coherence and subjective satisfaction. Sixty participants and a dataset of 30 videos of different genres were used in the experiments. Each participant voted 15 pairwise video comparisons. Three approaches were considered: Anisotropic Kuwahara, Coherent Enhancement and Extended Difference of Gaussians (XDoG). The results of the study, under 99% statistical confidence, indicated that the Anisotropic Kuwahara approach achieved the best results for most of the criteria. Coherent Enhancement and XDoG approaches came next. These results indicate that the selected approaches were considered satisfactory, when applied to video, for most considered criteria.

Keywords: Non-Photorealistic Rendering, Video Abstraction, Image Filtering, Comparative Evaluation

Introduction

Non-photorealistic rendering is a research area that lies at the intersection of Computer Graphics, Digital Image Processing and Art. From photorealistic images and videos, the goal is to generate stylized and abstract representations of reality, with the purpose of visually communicating an idea or concept to the observer. NPR examples vary from the generation of cartoon-like representations from images and videos (Kang *et al.*, 2015; Wenhua *et al.*, 2015) to renderings in which the image or video is converted into specific styles, such as watercolor and oil-on-canvas (Zhang *et al.*, 2013).

When applied to video, NPR techniques that are originally designed for processing images may produce some visualization issues, such as (i) scintillation and (ii) loss of space-time coherence, as discussed by Chen *et al.* (2012). Nearby frames of a video tend to be strongly correlated, but usually present subtle variations because of illumination changes, camera or object motion. These issues are related to instability in the results of NPR techniques under such sources of variation.

Traditional researches in the NPR area are usually validated by subjective/visual experiments, as in the research of Kang *et al.* (2015), Wenhua *et al.* (2015), Zhang *et al.* (2013), Baugh and Kokaram (2010), Benard *et al.* (2012), Borawski (2014), Cao *et al.* (2011), Chen *et al.* (2012), Gangopadhyay *et al.* (2016) and Wang *et al.* (2017). However, purely subjective/visual evaluation, without the use of external objective criteria and the lack of a representative

sample of participants, weakens the statistical validity of the analyses performed.

There is very little research that has adopted an evaluation approach based on polling the subjective opinion of a sample of potential NPR consumers, when considering a set of criteria measured according to a given numerical scale. An example of such work was performed by Arruda (2014).

In this context, this article aims to identify whether it is possible to employ static NPR to generate NPR videos that are pleasing to observers and if it is also possible to obtain statistically valid results.

The main contribution of this research is the implementation of an experimental statistical evaluation of the results of 3 state-of-the-art static NPR approaches when applied to video. It was used in the experiments a sample of 60 participants and 30 videos from different genres.

The structure of the remaining sections are as follows. The next section contains a description of the proposed evaluation methodology. It is followed by a description of the evaluated approaches. Then, the obtained results are shown and discussed. Finally, the last section contains the conclusions and proposals of future work.

Evaluation Methodology

The experimental evaluation methodology consisted of polling the opinion of human observers about three different NPR versions of videos. It is inspired on the methodology proposed by Arruda (2014).

The same criteria adopted by Arruda (2014) were considered, which take into account the characteristics of cartoons, according to the research of Ianeva *et al.* (2003), Medhi *et al.* (2007) and Junior *et al.* (2013).

Ianeva *et al.* (2003) defined cartoons as images composed of a reduced amount of colors, delimited by black borders. Medhi *et al.* (2007) and Junior *et al.* (2013) corroborated the definitions provided by Ianeva *et al.* (2003) by emphasizing the following cartoon characteristics:

- (I) Cartoons must contain regions of uniform color, delimited by dark outlines
- (II) The theme of the scene must not have serious flaws or imperfections, e.g., missing parts
- (III) The semantic content of the original image must be recognized in the rendered image
- (IV) The color regions must have high saturation
- (V) The background of the scene should be viewed in a simplified way and with little detail
- (VI) The cartoon must maintain a pleasant visual aspect

Taking into account the previous statements, Arruda (2014) developed the following criteria, which are adopted in the present article:

- (I) Simplicity: Uniformity of colored areas with darker contours
- (II) Preservation of content: Absence of missing parts or flaws that de-characterize the video theme
- (III) Similar to the original content: The cartoon-like video should have the same semantic content as the original video, being easily recognizable
- (IV) Sharpness: The cartoon-like video should have the theme presented in a clear way
- (V) Temporal Coherence: No flickering or unwanted artifacts present in the cartoon-like video
- (VI) Subjective Satisfaction: The cartoon-like video should have a visual appearance that satisfies the participant

The six criteria adopted for the evaluation of the videos were measured by 6 questions. Therefore, each participant voted in 6 scales for each of the 15 videos (each having an average duration of 10 seconds), totaling 60 polls per test (6 scales 15 videos). The questions that summarize all the six aforementioned criteria are presented next, respectively:

- (I) The cartoon-like video presents homogeneous colorful areas with dark contours
- (II) The cartoon-like video does not show flaws (missing parts) on some objects

- (III) The main idea conveyed by the original video was preserved in the cartoon-like video
- (IV) The objects and scenes portrayed in the cartoon-like video are sharp and easy to perceive
- (V) The cartoon-like video does not present frequent brightness and/or color variations (flashing effect)
- (VI) The cartoons-like video has a nice visual aspect according to the observer

The evaluation carried out by the participants used a 5-point Likert scale ranging from totally disagree to totally agree. According to Hogg *et al.* (1977), a large sample size ($n > 30$) was considered in this study. So, sixty (60) participants were recruited, with the aim of allowing more robust statistical tests.

The test videos were chosen from the derf's collection video dataset¹, made available by the Xiph.Org Foundation, a non-profit corporation dedicated to promote the development of the internet multimedia area with free/open protocols and software. Figure 1 contains the first frame of each used video.

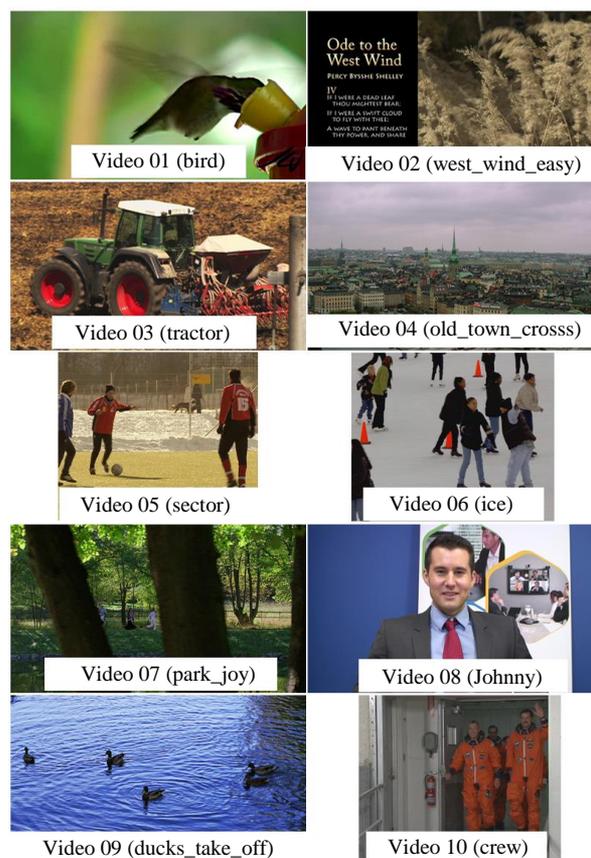


Fig. 1: First frame from each test video, which are labeled as Video 01, Video 02, ..., Video 10 in this research. The original derf's Collection label pears within parenthesis

¹<https://media.xiph.org/video/derf/>

The study included volunteers of both genders aged 18 and over, which facilitated recruiting volunteers from a population of university students. The sample was not stratified by gender or age. It was decided taking into account, on one hand, the widespread dissemination of access and consumption of digital video contents of the most varied genres available on the World Wide Web. On the other hand, it is a fact that digital video users come from all ages and genders. Users virtually come from all age groups, social and educational levels.

Each test session comprised the viewing 15 pairs of videos (out of 30 available videos generated), with an average duration of 10 sec. Each video pair consisted of the original source and the resulting rendering of an approach. The distribution of the videos to the participants was done in an equitable way. The transition between a pair of videos to the next pair only occurred after collecting the votes for all six evaluated criteria. Audio was not played to avoid distractions and help participants to focus on the assessment task.

After the voting sessions, it was performed the data collection and analysis. The normality of data was tested with the Shapiro-Wilk test to choose an appropriate inferential statistical test. The test revealed the experiment data did not follow a normal distribution, thus, the Kruskal-Wallis method was chosen. This method was applied to each evaluated criterion to verify the existence of statistically significant differences between the evaluated approaches, which was confirmed. It was then necessary to apply the Wilcoxon test on pairs of approaches in order to determine the approaches that were best evaluated in relation to each one of the six criteria.

Chosen Approaches

The present research involved evaluating experimental statistical results of NPR approaches when applied to video. Three open source state-of-the-art static NPR approaches were chosen. For comparison, their parameters were set to achieve similar rendering results on videos of the same database. The chosen approaches are described as follows.

The anisotropic Kuwahara filter by Kyprianidis *et al.* (2009) is based on a generalization of the Kuwahara filter by Pham (2006), in which the weighting functions defined over sectors of a disc are replaced by weighting functions defined over ellipses, to minimize artifacts. Shape, scale and orientation of these ellipses are adapted to the local structure of the input, to preserve and emphasize directional image features.

Required information about the local orientation as well as a measure for the anisotropy are derived from a smoothed structure tensor. By considering $f: R^2 \rightarrow R^3$ as their input image and $x_0 \in R^2$ as a point input image, Kyprianidis *et al.* (2009) let ϕ be the local orientation and A be the anisotropy at x_0 . The method proposed by Pham (2006) was used by Kyprianidis *et al.* (2009) to define an elliptical filter shape. So to adjust the eccentricity in terms of the amount of anisotropy, Kyprianidis *et al.* (2009) define S according to Equation (1):

$$S = \begin{pmatrix} \frac{\alpha}{\alpha + A} & 0 \\ 0 & \frac{\alpha + A}{\alpha} \end{pmatrix} \quad (1)$$

Where:

- $\alpha > 0$ is a tuning parameter
- A is the anisotropy

Figure 2 contains the first frame of test videos after applying anisotropic Kuwahara filter.

The core of the Coherence-Enhancing Filter (Kyprianidis and Kang, 2011) is the adaptive line integral convolution combined with directional shock filtering smoothing. The number of iterations of the filter controls the strength of the abstraction. According to the authors, a flow field derived from the structure tensor guides both smoothing and sharpening processes. The first one regularizes directional image features while the shock filter provides the second one. The local structure is computed twice for each iteration, namely before the adaptive smoothing and before the sharpening.

In each iteration, the structure tensor is calculated from the previous one, except for the first iteration, where a relaxation of the structure tensor is performed. The authors argue that their solution produces a result closer to a piecewise constant image with large smooth or even flat regions. Finally, edges are smoothed by using the adaptive smoothing with a small standard deviation. The structure tensor used by Kyprianidis and Kang (2011) is the rotational symmetric derivative filter (Jähne *et al.*, 1999; Weickert and Scharr, 2002), as given by Equation (2):

$$D_x = \frac{1}{2} \begin{pmatrix} p_1 & 0 & -p_1 \\ 1-2p_1 & 0 & 2p_1-1 \\ p_1 & 0 & -p_1 \end{pmatrix}, D_y = D_x^T \quad (2)$$

where, $p_1 = 0.183$.

Figure 3 contains the first frame of each used video after applying Coherence-Enhancing filter.

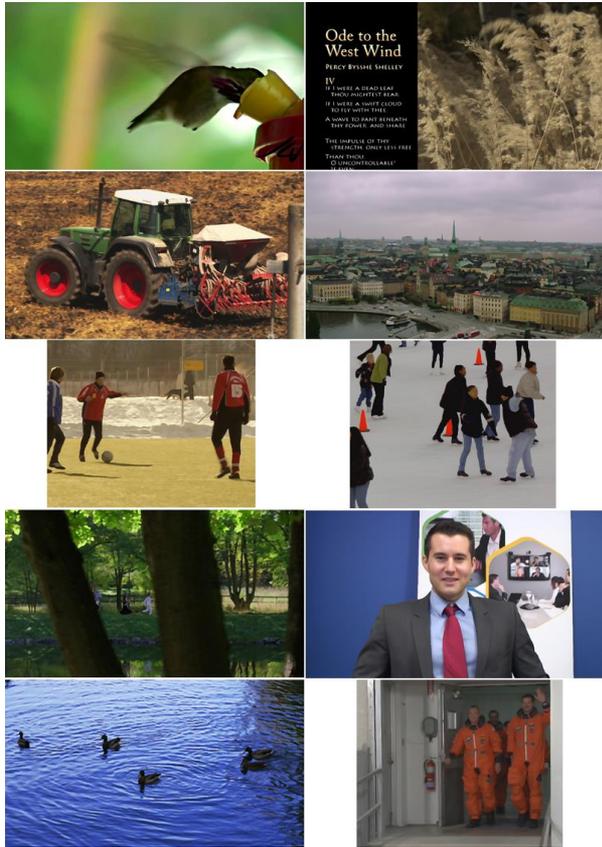


Fig. 2: First frame from each video using anisotropic filtering

In essence, Difference of Gaussians (DoG) is a grayscale image filtering approach that consists of processing an original image twice with two different Gaussian filters, after which one is subtracted from the other (Kyrianiadis *et al.*, 2012; Akenine-Möller *et al.*, 2018). The DoG filter generates a two-tone edge image from the original image (I) by describing which high-frequency details have to be added to the low-pass filtered image, as given by Equation (3):

$$G(\sigma) * I = G(k \cdot \sigma) * I + D(\sigma, k) * I \quad (3)$$

Where:

- G = The Gaussian filter with standard deviation σ
- K = A tuning constant to change the standard deviation.
- $*$ = The convolving operation
- D = Low pass operation parametrized by σ and k

In such a way, the sign of the DoG filter's response describes whether capturing the shape and structure of any nearby edges requires making each pixel darker or brighter than most of its neighbors.

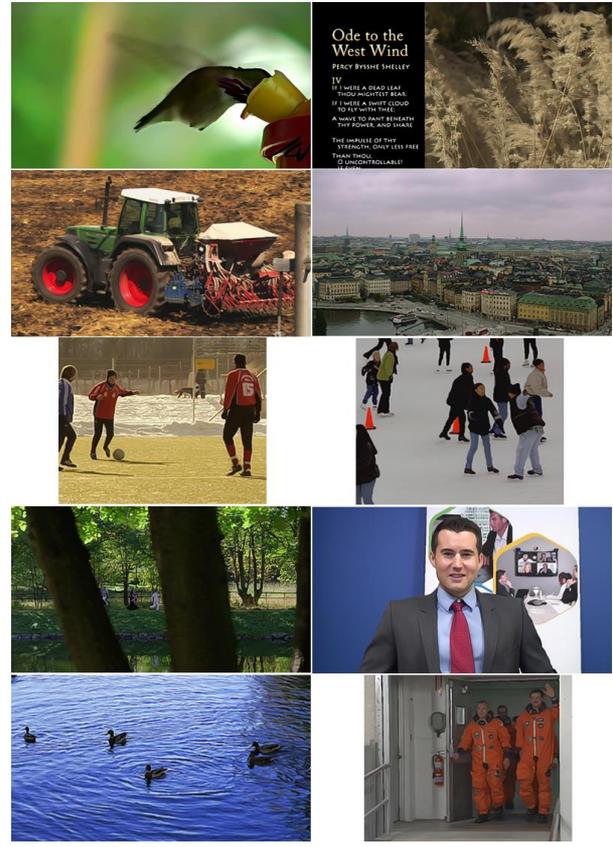


Fig. 3: First frame from each video using coherence filtering

An extension of DoG filtering operator, named Extended Difference of Gaussians (XDoG), is described by Winnemöller (2011). The XDoG filtering operator, given by Equation (4), takes an input image I and convolves it with two separate Gaussian filters, with standard deviations σ and $k \cdot \sigma$, as given by Equation (4).

$$D_x(\sigma, k, \tau) = G(\sigma) - \tau \cdot G(k \cdot \sigma) \quad (4)$$

Where:

- D_x = DoG result in x coordinates
- σ = Controls the scale
- τ = Controls tone mapping

After the previous operation, a sigmoidal function is applied to the difference of these two images, as specified by Equation (5):

$$E_x(\sigma, k, \tau, \varepsilon, \varphi) = \begin{cases} 1, & \text{if } D_x(\sigma, k, \tau) < \varepsilon \\ 1 + \tanh(\varphi \cdot D_x(\sigma, k, \tau)), & \text{otherwise} \end{cases} \quad (5)$$

Where:

- ε shifts the detection threshold to control sensitivity (large values make the edge detection less sensitive, while small values increase detection sensitivity)
- τ changes the relative weighting between the larger and smaller Gaussians, thereby affecting the tone-mapping response of the operator
- \tanh is a function that adjusts the transition between border and non border values

- φ controls transition steps

Figure 4 contains the first frame of each used video after applying the XDoG filter.

Public implementations of the approaches considered in this comparative study can be downloaded from the corresponding project websites (Table 1).

The metadata for each video used in the evaluation can be seen in Table 2.

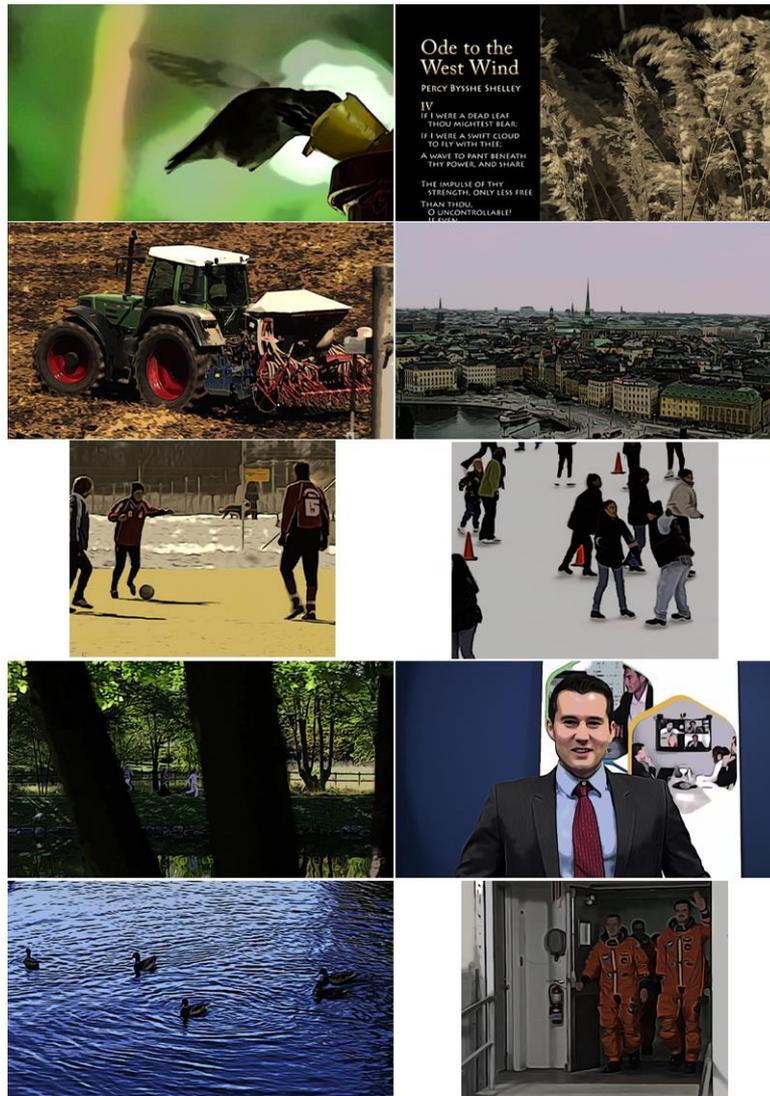


Fig. 4: First frame from each video using XDoG filtering

Table 1: Chosen approaches

Approaches	Hyperlink
Kuwahara by Kyprianidis <i>et al.</i> (2009)	https://code.google.com/archive/p/gpuakf/downloads
Coherence by Kyprianidis and Kang (2011)	https://code.google.com/archive/p/cefabs/downloads
XDoG by Winnemoller (2011)	https://code.google.com/archive/p/xdog-demo/downloads

Results and Discussion

The votes cast after the 60 trials are summarized in Fig. 5 to 10. The x-axis of these figures represents the evaluated approaches. The y-axis corresponds to the degree of agreement (Likert scale) of the participants with the evaluated criteria, varying from where 1 (Fully Disagree) to 5 (Fully Agree).

In the Fig. 5, the XDoG approach had a better evaluation for the first criterion.

In the Fig. 6, Kuwahara was better evaluated for the second criterion.

In the Fig. 7, Kuwahara had a better evaluation for the third criterion.

In the Fig. 8 and 10, regarding the fourth and sixth criteria, respectively, the XDoG approach had a less favorable evaluation compared to the other two approaches.

In the Fig. 9, regarding the fifth criterion, Kuwahara was better evaluated.

The data gathered from the participants of the experiment was processed according the methodology presented next.

First, the Shapiro-Wilk test was applied to each of the criteria in order to verify the normality of the data. The

p-value for all criteria was less than $2, 2 \times 10^{-16}$, which indicates that none of the data distributions fit a normal distribution, considering a confidence level of 99%.

In order to test the existence of possible statistical differences between the evaluated approaches, six (6) Kruskal-Wallis tests were applied (one test for each criterion, 3 approaches compared). The p-value resulting from each of the tests, when considering a confidence of 99%, was lower than $2, 2 \times 10^{-16}$. Therefore, it concludes that there are significant statistical differences between the approaches in the different evaluated criteria.

Table 2: Metadata information of video database

Video	Dimensions	Duration(s)	FPS
Video 01	1280×720	14	60
Video 02	1920×1080	19	29
Video 03	1920×1080	27	25
Video 04	1280×720	9	50
Video 05	704×576	9	60
Video 06	704×576	7	60
Video 07	1280×720	9	50
Video 08	1280×720	10	60
Video 09	1280×720	9	50
Video 10	704×576	9	60

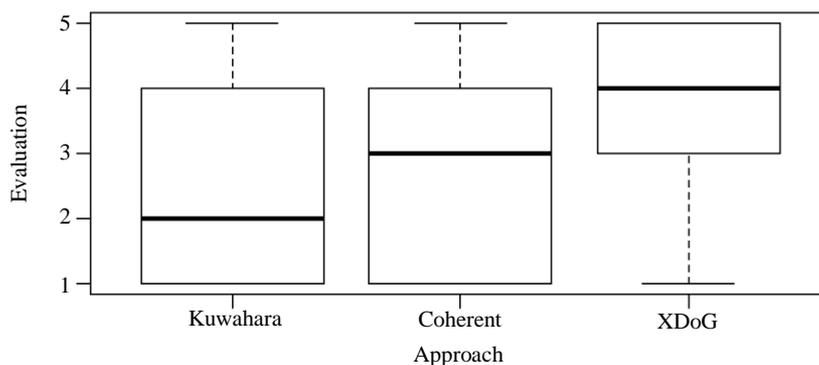


Fig. 5: Boxplots of the votes cast for the first criterion, distributed among the evaluated approaches

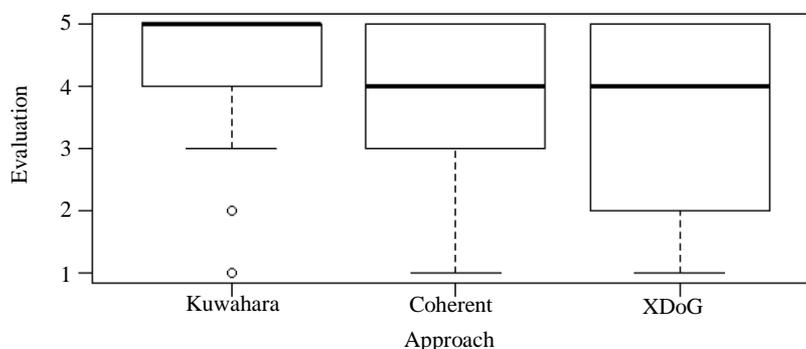


Fig. 6: Boxplots of the votes cast for the second criterion, distributed among the evaluated approaches

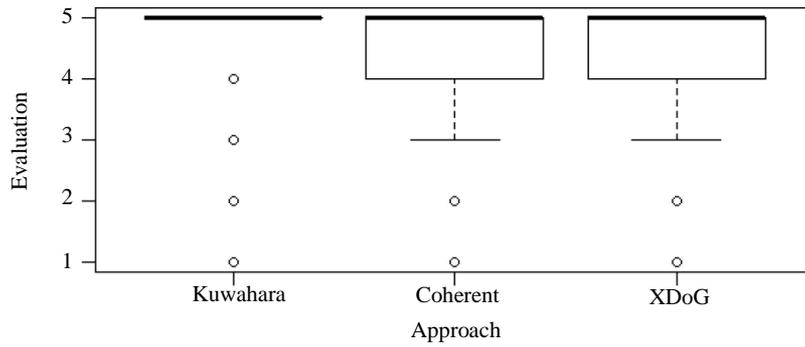


Fig. 7: Boxplots of the votes cast for the third criterion, distributed among the evaluated approaches

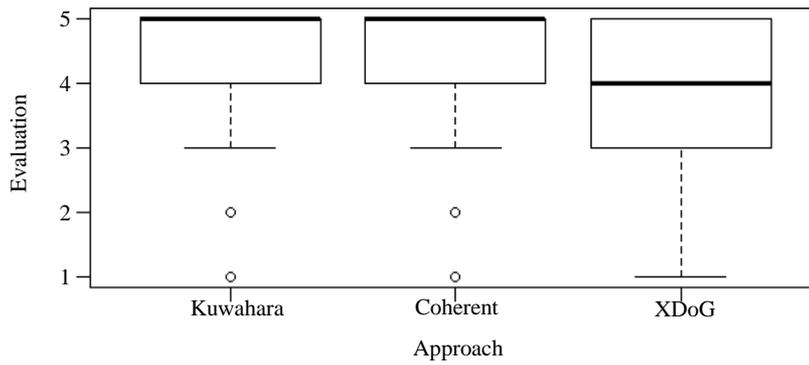


Fig. 8: Boxplots of the votes cast for the fourth criterion, distributed among the evaluated approaches

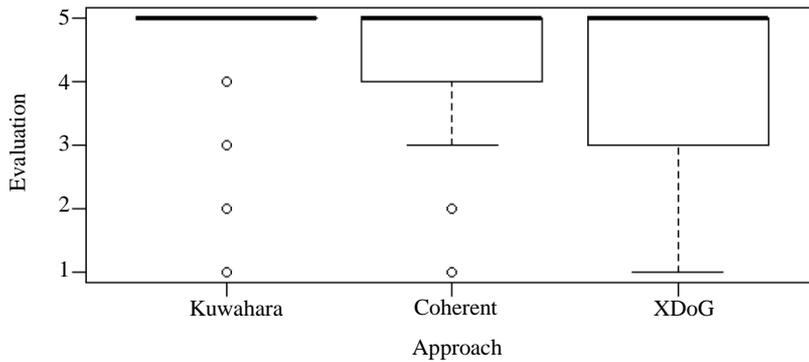


Fig. 9: Boxplots of the votes cast for the fifth criterion, distributed among the evaluated approaches

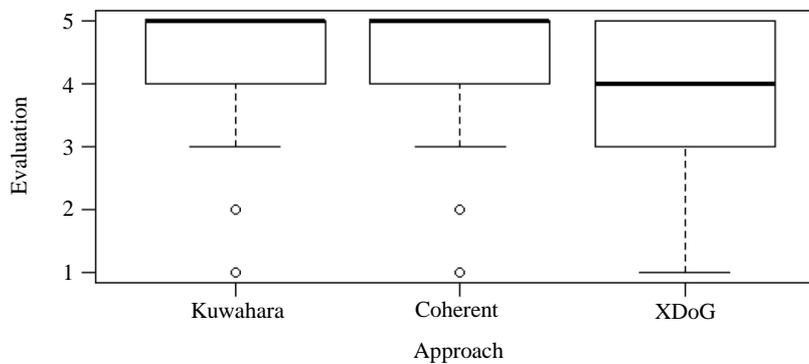


Fig. 10: Boxplots of the votes cast for the sixth criterion, distributed among the evaluated approaches

A pairwise comparison of approaches is then performed using the Wilcoxon test. The p-values resulting from the comparison for each criterion can be seen in Tables 3 to 8.

The p-values in bold indicate that, statistically, there are differences between approaches. For the third criterion, the compared approaches are statistically equivalent. For this reason, further analysis of the third criterion is not performed.

Next, the approaches that presented statistical differences were further analyzed, aiming to determine which one was better evaluated, under 99% confidence. Table 9 shows the p-values of the comparison between approaches regarding the first criterion.

Table 9 indicates that, according to the first criterion, XDoG performed best, followed by the Coherent and Kuwahara approaches, in that order. Table 10 shows the p-values of the comparison between approaches for the second criterion.

For second criterion, the Kuwahara and Coherent approaches performed best, followed by the XDoG approach. Table 11 presents the p-values of the comparison between approaches for the fourth criterion.

For the fourth criterion, the Kuwahara approach performed best, followed by the Coherent and XDoG approaches, in that order. Table 12 presents the p-values of the comparison between approaches for the fifth criterion.

Results on Table 12 indicate that, for criterion 5, the Kuwahara approach was better evaluated, followed by the Coherent and XDoG approaches. Table 13 presents the p-values of the comparison between approaches for the sixth criterion.

Results on Table 13 indicate that, for the sixth criterion, the Kuwahara approach is better evaluated, followed by the Coherent and XDoG approaches, in that order.

This case study aimed to verify the subjective satisfaction of the participants regarding the results of non-photorealistic rendering approaches originally intended for still images, but applied to videos in this research.

Table 3: Differences resulting from the Wilcoxon test for the first criterion

Approaches	Kuwahara	Coherence
Coherence	2×10^{-04}	-
XDoG	2×10^{-16}	5.1×10^{-11}

Table 4: Differences resulting from the Wilcoxon test for the second criterion

Approaches	Kuwahara	Coherence
Coherence	0.031	-
XDoG	2.6×10^{-09}	8.6×10^{-05}

Table 5: Differences resulting from the Wilcoxon test for the third criterion

Approaches	Kuwahara	Coherence
Coherence	0.4252	-
XDoG	0.0015	0.0106

Table 6: Differences resulting from the Wilcoxon test for the fourth criterion

Approaches	Kuwahara	Coherence
Coherence	0.00022	-
XDoG	4.1×10^{-13}	8.6×10^{-15}

Table 7: Differences resulting from the Wilcoxon test for the fifth criterion

Approaches	Kuwahara	Coherence
Coherence	0.004	-
XDoG	3.6×10^{-07}	0.015

Table 8: Differences resulting from the Wilcoxon test for the sixth criterion

Approaches	Kuwahara	Coherence
Coherence	3.2×10^{-06}	-
XDoG	2×10^{-16}	3.2×10^{-06}

Table 9: P-values resulting from the Wilcoxon test for the best approach for the first criterion

Approaches	Kuwahara	Coherence
Coherence	0.9999	-
XDoG	1	1

Table 10: P-values resulting from the Wilcoxon test for the best approach for the second criterion

Approaches	Kuwahara	Coherence
XDoG	4.329×10^{-10}	2.876×10^{-05}

Table 11: P-values resulting from the Wilcoxon test for the best approach for the fourth criterion

Approaches	Kuwahara	Coherence
Coherence	0.000109	-
XDoG	6.867×10^{-14}	2.855×10^{-05}

Table 12: P-values resulting from the Wilcoxon test for the best approach for the fifth criterion

Approaches	Kuwahara
Coherence	0.00133
XDoG	6.021×10^{-08}

Table 13: P-values resulting from the Wilcoxon test for the best approach for the sixth criterion

Approaches	Kuwahara	Coherence
Coherence	1.399×10^{-06}	-
XDoG	2.2×10^{-16}	1.622×10^{-06}

Table 14: Average processing time (in seconds) of the investigated approaches, for processing 10 video frames

Time vs approach	Kuwahara	Coherence	XDoG
Time (s)	10	8	15

Looking at these results, when considering a 99% confidence level, the Kuwahara approach was better evaluated for most criteria, seconded by the Coherent and the XDoG approaches. In conclusion, the statistical analysis indicate that there is greater preference for the rendering results of the Kuwahara approach.

In addition to identifying the best evaluated approach in this study, it was found out that viewer-friendly video renderings can be produced, despite the problems discussed in the Introduction section. Moreover, the processing times of the approaches discussed in the Chosen Approaches section are not prohibitive when they are applied to videos (Table 14).

The average processing times presented on Table 14 were obtained by processing 10 consecutive video frames. An Intel core i7 computer, with 8Gb of RAM and 1Tb hard drive, running the Windows 10 operating system, was employed in this evaluation. Processing time was not directly related to rendering quality. For instance, the Kuwahara filter was best evaluated but had an intermediate processing time, when compared to the other two approaches.

The videos generated in this study can be accessed at: <https://drive.google.com/drive/folders/1u4zqQsNhF2SOt3pQh2vTWa-kbdLF6yiY?usp=sharing>.

Conclusion

It is a fact that two-dimensional (i.e., image based) NPR has proliferated in the past two decades. In this study, aiming at verifying the effect of applying to videos NPR state-of-the-art approaches originally developed for static images, a comparative study was reported. An experimental evaluation methodology based on polling the opinion of human observers about videos rendered with three NPR approaches was given. The subject votes, according to Fig. 5 to 10, revealed the successful use of the NPR approaches considered in this study for most of the adopted criteria. It means that, from the point of view of the observers, those approaches can produce pleasant and satisfactory results when applied to videos.

There are two key academic and artistic contributions of this study. First, a methodological approach was given that can be applied to academic NPR related studies. Second, the digital art industry can make use of such approach to develop more pleasant advertising campaigns, presentations and other related products to their consumers.

Some possibilities for future work remain. In addition to the possibility of extending the methodology to other approaches, the possibility of increase the video database size is also envisaged in order to improve the research conclusions.

Regarding the adopted criteria, the statistical results obtained did not indicate differences between some pairs of the compared approaches. The recruitment of a larger number of participants in an extension of this study may allow verifying if the population size influences such results.

Finally, the use of samples involving factors such as age, gender and level of study of the participant seems promising in order to identify if such factors influence the experimental results.

Author's Contribution

Alberto Fagner Ferreira de Barros: Searching, finding and applying the approaches used in this study. Analysing the resulting data and reviewing the manuscript.

Jose Eustaquio Rangel de Queiroz and Herman Martins Gomes: Supervising the comparative experiment, reviewing the manuscript iteratively and giving recommendations and suggestions.

Ethics

This study is original and contains unpublished material. The corresponding author confirms that all of the others authors read and approved this manuscript. Additionally, the research was conducted with approval by and in accordance with the Ethics Committee of Federal University of Campina Grande (Paraíba State - Brazil).

References

- Akenine-Möller, T., E. Haines and N. Hoffman, 2018. Real-Time Rendering. 4th Edn., AK Peters/CRC Press, Boca Raton, ISBN-10: 1351816152, pp: 1198.
- Arruda, 2014. Uma Abordagem Híbrida para a Renderização Não Fotorrealística de Vídeos Digitais. PhD Thesis, Universidade Federal de Campina Grande.
- Baugh, G. and A. Kokaram, 2010. Semi-automatic motion based segmentation using long term motion trajectories. Proceedings of the International Conference on Image Processing, Sept. 26-29, IEEE Xplore Press, Hong Kong, China, pp: 3009-3012. DOI: 10.1109/ICIP.2010.5653946
- Benard, P., J. Lu, F. Cole, A. Finkelstein and J. Thollot, 2012. Active strokes: Coherent line stylization for animated 3d models. Proceedings of the Symposium on Non-Photorealistic Animation and Rendering, (PAR' 12), Eurographics, pp: 37-46. DOI: 10.2312/PE/NPAR/NPAR12/037-046

- Borawski, M., 2014. Non-photorealistic rendering with the use of short segments of straight lines in a vector space of increments. *Control and Cybernetics*.
- Cao, C., S. Chen, W. Zhang and X. Tang, 2011. Automatic motion-guided video stylization and personalization. *Proceedings of the 19th ACM International Conference on Multimedia, (ICM' 11)*, ACM, pp: 1041-1044. DOI: 10.1145/2072298.2071933
- Chen, J., G. Turk and B. MacIntyre, 2012. A non-photorealistic rendering framework with temporal coherence for augmented reality. *Proceedings of the International Symposium on Mixed and Augmented Reality*, Nov. 5-8, IEEE Xplore Press, Atlanta, GA, USA, pp: 151-160.
DOI: 10.1109/ISMAR.2012.6402552
- Gangopadhyay, A., S. Pachori and S. Raman, 2016. Automatic silhouette photography. *Proceedings of the 20nd National Conference on Communication*, Mar. 4-6, IEEE Xplore Press, Guwahati, India, pp: 1-6. DOI: 10.1109/NCC.2016.7561153
- Hogg, R.V., E.A. Tanis and D.L. Zimmerman, 1977. *Probability and statistical inference*. Macmillan New York.
- Ianeva, T.I., A.P. de Vries and H. Rohrig, 2003. Detecting cartoons: A case study in automatic video- genre classification. *Proceedings of the International Conference on Multimedia and Expo*, Jul. 6-9, IEEE Xplore Press, pp: I-449.
DOI: 10.1109/ICME.2003.1220951
- Jähne, B., H. Schar and S. Körkel, 1999. Principles of filter design. *Handbook of computer vision and applications*, 2: 125-151.
- Junior, J.D.O.A., J.E.R. de Queiroz and H.M. Gomes, 2013. An approach for non-photorealistic rendering that is appealing to human viewers. *Proceedings of the 26 Conference on Graphics, Patterns and Images*, Aug. 5-8, IEEE Xplore Press, Arequipa, Peru, pp: 242-249.
DOI: 10.1109/SIBGRAP.2013.41
- Kang, D., P. Kong, K. Yoon and S. Seo, 2015. Directional texture transfer for video. *Multi. Tools Applic.*, 74: 245-258. DOI: 10.1007/s11042-013-1759-3
- Kyprianidis, J.E. and H. Kang, 2011. Image and video abstraction by coherence-enhancing filtering. *Comput. Graph. Forum*, 30: 593-602.
DOI: 10.1.1.372.3200
- Kyprianidis, J.E., H. Kang and J. Döllner, 2009. Image and video abstraction by anisotropic kuwahara filtering. *Comput. Graph. Forum*, 28: 1955-1963.
DOI: 10.1111/j.1467- 8659.2009.01574.x
- Kyprianidis, J.E., J. Collomosse, T. Wang and T. Isenberg, 2012. State of the art: A taxonomy of artistic stylization techniques for images and video. *Trans. Visualiz. Comput. Graph.*, 19: 866-885.
DOI: 10.1109/TVCG.2012.160
- Medhi, I., C.R. Patabandhige and K. Toyama, 2007. Cartoon generation for text-free user interfaces. *Proceedings of the International Workshop on Computer Vision Applications for Developing Regions, (ADR' 07)*.
- Pham, T., 2006. Spatial adaptivity in super-resolution of undersampled image sequences. PhD Thesis, Delft University. Technology.
- Wang, Q., D. Chen, S. Li, Q. Wu and Q. Zhang, 2017. An adaptive cartoon-like stylization for color video in real time. *Multi. Tools Applic.*, 76: 16767-16782.
DOI: 10.1007/s11042-016-3951-8
- Weickert, J. and H. Schar, 2002. A scheme for coherence-enhancing diffusion filtering with optimized rotation invariance. *J. Visual Commun. Image Representation*, 13: 103-118.
DOI: 10.1006/jvci.2001.0495
- Wenhua, Q., X. Dan, Y. Kun and G. Zheng, 2015. Image abstraction painting of flow-like stylization. *Tehnički vjesnik*, 22: 837-844.
DOI: 10.17559/TV-20150612175259
- Winnemöller, H., 2011. XDoG: Advanced image stylization with extended difference-of-Gaussians. *Proceedings of the ACM SIGGRAPH/Eurographics Symposium on Non-Photorealistic Animation and Rendering, (PAR' 11)*, ACM, pp: 147-156.
DOI: 10.1145/2024676.2024700
- Zhang, W., C. Cao, S. Chen, J. Liu and X. Tang, 2013. Style transfer via image component analysis. *Tran. Multi.*, 15: 1594-1601.
DOI: 10.1109/TMM.2013.2265675