Comparing Image-Processing Operators by Means of the Visible Differences Predictor

Martin Čadík
Czech Technical University in Prague
Faculty of Electrical Engineering
Karlovo náměstí 13
121 35 Prague, Czech Republic
cadikm@fel.cvut.cz

Pavel Slavík
Czech Technical University in Prague
Faculty of Electrical Engineering
Karlovo náměstí 13
121 35 Prague, Czech Republic
slavik@fel.cvut.cz

ABSTRACT

Utilization of non-photorealistic techniques (NPR) is beneficial in many cases, in comparison with the use of traditional rendering methods in computer graphics. The observer's sensation is often straighter, clearer, or even more valuable. There exists plenty of various NPR techniques in computer graphics, however application of one technique to the specific problem is not necessarily as providential as usage of another one. There arises a strong need to classify NPR techniques with respect to their applicability, to find a mechanism able to compare NPR techniques automatically. The search for such a mechanism will be a long term goal. We present our first steps towards the solution to this problem – the comparison of 2D-based NPR techniques using Daly's Visible Differences Predictor (VDP). Results from our experiments are reported and will be used to improve communication in graphical user interfaces.

Keywords

Non-Photorealistic Rendering, Image-Processing, Visible Differences Predictor, Human Visual System, Human–Computer Interaction.

1. INTRODUCTION

In recent years a lot of effort has been spent to the research of incorporation of human perception into the computer graphics methods. Thanks to this we have seen a big progress in several areas, e.g. in the field of comparison of images by computer. It is well known that classical metrics like Root Mean Square (RMS) error are not sufficient when applied to the comparison of images, because the RMS error poorly predicts the differences as perceived by the human observer. To solve the problem properly the visual differences predictors have evolved [Dal93a, Lub95a]. These predictors have been formerly applied for various purposes [Mys98a, Zho02a] mainly in the field of the realistic image synthesis.

WSCG POSTERS proceedings
WSCG'2004, February 2-6, 2004, Plzen, Czech Republic.

In this paper we use the techniques mentioned above for comparison of images acquired by various image-processing operators, or let us say 2D NPR techniques [Goo01a]. The utilization of a NPR technique is in many practical cases more advantageous than the use of a classical photorealistic technique. Indeed, NPR methods allow us to emphasize or omit details in order to communicate information more effectively [Goo02a]. For example, sketch rendered images (one group of NPR techniques) of architectural scenes have shown better result in appreciation between architects and clients, compared with that obtained through realistic rendering [Sch96a].

Nevertheless not as much attention as to comparison of classical rendering methods was given to the comparison of output images of NPR techniques. The existence of a mechanism or a metric able to compare individual NPR outputs is urgently required in many areas where the NPR techniques are used. The motivation for such a research is a suitability of individual NPR techniques for rendering specific scenes and objects.

We assume the way towards the solution to this problem has two stages. The first stage is the "low-level" perception stage that we partly address in this paper by the means of the Daly's Visible Differences Predictor. Second stage is the "semantic-level" perception where the phenomena like the meaning of a scene, semantics or context are treated.

2. BACKGROUND

Since we use the Daly's Visible Differences Predictor extensively in this work, we first give a brief description of its principles.

2.1. Visible Differences Predictor

The Visible Differences Predictor (VDP, see [Dal93a] for more details) is an algorithm for describing the human visual response. The goal of the VDP is to determine the degree to which physical differences between images become visible.

The VDP is a relative metric because it does not describe an absolute value of image quality but instead addresses the problem of differences between two images. It consists of components for calibration of the input images, a human visual system (HVS) model and a method for displaying the HVS visible differences. The input to the algorithm includes two images and parameters for viewing conditions, whereas the output is a map describing the visible differences between them. The output map defines the probability of detecting the differences between these two images as a function of their location in these images.

Unlike most image-quality metrics the VDP does not reduce its output into a single number. An advantage is that we can see the nature of the difference and then use the information to further improvement of the design. However we want to quantify the difference by a single number as well, so we can follow the approach described by Myszkowski in [Mys98a]: the difference between images $D_{0.75}$ is the percentage of pixels for which the probability of difference detection is greater than 0.75. It is assumed, that the difference can be perceived for a given pixel when the probability value is greater than 0.75 (75%), which is the standard threshold value for discrimination tasks.

3. COMPARISON OF IMAGE-PROCESSING OPERATORS

Since the visualization by means of NPR techniques is commonly not used in cases where we want to obtain image as close to the reality as possible, our problem has two stages. As we already declared in the introduction, the first stage is the "low-level" perception stage. In this section we describe the insight in this stage via comparison by means of the VDP method.

The goal is to have a tool that will determine the most suitable technique for the given class of objects. Suitability in our case means the lowest information loss typical for particular NPR method. In such a way we can substantially improve visual communication with computer systems.

3.1. Input scenes

We compared the techniques mentioned below on several typical input images. These images included a natural photograph of a tree, a computer-generated bust, a classical radiosity scene (cornell box), a ray-traced scene, and several other images [Cad03a]. The radiosity scene contained soft shadows, while the ray-traced scene encompassed only sharp-edged shadows.

3.2. Tested techniques

There exists notable amount and variation of current techniques in non-photorealistic rendering (see the book by Gooch and Gooch [Goo01a] for an overview). With respect to this fact and to make our results reproducible we reduced our scope to the image based techniques, ordinary available with the program Adobe®Photoshop®6.0, although other possibilities are also open [Hal03a]. We investigated following 27 techniques divided into 6 groups:

brush Strokes: Angled Strokes, Crosshatch, Ink Outlines, Spatter,
sketch: Bas Relief, Graphic Pen, Chalk & Charcoal, Charcoal, Note Paper, Photocopy,
artistic: Colored Pencil, Cutout, Dry Brush, Paint Daubs, Poster Edges, Smudge Stick, Sponge, Underpainting, Watercolor,
stylize: Diffuse, Emboss, Find Edges,
other: Crystallize, Add Noise (12.5%), Pointillize, Sharpen, Smart Blur,
and certainly the unchanged **original** picture.

3.3. Comparison of the techniques

We compared the output image of every technique with the output image of every other technique, so we obtained 756 variations of difference maps for each input image (27 techniques plus the original image). For each difference map we computed the difference value $D_{0.75}$.

For all of the techniques we obtained 28 difference values. These difference values were treated as a discrete difference fuction $D_{0.75}(i)$, where the variable i stands for a technique ordered as in the section 3.2. These functions were plotted in planar and 3D graphs and examined for correlations.

We quantified the differences between two difference functions $D_{0.75}(i), D_{0.75}(j)$ by the absolute metric

$$\rho_{0.75}(i,j) = max|D_{0.75}(i) - D_{0.75}(j)|,$$

where i, j denote technique i or j respectively.

4. RESULTS

All of the results were depicted as ordinary planar charts to inspect the results. The investigated techniques were ploted on the horizontal axis, one may notice the gaps on the depicted functions that separate groups of the techniques (see Figure 1).

4.1. Absolute values of differences

The difference of the original image and the image produced by some image-based NPR technique is typically considerable. Therefore the difference is easy to detect for an observer and the absolute value of difference between images $D_{0.75}$ is high. This is perfectly true for *Ink Outlines* technique. The probabilities of the difference detection always exceed 82%. We have observed such a behaviour for any of the input images, because the *Ink Outlines* technique changes the input image massively and the difference is therefore straightforward.

Very high difference values were observed also for methods *Bas Relief, Graphic Pen, Note Paper* and *Pointilize*. Arithmetical averages of the difference values $D_{avg} = \frac{1}{28}\sum_{i=1}^{28} D_{0.75}(i)$ exceed the 90% threshold for all of these methods. The $D_{avg} = 90\%$ means for a particular technique that in average case a human oberver would be able to distinguish 90% of the image area when comparing to the image obtained by another technique.

For the synthetic images the absolute values of the differences were generally lower than for the photos. This is due to the fact that some of the technique's naturally added distortions are not exerted for the synthetic images.

4.2. Coherences

Consecutively we have investigated the coherences between the tested techniques to be able to classify them. Coherences between discrete functions $D_{0.75}(i)$ were examined using previously defined absolute metric ρ. This metric was evaluated for each pair of functions $D_{0.75}(i)$, $D_{0.75}(j)$ and was depicted as an ordinary planar graph. Below, we consider as coherent all of the cases when the absolute value of difference ρ does not exceed the 5% threshold.

4.2.1. General Coherences

We observed a strong coherence for following groups of techniques:

1) *Diffuse, Dry Brush, Original, Sharpen, Smart Blur,*
2) *Noise, Pointilize, Sponge,*
3) *Colored Pencil, Crystalize, Paint Daubs, Photocopy, Spatter.*

These coherences were independent on the type of the input image. Average values of ρ for several pairs from these groups are recorded in the Table 1.

	Technique–Technique	$\rho_{\mathbf{avg}}[\%]$
Group 1)	**Original–Smart Blur**	0.5799
	Original–Diffuse	2.639
	Dry Brush–Sharpen	4.949
Group 2)	**Noise–Sponge**	1.025
	Pointilize–Noise	3.897
	Pointilize–Sponge	4.346
Group 3)	**Color Pencil–Photocopy**	3.206
	Spatter–Crystalize	0.9
	Crystalize–Paint Daubs	4.772

Table 1: Average values of ρ for selected pairs of techniques.

Just mentioned groups reflect our vague knowledge of common properties of the given techniques. In the group 1 there are techniques that do not distort the image too much, they are just "improving" the input image in some sense. Group 2 consists of "point-based" techniques and finally in the group 3 there are techniques producing similar "shake" distortion.

4.2.2. Coherences between the groups of techniques

In the scope of the groups of techniques (brush strokes, sketches, artistic effects, stylize, other techniques) we primarily expected good coherences between the techniques. There is sometimes a noticable correlation of the results, however, generally we have noticed good coherences only in the group of sketches (after excluding the *Note Paper* method), where the average values of ρ do not exceed the 6% threshold. In the other groups there are no conspicuous coherences between apropriate techniques.

4.2.3. Coherences dependent on type of input image

Apart from the general-valid coherences described above, we found also the other group of coherent techniques dependent on the type of the input image. We have observed that for the synthesised images with the uniform-colored faces there is a good coherence between *Emboss, Colored Pencil, Photocopy, Bas Relief,* and *Charcoal* techniques. This is especially distinctive in the case of the *Cornell box* scene, where these techniques produce similar edge enhancement. The graphs in Figure 1 exhibit the coherence between the mentioned techniques. Note the minimal differences in the first and third parts of the charts, that represent very good coherence with the *Brush Strokes* and *Artistic* groups respectively.

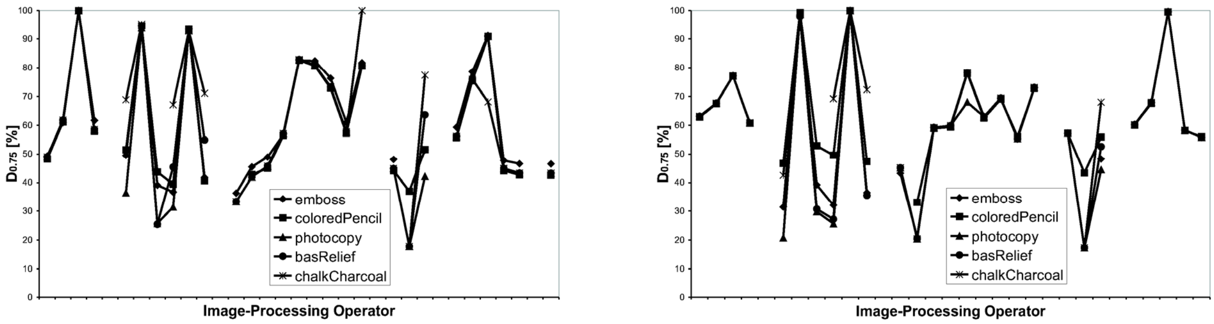

Figure 1: Coherences of techniques *Emboss, Colored Pencil, Photocopy, Bas Relied* and *Charcoal* for the *Cornell box* (left) and *Ray-traced* (right) input image.

5. CONCLUSIONS & FUTURE WORK

In this paper we have described our first steps towards finding a mechanism which would be able to automatically compare NPR images. We have shown that by such a low-level mechanism like the VDP is (from the point of view of the complexity of the human perception of the NPR images) we are able to distinguish some of the naturally vague defined groups of images with similar properties. Next, we have observed that the absolute values of differenecs are inherently high for several techniques and these values are generally lower for synthetic images than for the photographs.

However, it is evident that the VDP-like mechanisms are just a first stage in the field of comparing of the NPR techniques. NPR techniques are often utilized in such cases where we want to highlight fundamental semantic informations incorporated in the image, which principles of "low-level" perception are unable to catch.

In the future, we will carry on psychophysical experiments on human observers in order to validate the presented results. We will compare our results with other algorithms describing the human visual response, and especially we want to interpret our knowledge in the context of the work of Duke et al. [Duk03a] to design an algorithm, the result of which will correspond with "semantic sensation" of a human observing a NPR image.

Acknowledgements

This project has been partly supported by the Ministry of Education, Youth and Sports of the Czech Republic under research program No. Y04/98: 212300014 (Research in the area of information technologies and communications).

6. REFERENCES

[Cad03a] Čadík, M. Comparing Image–Processing Operators by Means of the Visible Differences Predictor. Project homepage, http://www.cgg.cvut.cz/~cadikm/papers/wscg04/.

[Dal93a] Daly, S. The visible differences predictor: an algorithm for the assessment of image fidelity, In Digital images and human vision, MIT Press, Cambridge, MA, 1993.

[Duk03a] Duke, D. J., Barnard, P. J., Halper, N., Mellin, M. Rendering and Affect. In Eurographics '03, UK, 2003.

[Goo01a] Gooch, B., Gooch, A. Non-Photorealistic Rendering. AK Peters Ltd, 2001.

[Goo02a] Gooch, A. A., Willemsen, P. Evaluating Space Perception in NPR Immersive Environments. Non-Photorealistic Animation and Rendering (NPAR'02), France, 2002.

[Hal03a] Halper, N., Isenberg, T., Ritter, F., Freudenberg, B., Meruvia, O., Schlechtweg, S., Strothotte, T. OpenNPAR: A System for Developing, Programming, and Designing Non-Photorealistic Animation and Rendering, In Proceedings of Pacific Graphics 2003, CA, 2003.

[Lub95a] Lubin, J. A visual discrimination model for imaging system design and evaluation. Vision Models for Target Detection and Recognition, 1995.

[Mys98a] Myszkowski, K. The Visible Differences Predictor: Applications to global illumination problems. In Rendering Techniques '98, 1998.

[Sch96a] Schumann, J., Strothotte, T., Raab, A., Laser, S. Assessing the Effect of Nonphotorealistic Rendered Images in CAD. In Proc. CHI '96, ACM Press, 1996.

[Zho02a] Zhou, H., Chen, M., Webster, M., F. Comparative Evaluation of Visualization and Experimental Results Using Image Comparison Metrics, Proceedings of IEEE Visualization'02, Massachusetts, 2002.