Perceptually Based Image Comparison Method

Sergey I. Titov
Moscow, Russia

Abstract

In this work a new perceptually based method of image comparison is proposed. It is based on the colour comparison in a perceptually uniform colour space CIE Luv, and using Contrast Sensitivity Function to modify colour comparison thresholds, provided by CIE Luv space.

This method can be used to measure image distortion in case of lossy image compression, and steering image generation.

***Keywords**: Perceptually based image comparison, perceptually uniform colour space, contrast sensitivity function.*

1. INTRODUCTION

In this work we introduce new image comparison method. It is based on the perceptually based colour comparison and modelling eye perception of the non-uniform images with contrast sensitivity function.

Image comparison is widely used in many areas. It is used in image search engines, in databases in systems Query by Example. This area requires very high performance (less than 10^{-3} sec per image), but comparison can be pretty rough, and should insensitive to image shears, tilts, and rotations.

Another area is comparison in rendering systems and image quality control in compression systems. These tasks do not need so hard time restrictions, but they require much more precise comparison, that can be sensitive to shears and rotations of the image. These tasks also need detection of areas, where images look different.

The goal of my work was to create image comparison method, that could be used to control image compression quality, and to increase performance of rendering algorithms. Special requirements were following:

- this method should give correct results for any display;
- it should detect areas, where images are noticeably different;
- it should provide general characteristic of image dissimilarity.

1.1 Background

Several works were made last years in this area. In one of them, by Gaddipatti et al [1] it was proposed to select perceptually important elements of the image and to pay most attention to comparing of these elements. It was also proposed in this work to use Contrast Sensitivity Function (CSF) to compute saliency values, that could be compared using MSE metric. Also it was shown in this work, why MSE metric can't be applied for direct image comparison.

Another work, by Neumann et al. [2] proposed to compare mean colours in random rectangles. They used CIE XYZ space for mean colour computation and CIE Luv space for colour comparison. The size of rectangles was a random value, distributed according to CSF, so that common size corresponded to the maximum sensitivity of the eye. Another idea described in this paper dealt with image distortion measurement. It was proposed, that only areas, where the difference is noticeable should influence total image difference

One more article, by Rammasubramiani et al [3], describes a rendering system, that uses image comparison to increase productivity. They propose to compute maximum luminosity deviation, that still produces unnoticeable image distortion. The only drawback of this system is that it takes into account just luminosity fluctuations, and does not use colour information.

1.2 The concept of the method

The idea of the proposed method is based on the model of human visual system (HVS), that uses contrast sensitivity function for correct colour comparison, and FFT for spatial frequency computation. The model is described in part 2. Using this model, the images are compared and the Visible Error Map is generated. The element of this map shows whether colours of corresponding pixels look different. The general characteristic of image dissimilarity is obtained during processing of this map, that is described in part 3.

2. HUMAN VISUAL SYSTEM MODELLING

The goal of modelling human visual system (HVS) was to provide correct and accurate colour comparison, that should be independent from the type of display, and image uniformity.

The proposed HVS model consists of two parts. The first part provides device-independent correct colour reproduction, and the second part compensates the impact of the image non-uniformity, modifying colour comparison thresholds

2.1 Colour Reproduction

The first problem we encountered, was a problem of correct colour reproduction and colour comparison. It appeared because phosphors in different monitors have different emission spectra. Therefore one colour (defined in RGB space) can look different on different monitors. To solve this problem we used perceptually uniform colour space CIE Luv. It is derivative space from the standard colour space CIE XYZ, and therefore provides device-independent correct colour reproduction.

Since it is perceptually uniform space, the distance between colours may be obtained by formula:

$$\Delta E_{Luv} = \sqrt{\Delta L^2 + \Delta u^2 + \Delta v^2}\ , \qquad (1)$$

where Δ – difference between corresponding components.

There are also 2 thresholds defined in this space. They help to determine whether one can notice difference between two colours. If the distance is lower than 1, than colours look like each other. If the distance is greater than 3, than difference can be easily noticed. If the distance is between 1 and 3, than difference between colours is very small, and can be noticed only at good viewing conditions.

To learn more about colour space conversion, see book by David Travis [4] or Poynton's Colour FAQ [5].

2.2 Threshold computation

The problem with colour comparison thresholds appears because they were determined for "ideal" cases, where the spot of one colour was painted on the background colour. In real cases there are many factors, that impair ability to distinguish colours. One of the main factors is the dependence of the eye sensitivity from the uniformity of the image.

This effect can be modeled with a Contrast Sensitivity Function.

2.2.1 Contrast Sensitivity Function

Contrast sensitivity function shows the ratio between percepted contrast of the image, and real contrast of the image, as a function of the spatial frequency of the image.

If L is an amplitude of a periodic signal, than it's contrast equals to: $C = \frac{L_{\max} - L_{\min}}{L_{\max} + L_{\min}} = \frac{\Delta L_{peak}}{L_{mean}} \approx \frac{\Delta L_{peak}}{L_{background}}$

Spatial frequency of a periodic signal equals to the number of cycles, that one can see under the angle of 1 visual degree. Therefore it can be used as a measure of image non-uniformity.

We decided to use contrast sensitivity function, proposed by Mannos and Sacrison [6], since it is one of the most popular CSFs, used in contemporary works.

Their CSF can be computed by following formula:

$$CSF(f) = 2.6 \cdot (0.0192 + 0.144 \cdot f) \cdot \exp\left(-(0.144 \cdot f)^{1.1}\right)$$

It has maximum at point 7.9 and reaches value about 0.98. Therefore it requires normalization, when using it as a weight coefficient (see formula 4).

2.2.2 Spatial Frequency Computation

Since definitions of the contrast and spatial frequency were given for periodic signals, we need to represent image as a sum of periodic signals to use CSF. The simplest periodic signals are sine and cosine. Therefore it was decided to use discrete Fourier transform. To increase effectiveness the algorithm of Fast Fourier Transform was used. As FFT is not local transform, to get local features of the image, it was subdivided into overlapping squares, and the FFT was applied to each square. The size of the square was selected so, that CSF could reach values between 1 and 0.5.

CSF was developed to measure contrast perception. If we want to use CSF to adopt colour comparison thresholds, then spatial frequency should reflect distortion, produced by fluctuations of u and v components of Luv space. It is proposed that FFT should be applied to the array of values:

$$E_{Luv} = \sqrt{L^2 + u^2 + v^2} .$$

To compute the spatial frequency, we need to estimate the influence of each frequency to the percepted non-uniformity of the image. The influence depends on the amplitude, and the eye sensitivity to this frequency. Thus we can write a formula for the spatial frequency:

$$F_{sp} = \frac{1}{N} \sum_{i,j} \frac{f_{r_{i,j}}}{l} \cdot cpd \cdot A_{ij} \cdot W_{ij} \text{, where} \qquad \textbf{(2)}$$

- $f_{r_{ij}}$ [cycles/square] – radial frequency corresponding to X_{ij} FFT coefficient;
- A_{ij} – real part of the X_{ij} FFT coefficient;
- W_{ij} – weight coefficient, that shows the influence of each frequency to the perceptable image non-uniformity;
- l [pixels] – square side, to which FFT is applied;
- cpd [pixel/degree] – cycle/pixel to cycle/degree conversion coefficient;
- N – this coefficient is computed from the equation: $\frac{1}{N} \sum_{i,j} A_{ij} \cdot W_{ij} = 1$

The radial frequency, corresponding to FFT coefficient is computed by the formula:

$$f_{r_{i,j}} = \sqrt{f_{h_i}^2 + f_{v_j}^2} \text{, where}$$

f_{h_i} and f_{v_j} – are horizontal and vertical frequencies corresponding to X_{ij} FFT coefficient.

The cpd value is computed according to formula:

$$cpd = \frac{360}{\pi} \cdot \arctan\left(\frac{\sqrt{w_p^2 + h_p^2}}{2D}\right) \text{, where} \qquad \textbf{(3)}$$

- D [cm] is a distance between eye and display;
- h_p, and w_p [cm] are height and width of the pixel respectively;

The weight coefficient W_{ij} shows human eye sensitivity to fluctuations, produced by FFT frequency. Therefore it is equal to:

$$W_{ij} = NormCSF\left(\frac{f_{r_{i,j}}}{l} \cdot cpd\right) \cdot , \qquad \textbf{(4)}$$

where $\mathrm{NormCSF}(x) = \frac{\mathrm{CSF}(x)}{\max\limits_{y \in [0;\infty)} \mathrm{CSF}(y)}$ – is a normalised Contrast Sensitivity Function.

2.2.3 Threshold Computation

After the computation of the spatial frequency it becomes possible to compute colour comparison thresholds. We applied CSF for this job.

Since CSF shows the ratio between percepted image contrast and real image contrast. Thus, to determine whether one can distinguish 2 colours, it is required to multiply the distance between colours by the CSF value, and to compare this value with thresholds 1 and 3. So it is more convenient to use as a first threshold at point (i,j) the value T_{ij}, and as a second threshold – 3 T_{ij}

$$T_{ij} = \frac{1}{CSF\left(F_{sp\,ij}\right)}, \text{ where} \quad (5)$$

$F_{sp\,ij}$ – spatial frequency at point (i,j);

All values T_{ij} for one image give us individual threshold map. To get joint threshold map of 2 images we should take minimal of 2 thresholds from individual maps, because it corresponds the higher eye sensitivity.

At this stage we can compute visible error map, that shows whether images look different at particular point. The element of the Visible Error map equals to the distance between colours, if it exceeds value T_{ij}, and is zero otherwise.

3. VISIBLE ERROR MAP PROCESSING

Following the ideas, described in [3], we propose to use as a general characteristic of image dissimilarity mean value of visible error map.

$$MVE = \frac{1}{H \cdot W} \sum_{1 \le i \le W} \sum_{1 \le j \le H} \Delta E_{Luv\,i,j} \cdot I_{i,j} \text{ , where}$$

H, W – height and width of the image respectively;

$\Delta E_{Luv\,i,j}$ – distance between colours of pixels (i,j) (see **formula 1**)

$$I_{i,j} = \begin{cases} 1 \ : \ \Delta E_{Luv\,i,j} > T_{i,j} \\ 0 \ : \ \Delta E_{Luv\,i,j} \le T_{i,j} \end{cases}$$

T_{ij} – element of a joint threshold map

This allows to eliminate the influence of the areas, where images are different, but the difference is unnoticeable. It also allows to suppress the influence of the isolated pixels with noticeably different colours, when their perception is limited by the resolution power of the eye.

The experiments showed that the difference between images becomes noticeable, when MVE becomes greater than 0.8-0.9 for grayscale images, and 1.2-1.4 for colour ones.

Another way to present results of comparison is to visualize Visible Error Map.

According to the properties of Luv colour space, the image was separated into three areas. The first area, where visible error was lower than T_{ij}, was painted with black. The second, where visible error was between thresholds, was painted with green (gray in print version). And the third, where the error was greater than $3T_{ij}$, was painted with red (white in print version)

4. CONCLUSION AND FUTURE WORK

Proposed method was implemented in an image comparison program, and showed both, precise image comparison, and good performance. High enough to use it in image compression systems, and rendering systems.

This program provided user friendly interface for setting up viewing conditions, and visual image comparison. It was used in the project, devoted to the development of a new lossy image compression method.

The examples of image comparison are presented in appendix 1.

It seems to be important in future, to take into account that resolution power of the eye is limited. I hope this should solve a problem of isolated pixels with different colours. Another research directions are experiments with different Contrast Sensitivity Functions, described in [7].

5. ACKNOWLEDGEMENTS

This work was supported in part by the Intel Corporation.

I am grateful to Prof. Yuri Bayakovsky, who told me about this area, and provided me with required papers, and Leonid Levkovich, Anton Pereberin, and other members of MSU Media Lab for many discussions and constant support of my work.

Especially I want to thank Konstantin Kolchin for discussions, and many advises, he gave me at all stages of my work.

6. REFERENCES

[1]*Mahesh Ramasubramanian, Sumanta Pattanaik, Donald Greenberg* "A Perceptually Based Physical Error Metric for Realistic Image Synthesis" // SIGGRAPH-99 Proceedings.

[2]*Ajeetkumar Gaddipatti, Raghu Machiraju, Roni Yagel* "Steering Image Generation with Wavelet Based Perceptual Metric" // EuroGraphics-97, Volume 16, Number 3, pp.C-241 – C-251.

[3]*Laszlo Neumann, Kresimir Matkovic, Werner Purgathofer* "Perception Based Colour Image Difference" // EuroGraphics'98 Vol. 17(98) #3 pp.233-241.

[4]*David Travis* "Effective Colour Displays. Theory and practice" // pp.89-100; Academic Press 1991 ISBN 0-12-697690-2.

[5]*Charles Poynton FAQ on Colour and Gamma // http://198.53.144.31/~poynton/ColorFAQ.html http://198.53.144.31/~poynton/GammaFAQ.html*

[6]*J.L. Mannos, D.J. Sakrison* "The Effects of Visual Fidelity Criterion on the Encoding of Images" // IEEE Transactions on Information Theory IT-20(4) 1974, pp. 525-536.

[7]*G.C. Higgins* "Image Quality Criteria" // Journal of Applied Photographic Engineering. 1977 3(2), pp. 53-60

About the author

Sergey Titov is a postgraduate student of Computational Mathematics and Cybernetics Dept., MSU

E-mail: tit@motor.ru

Аннотация

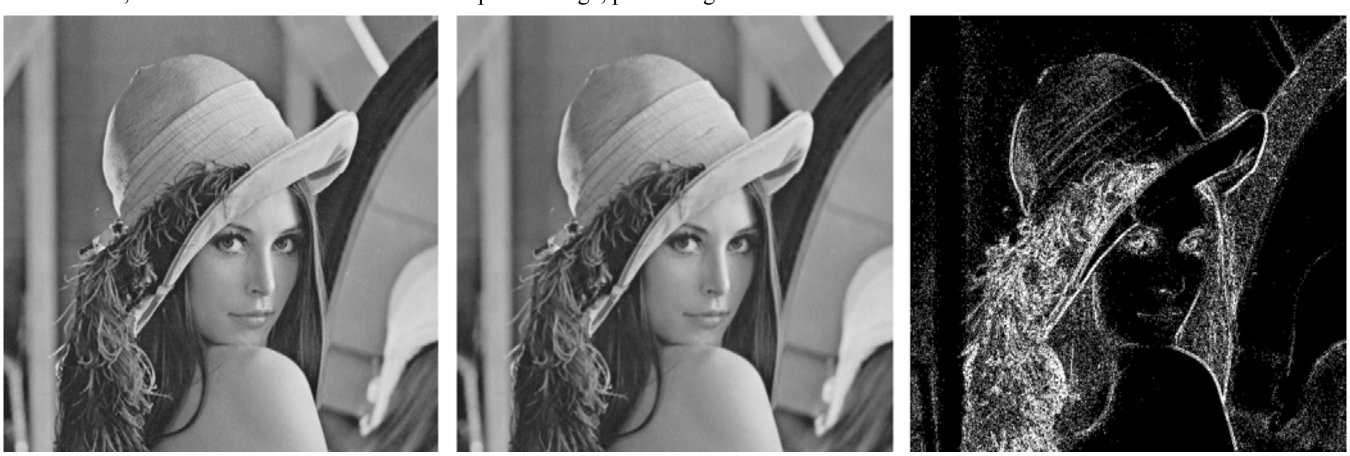
В этой работе предлагается новый метод сравнения изображений, учитывающий особенности человеческого восприятия. Для этого используется адаптация порогов различимости цветов в пространстве CIE Luv с помощью функции чувствительности контраста.

Предложенный метод может быть использован для управления синтезом изображений по геометрической модели, а также для контроля качества изображений, при сжатии изображений с потерями.

7. APPENDIX 1. EXAMPLES

Original (left), 12 times compressed “Portrait” image, and visualization of the Visible Error Map (right).

MVE – 2.76, PSNR – 29.551 dB   102x160 pixels image, processing time – 3.7 sec



Original (left), blurred “Lena” image, and visualization of the Visible Error Map (right).

MVE – 0.54, PSNR – 35.539 dB   512x512 pixels image, processing time – 13.7 sec

All experiments were conducted at Pentium-III-500 128M Ram workstation, with 17 inch display, under Win NT 4.0.
Viewing distance – 60 cm.

Black color at right pictures corresponds to pixels, where colour difference is less T_{ij},, gray to pixels, where difference is between thresholds, and white to pixels, where difference is greater than $3T_{ij}$.