The semantics of image quality

T.J.W.M. Janssen, F.J.J. Blommaert
IPO, Center for Research on User–System Interaction
Eindhoven, The Netherlands

Abstract

Using an "information-processing" approach we give a semantic description of image quality. Experimental evidence for this description, which allows one to meaningfully characterize the quality of an image as the degree to which the image can be successfully exploited by the observer, will be discussed.

1. Introduction

At present, research concerning objective image quality metrics is mostly based upon an approach which can be characterized as a "signal evaluation" approach. In this approach, the image is regarded as a complex signal which deviates to a certain extent from the complex signal that represents the "ideal" or "original" image. Images are defined in the physical or perceptual domain—in the latter case using models of the earliest stages of visual perception—and quality measures are defined as distances in an appropriate function space, e.g., Euclidian distance between actual and original image.

In contrast, we regard the processing of images by the visuo-cognitive system not as the evaluation of complex signals but instead as the processing of visual information. This "information-processing" approach enables us to characterize the quality of an image in a more meaningful manner as the degree to which the image can be successfully exploited by the observer.

2. Subject of research

The subject of research is the quality of images of natural scenes as perceived by observers. Specifically, we want to understand the fundamental processes that underly image quality. Our aim is to use this knowledge in order to arrive at a computational theory for image quality. Algorithms based on this theory will enable us to evaluate and optimize the quality of reproduced images.

3. Approach

We adopt the general viewpoint in computational cognition[1] that, first, the visuo-cognitive system can be considered to be an information-processing system and, second, information-processing systems can only be completely understood when they are understood at three distinct levels. These levels are the *semantic level*, which is the level describing the system in terms of computational goals and strategies, the *algorithmic level*, which is the level describing the implementation of the computational theory into algorithms and associated representations, and the *level of physical implementation*, which is the level describing the physical implementation of these algorithms and representations.

In this research we use the above mentioned structure in a strict top–down approach in which we first try to formulate a theory of image quality at the semantic level. In a second stage, we will implement this semantic theory into algorithms for the evaluation and optimization of image quality. These implementations will allow us to thoroughly test the semantic theory in experiments.

4. Theory

An important ingredient of the theory of image quality outlined here is the notion that visual processing serves the higher goal of a proper interaction between an observer and his environment. In the interaction between observer and environment three main processes can be distinguished (figure 1): *perception*, the acquirement of information from the outside world and the construction of an internal representation from it, *cognition*, the interpretation of this internal representation, and *action*, an appropriate response according to this interpretation. It quite directly follows from this description that in order to ensure a proper response to outside world occurrences, the observer's interpretation of what is "there" in the outside world should under no circumstance conflict with what really is there. Within this context, we characterize the quality of an image as the degree to which this image can be successfully exploited by an observer.

Looking more closely at the visual processing of images (figure 2) one quite directly arrives at the conclusion that an image of "good" quality should allow for a proper interpretation of the scene represented in it, and should therefore satisfy two main criteria: first, the visual representation of the image should be precise and, second, the visual representation of the image should correspond to

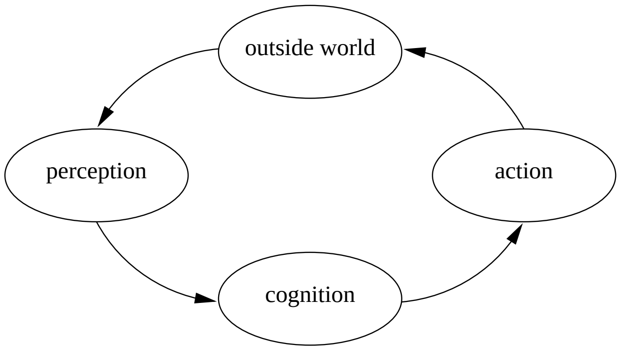

Figure 1: *The interaction between an observer and his environment consists of acquiring information from the outside world and constructing an internal representation from it (perception), interpreting the obtained internal representation (cognition), and responding appropriately according to this interpretation (action). A proper interaction between observer and environment requires that the observer's interpretation of what is "there" in the outside world should under no circumstance conflict with what really is there.*

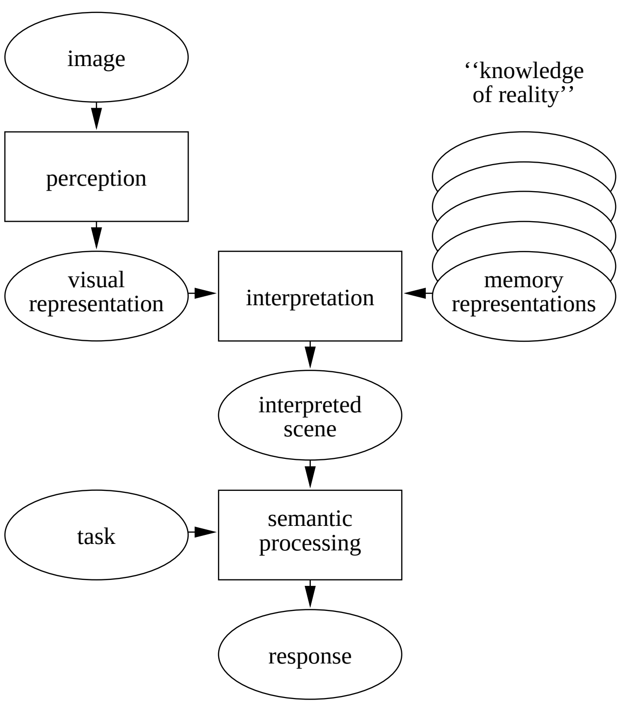

Figure 2: *Visuo-cognitive processing of images. Ellipses denote representation of information, and rectangles denote processes transforming one representation into another. Ensuring a proper interpretation of scene content imposes two requirements upon the image: its visual representation should be precise, and its visual representation should correspond to "knowledge of reality" as stored in memory.*

knowledge of reality as stored in memory. These requirements can be related to the attributes "usefulness" (i.e., the internal signal–to–noise ratio) and "naturalness" (i.e., the degree of correspondence to knowledge of reality) of the visual representation of the image.

In general, the sets of requirements maximizing either usefulness or naturalness will not coincide. Detection or discrimination of objects or details in an image, for example, may require enhancement of certain features in this image, which would then be reproduced in a somewhat unnatural way. The quality of an image will hence be determined by the degree to which both sets of requirements are satisfied *simultaneously*. Based on this observation, we describe image quality in terms of a compromise between partially conflicting demands maximizing either usefulness or naturalness.

5. Applications in the colour domain

Results of preliminary research involving global colour manipulations of natural scenes[2] have shown that the perceived quality of an image correlates well with the naturalness of that image. Nevertheless, small but systematic deviations between optimal parameter settings for naturalness and quality were found. Specifically, subjects' judgments of image quality peaked at slightly higher values of average chroma than judgments of image naturalness.

Recently, a series of experiments was performed to test the above definition of image quality in terms of a compromise between partially conflicting demands maximizing either usefulness or naturalness[3]. To this end, two kinds of global colour transformations were selected to increase or decrease the colourfulness of images of natural scenes. These manipulations, i.e., scaling chroma (C^*_{uv}) by a factor ranging from 0.5 to 2.0 and varying the colour temperature (T_c) of the reference white from 4650K to 10300K in seven steps of perceptually equal size, were expected to affect the naturalness and the usefulness of the images in distinctly different ways. Specifically, scaling chroma was expected to influence both naturalness and usefulness, whereas changing the colour temperature of the reference white was expected to influence only the naturalness of the images.

In separate sessions subjects were asked to judge the quality*, colourfulness† or naturalness‡ of the manipulated images on an eleven-point numerical scale (i.e., 0–10). The results, after z-correction and averaging over three repetitions, seven subjects and four scenes, are shown in figures 3 and 4. These results show that subjects' judgments of image quality are shifted with respect to subjects' judgments of image naturalness towards higher values of judged col-

* "The degree to which you like the colours in the image."
† "The presence and vividness of the colours in the image."
‡ "The degree to which the colours in the image seem realistic to you."

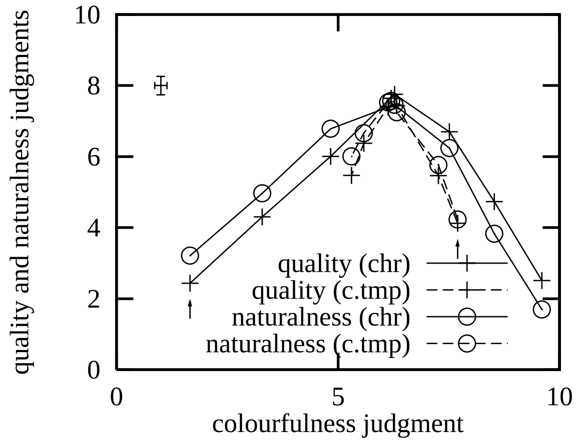

Figure 3: *Quality judgments (plusses) and naturalness judgments (circles) versus colourfulness judgments for the conditions chroma (solid lines) and colour temperature of the reference white (dashed lines). The arrows denote the images with smallest chroma multiplier and lowest colour temperature of the reference white, and the error cross denotes twice the standard error of the mean.*

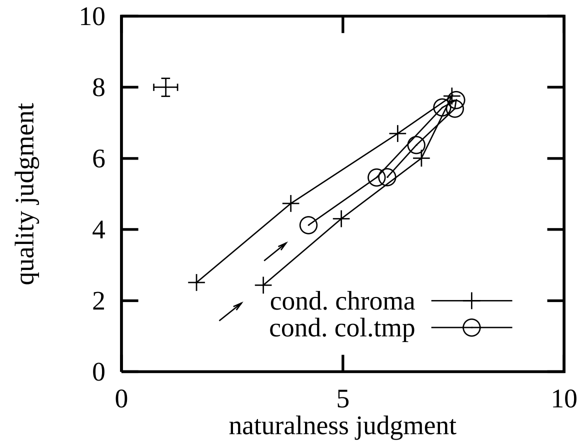

Figure 4: *Quality judgments versus naturalness judgments for the conditions chroma (plusses) and colour temperature of the reference white (circles). For the first condition a U-formed shape results from the shift between quality judgments and naturalness judgments. The arrows denote the images with smallest chroma multiplier and lowest colour temperature of the reference white, and the error cross denotes twice the standard error of the mean.*

ourfulness *only* for the condition chroma. A similar shift does *not* occur for the condition colour temperature of the reference white.

The shift between quality judgments and naturalness judgments for the condition chroma can be explained by viewing the image as a cloud of dots in the CIELUV colour space. In CIELUV, changing chroma can be thought of as a radial expansion or contraction of this cloud of dots away from or towards the reference white, while changing the colour temperature of the reference white can be thought of as a displacement of the entire cloud along the yellow–blue direction. Hence, changing chroma will influence distances between the dots in the cloud, whereas changing the colour temperature of the reference white does not. Since changing distances between points in a perceptually uniform space is—at a presumed, constant level of internal noise—equivalent to changing a perceptual “signal–to–noise” ratio, only the first manipulation will influence the usefulness of the image. The resulting variation in usefulness manifests itself as a shift between quality judgments and naturalness judgments, which would otherwise have coincided.

In an analogous series of experiments, results resembling those for the condition chroma were found when the brightness contrast of monochromatic images of natural scenes was varied by means of applying a sigmoid-shaped transformation to lightness (L^*; figure 5). Here, too, the applied transformation is influencing both the naturalness and the usefulness of the resulting images.

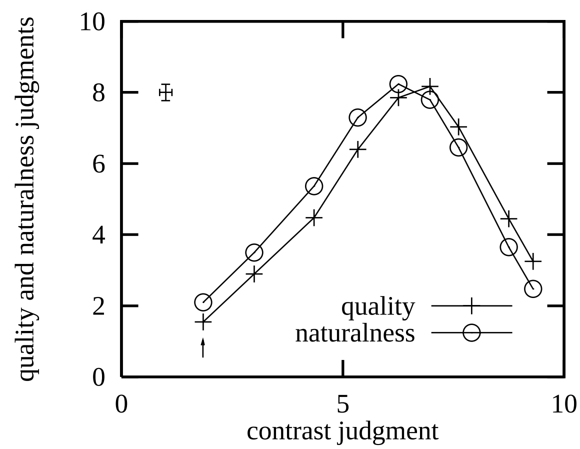

Figure 5: *Quality judgments (plusses) and naturalness judgments (circles) versus contrast judgments. The arrow denotes the image with smallest steepness of the inflexion point of the sigmoid-shaped transformation, which varied from 0.5 to 2.0. The error cross denotes twice the standard error of the mean.*

6. Naturalness of colour reproduction

A successful definition of the naturalness of object colours has been given in terms of "closeness to memory prototypes"[4]. Using this definition, the naturalness of the reproduction of an object colour can be described by a monotonically decreasing function (e.g., a Gaussian) of the distance in CIELUV colour space between the perceived object colour and the prototypical object colour stored in memory.

Within the theoretical concept already presented, the prototypical object colour can be viewed as a probability-density function of how the object's colour appeared to the observer in the past. Hence, the naturalness of an object colour as reproduced in an image can be described as the probability of the object colour appearing as it does in the image. Since the appearances of the object colour in the past and of the object colour in the image are subject to random fluctuations in, e.g., object reflectance properties, light source properties, and noise in the visual pathway, a Gaussian relation between naturalness of object colour and location in CIELUV colour space is expected for most object colours. Nevertheless, deviations from this general shape can be expected for objects showing a range of colours (e.g., rock) and objects showing multiple "discrete" colours (e.g., peppers).

7. Colourfulness

A question which we try to answer, and which is related to the above mentioned research, is whether the attribute colourfulness of an image is a scene-independent attribute, i.e., a statistic of the image which is judged independently from (cognitive) scene content, or a relative attribute, i.e., a statistic of the image which is judged relative to the statistic of the memory prototype of the displayed scene. To this end, two series of pairwise-comparison experiments have been conducted in which eight subjects were asked to judge the colourfulness difference, on a five-point numerical scale (i.e., -2–2), between twelve original, unprocessed scenes (first experiment) and between spatially scrambled versions of those scenes (second experiment) in which all cognitive scene content is eliminated but colour statistics are preserved[5]. A third experiment aimed at identifying the memory prototype of the scenes represented in the first experiment by means of scaling chroma and asking subjects to judge the naturalness of the manipulated images on an eleven-point numerical scale. From subjects' naturalness judgments, chroma multipliers maximizing naturalness were obtained by means of interpolating the data and determining the chroma multipliers for which the interpolation was maximal.

Difscal analysis[6] of subjects' judgments in the first two experiments showed that, first, subjects' judgments of colourfulness for the original, unscrambled images correlate highly ($r = 0.93$) with subjects' judgments of colourfulness for the scrambled images (see figure 6) and, second, subjects' judgments of colourfulness judgments correlate well ($r = 0.89$) with a root–mean–square measure of saturation (s_{uv}; figure 7). These findings seem to indicate that colourfulness is judged primarily independent of cognitive scene content.

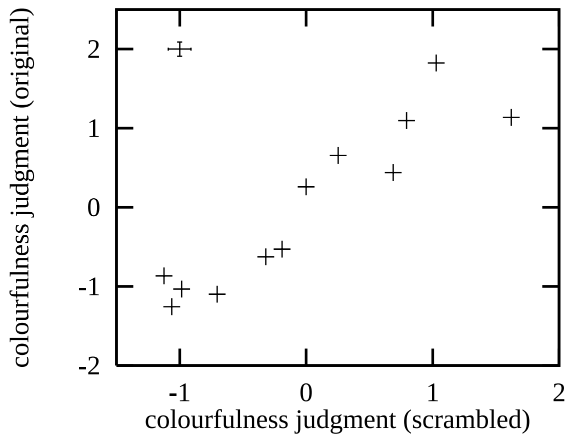

Figure 6: *Colourfulness judgments of the original scenes versus colourfulness judgments of the scrambled versions. Correlation* $r = 0.93$. *The error cross denotes twice the standard error of the mean.*

The above conclusion is confirmed by the results from the third experiment, where a negligibly small correlation ($r = 0.26$) between chroma multipliers maximizing naturalness and subjects' judgments of colourfulness for the original, unscrambled images was found. According to the hypothesis of colourfulness being a "relative" attribute, an image should be judged as colourful when chroma values of its prototype are small compared with chroma values of the image which is being judged. Hence, the chroma multiplier which maximizes the naturalness of that image is expected to be relatively small. In general, therefore, a significant negative correlation between chroma multipliers maximizing naturalness and subjects' judgments of colourfulness would then have been expected.

8. Concluding remarks

We have argued that image quality is, to the observer, a *useful* attribute of an image; an attribute which expresses how well the observer is able to employ the image as a source of information about the outside world. This view of image quality is strikingly different from the "perceived distance to the original"–philosophy often employed in image

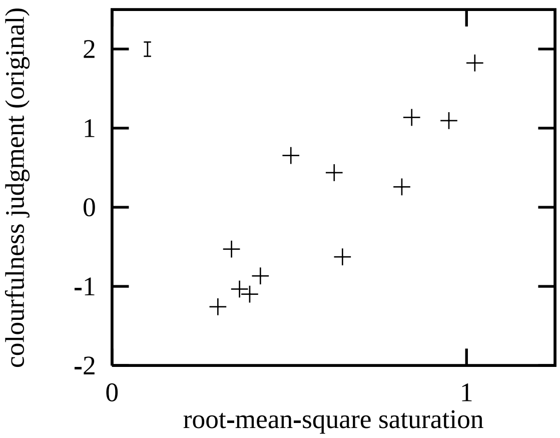

Figure 7: *Colourfulness judgments of the original scenes versus the root–mean–square of saturation (s_{uv}). Correlation $r = 0.89$. The error bar denotes twice the standard error of the mean.*

quality research.

The results of the experiments we discussed, support the concept that the quality of an image can be described in terms of a compromise between the naturalness and the usefulness of that image. The next step to proceed from this point onwards is to more thoroughly specify the naturalness and usefulness requirements imposed upon an image, e.g., by means of formulating algorithms. Implementations of such algorithms will enable the development of instrumental measures for the prediction of image quality and the estimation of parameter settings which optimize the quality of images.

9. References

1. A. Newell and H.A. Simon, *Human problem solving*, Prentice-Hall, Englewood Cliffs, (1972).
2. E.A. Fedorovskaya, H. de Ridder and F.J.J. Blommaert, Chroma variations and perceived quality of color images of natural scenes, *Color: Research and Application*, **22**, (1997).
3. T.J.W.M. Janssen and F.J.J. Blommaert, Image quality semantics, *The Journal of Imaging Science and Technology*, (to appear).
4. S.N. Yendrikhovskij, F.J.J. Blommaert, H. de Ridder and S.T.J. van Helvoirt, Naturalness judgments of object colours, *Proceedings of the AIC Color Conference*, Kyoto, Japan, (1997).
5. S.N. Yendrikhovskij, H. de Ridder, E.A. Fedorovskaya and F.J.J. Blommaert, Colourfulness judgments of natural scenes, *Acta Psychologica*, (to appear).
6. M.C. Boschman and J.A.J. Roufs, Text quality metrics for visual display units: II. An experimental survey, *Displays*, **18**, (1997).