

Psychophysical Metrics for Perceptible Image Quality

Vitali V. Gavrik, Imaging Consultant, Cologne

Annotation

Psychophysical grounds are discussed for economical quality metrics. The perceptible image quality recently proved to be related subconsciously to only a few important scene objects. Comparable distortions in other objects are perceptible under artificial conditions (e.g., colorimetry) or by intentional drawing observers' attention to the naturalness or other image features. In studies on sharpness and noise, different image details were taken into account while judging either the image quality, or an image feature mentioned in an instruction, or the image quality regarding the image feature. A naturalness instruction seems to involve all the details whose photometric statistics has been impressed in memory. The perceptible image quality recently proved to be a product of the independent single-stimulus functions whose mathematics is characteristic to other psychophysical response-to-stimulus dependencies. They consist of a high-quality tolerance plateau and one or two portions of linear quality decrease with the log-stimulus. Every stimulus is a simple optical property of an image detail. Most of the details are not critical due to their broad non-influence plateaus. Their photometry attracts no attention although has imprinted in visual memory, and influence the expert evaluations only after an inappropriate instruction. Transfer properties of an imaging system should provide the stimuli to meet their psychophysical tolerances. The single-stimulus functions can be obtained directly in a single-stimulus test, or with a simple iteration procedure applied to a representative set of images with multi-stimuli variations.

Economical Quality Metrics. Few Dominant Image Details Make Imperceptible All the Rest

It appears to be a common knowledge that economical image quality metrics should be an integral approach taking into account the photometry of an image as a whole. All the image details are assumed then to be rendered within about even tolerances. Most investigations have been aimed to elaborate such integral optical correlates to the perceptible image quality.

The attempts have resulted in complex dependencies or rough correlation. The integral optical properties of image were often provisory defined and their relations to directly perceptible image properties remained not quite clear. Their quantitative characteristics often required laborious procedures to be calculated from directly measured data.

The situation was especially intricate when the characteristics adopted did not concern a concrete image but were the conventional transfer characteristics of a technical imaging system. The latter are often expected to be an image generalization tool to account provisory for the optics of all the details in the image. Nevertheless, their correlation to the perceptible image quality remained rough. The transfer properties are also commonly expressed in a conventional form, related rather to a standardized image detail that is to substitute a provisory image in total. For example, the image noise is usually characterized at a constant brightness level although the brightness differences of image details can highly influence an actual noise level and related visual impression. The sharpness is commonly characterized for only the scene objects in a focus plane, not for a lot of details out of the plane.

The search for the correlation between the perceptible image quality and integral optical characteristics of the image is an unrestricted field of experimental activities due to the enormous number of concrete scenes. The results crucially depend on the images under test.

The failures seem to have a psychophysical origin. There exist typical scene objects of particular importance for observers.[1-7] Among those are a human face (especially, cheeks, eyes) sunlit and shadowed taken separately, white and black objects, greenery, blue sky, specular highlights. Some recent data have shown a crucial difference between such details and all the others in the required accuracy of their optical rendition.[4-6]

The photometry of dominant group has to be rendered within short limits that are easily overstepped by the current imaging systems. Other details have no effect on the expert quality judgments of image even if their photometry varies within much larger limits. That results in the dramatically different photometric tolerances to provide the same level of total image quality. Generally, when the dominants lie within their tolerances of the highest quality (and so do not affect the image quality), the other details find themselves automatically within their much wider tolerances.[8,9]

That does not mean that the optical differences in rendering the details could not be appreciated by vision. Under some artificial conditions, that can provide a direct comparison of two images of the same object, they may become quite perceptible.[4,6] An example may be the conventional colorimetry. It is aimed to determine the vision sensitivity to the smallest color differences under special viewing conditions that make them most perceptible and have nothing in common with the image observation. Under image viewing conditions, even much larger differences may become not only insignificant but often actually imperceptible.[4,6]

It seems to be the time to give up the thought that every object of a scene should be rendered by a perfect image with a definite optical properties. Each of them should lie within its own, often highly extensive tolerances, with no influence on the perceptible image quality. That is bound with the psychological and psychophysical factors.

Under the circumstances, the averaged rendition of all the image details within some even tolerances as it would be provided by the integral image characteristics may lower the total image quality for the dominant details may find themselves out of their short high-quality tolerance plateaus. The transfer properties of an imaging system should only provide the corresponding stimuli to meet their psychophysical tolerances.

Psychophysically Stipulated Mathematical Formalism for the Perceptible Image Quality

One considers the perceptible image quality, PQ, as an average of quality judgements of an image by a sufficient number of experts. It is normalized to the maximum value on the quality scale in order to vary from 0 to 1. To represent the total PQ with several PQ of dominant details, DPQ, may seem more complex than to correlate it approximately with an image feature integrated along the image area.

The new approach is distinctive, however, of exhausting the entire spread of experimental quality evaluations out of the error.[5,8,9] That has never been achieved before. The multiplicative interaction formalism of DPQs to describe the total image quality is quite simple and similar to that of the probabilities of independent events[8,9] in accordance with the probabilistic nature of learning and memorizing processes.[10]

Another advantage is that every DPQ stimulus of those already detected is a simply measurable optical property of an image detail. They are directly perceived by vision, e.g., the reflection ratios of some details. In the lightness impression (brightness scale) rendering, the decisive ratios are those of a dominant detail to a white object in the image.[3,4-7] In sharpness rendition, it seems to be the size ratio of eye iris image to the resolution.[11,12]

The third advantage is an obvious psychophysical sense of single-stimulus quality functions, SF, detected. Each of them found itself a tolerance function of the simplest kind. It consisted of a high-quality tolerance plateau and one or two portions of quality decrease with varying the stimulus outside the plateau.[6,8,9,11,12]

The plateaus have appeared to be more or less long, depending on the actual photometric statistics of an object and on the apportionment of discernible object types as they had been recorded in the memory of the typical observer.[8,9] An example is the known memory colors: skin, greenery and sky. The shortest plateau was observed for the white objects due to their especial psychophysical referent role.[1,3,7]

The plateaus seem to depend not only on the psychological importance of an object type but also on the psychological preferences of what way they would look most beautifully: the healthy pink skin, lush greenery, blue not dull sky, clean white. The consumer prefers rather a standardized color than an actual one.

The psychophysical nature of the approach is displayed clearly by the semi-logarithmic quality decrease.[9,11-13] The SF portions are linear on a log-stimulus scale: the difference of optical densities,[8,9] the log-resolution,[11] the log-granularity,[12] the log-ratio of total screen darkening by two subsequently shown slides.[13] That is characteristic to other psychophysical response-to-stimulus dependencies.

There appears to be no obvious restrictions to apply the formalism to every image feature of importance. In practice, the PQ not often reach the highest, impeccable level. The best images in a set lie generally at a lower level, reflecting the action of dominant details bound with other image features if they are out of their non-influence plateaus and do not vary in images considered.[9] The PQ interaction of different image features should be also multiplicative. The different features (e.g., color features, noise, sharpness) may be bound with different dominant details. The eye image size is critical for a total sharpness impression,[11,12] whereas the reflection coefficients of a cheek and a white cloud are critical for the brightness rendering.[8,9]

Most of image details are not critical for image quality impression due to their broad non-influence plateaus. Their photometry attracts no attention although it may be imprinted in visual memory. It may influence expert quality judgements after an inappropriate instruction. When observers are instructed not to judge the image quality as such („degree of image excellence“[14]), but that with respect to sharpness, noise, color or other image features, some previously insignificant details may act as the dominants.[12]

Observer's attention is then artificially attracted to the image details that highly display the influence of the image feature, directly mentioned in the instruction. Some other optical variations in extra details could become significant if the experts would be asked to judge not the quality but a perceptible degree of image sharpness, noise, color saturation or other features. The vision system is using then as a measuring device for a technical image property.

Even an instruction to judge image naturalness, "degree of correspondence between the reproduced image and reality (that is, the original scene as it is according to the viewer) instead of quality significantly influences the judgments.[14-16] The latter should involve into consideration not the same details as the instruction to judge the image quality. It provokes taking into account all the details whose photometric statistics has ever been impressed in memory, not only those of the highest psychological and psychophysical importance.[12]

Extracting a Complete Set of Single-stimulus Quality Functions

The SFs are of the same statistical nature as such psychophysical dependencies as the spectral sensitivity curve of human vision, or the perceptible thresholds of color differences at different wavelengths. They are capable of being statistically standardized on a representative set of human population. The number and properties of dominant SFs could be standardized for an image application field (mass consumer imaging, MCI, medical, environmental, documentation, educational or military imaging), and likely for an image type (portrait, landscape, animals in MCI).

The image type classifications should be limited by their practical frequency. For example, the images with people in different surroundings are 90% the MCI color images, flowers are 5%, and landscapes are still less. The practical considerations may be restricted by the two scene kinds: with people and without people. The mass consumer imaging is known to be predominantly adjusted for the first scene kind with people.

The absence of some dominants can dramatically affect the perception of other details. For instance, a landscape image of the same sharpness is judged considerably lower without a human face.[9] A sharp human face induces the highest quality judgements independently of background defocus.[11] Thus, the image kinds notable for the presence or absence of such unique psychological dominants as human face differ in the plateau length of their SFs.[4,6]

The SFs of dominant stimuli of a scene kind in an application field could be obtained either directly by single-stimulus tests,[5,9,12,13] or with a simple iterative procedure, applied to multi-stimuli data. If some stimuli do not vary within a representative set of images, they behave as a constant factor. It is equal to unity if all the stimuli lie within their high-quality tolerance plateaus and cannot influence the PQ under test at all.

If all the stimuli but one could be kept invariable within a test set, the PQ function should coincide with the SF of the single stimulus varied or differ from it by only a constant factor. Random image sets with single-stimulus image variations are seldom.[11-13] They could be composed, however, from the multi-stimuli data sets in the way below.[8,9]

A routine iterative procedure of SF extraction uses a large set of images with multi-stimuli image variations and needs generally several iteration steps.[8,9] It starts from the analysis of optical data spread in the images of the best quality. It appears obvious that the best images lie approximately on the plateaus of all the varying quality stimuli and their optical spread outlines some portions of the plateaus.

If all the optical stimuli but one are held within the range of the highest quality (only the images that possess the optical properties are considered), the experimental points should roughly outline the SF graph of the stimulus varied. One can do that way with another stimulus and hold the other stimuli, including those just considered, within the portions of their non-influence plateaus detected.[5,6]

The choice of optical stimulus in its psychophysically stipulated form is in part procedure-supported. The SFs of brightness scale rendering have transformed into the regular tolerance functions only after the stimuli had been expressed as the optical density differences between a dominant detail and a white object.[5,6] The form resulted also for the stimuli from other psychophysical experiments.[3]

The rough approximations obtained after the first iteration step can be made more accurate. One can repeat the same procedure with using now all the images of the set. To specify a SF needs then to use all the SFs but one, determined by the previous iteration, and to take into account their multiplicativity. The semi-logarithmic linearity of quality-decrease SF portions could be another helpful mean to reach soon the most accurate solution after few iterations.[8,9]

If some acting stimuli remain out of consideration, the entire exhaustion of substantial, out-of-error optical spread cannot be achieved. Such a situation has provoked a search for one more stimulus for the brightness scale investigation of color images.[8,9] The stimulus is bound with the specular highlights (the minimal densities in image). All the other SFs could not exhaust the entire out-of-error spread of the color images. Much lower fog in the monochrome images[5,6] has made the stimulus lie within its non-influence plateau.

To Predict Color Image and Produce Quality Simply and Precisely

The imaging aim in MCI is close to rendering a naturalness impression even if a definite difference between them has been shown.[14-16] In medical, technical or military imaging whose aim is to detect some details of not instinctive subconscious psychological preference. The imaging fields need to train reflex detail preferences. A professionally engaged observer seems to obey the same general psychophysical rules concerning the objects of professional interest, and the same quality mathematics.

Being of basic nature, the quality mathematics appears to be capable of being easily generalized on other measurable produce features and even other fields of quality evaluation along with the image features. One may so handle with the optical and mechanical properties of supports, with image size, frame shape, preferred image compositions (position of dominants, influence of large non-informative fields), image stability, ink and toner properties, and so on.

In the image quality, a pressing problem to solve is color rendition. The colorimetric approach has been a must at the initial stages of image color description for no alternative existed. The approach tried to express not the image quality as such but the image colors, that were observed in complex, psychologically influenced surroundings. It used the colorimetric formalism of artificial colors, observed in the simplified conditions of physical device. The latter crucially differ from the image viewing conditions, cannot allow for any psychological stimulus, and have little to do with actual image color impression.

The colorimetry is a perfect tool to evaluate the differential eye sensitivity to visible radiations of wavelengths compared that produce in a colorimeter two definite color impressions. It cannot explain the about normal complex multicolor images while acting on the eye with the light of only the two close wavelengths, 570 and 590 nm, under non-colorimetry, imaging conditions.[17]

The first step in the direction has been the rendition of lightness[8,9] as a color feature. A search in a similar way for the quality stimuli and the SFs of rendering hue and saturation should be the next step. The approach considered is quite capable to replace the infinite succession of rough colorimetry correlations to perceptible image quality by the precise mathematics of perceptible image quality as such.

Summary

A quality mathematics is shown to obey natural psychophysical laws and be based on proper rendering several typical dominant objects, each within its own photometric tolerances. It could provide an economical, accurate and comprehensive description and prediction of perceptible image quality, following from a few simple optical measurements.

References

1. E. Goldberg, *Der Aufbau des photographisches Bildes.* W.Knapp, Halle (Saale), S. 73 (1925).
2. C.N. Nelson. The Theory of Tone Reproduction. In *The Theory of the Photographic Process*, Eds C.E.K. Mees and T.H. James. Macmillan , N.Y., p.464 (1966).
3. C.J.Bartleson. Criterion for tone reproduction. *J. Opt. Soc. Am.,* **58**: 992-995 (1968).
4. V.V. Gavrik, G.L. Polovtseva, N.V. Lugovaya,and T.N. Muraviova. Constancy features of the best scene images with people. *Proc. S.I.Vavilov State Optics Inst.* **70**: #204, 105-114 (1988).
5. V.V. Gavrik, E.S. Ezerskaya and T.N. Muraviova. Photometric constancy properties of the best slide images with people. *Proc. S.I. Vavilov State Optics Inst.* **79**: #213, 97-104 (1991).

V.V. Gavrik. Principles of psychophysical and psychological correspondence in photographic scene imaging. *Sov. J. Opt. Technol.* (OSA) **58**: # 8, 688-692 (1991).

7. R.W.G. Hunt. Model of colour vision for predicting the appearance of colors under different viewing conditions. *Proc. SPIE,* **1909**: 2-14 (1993).
8. V.V. Gavrik. Optical tolerances for brightness scale reproduction in printing complex color images of natural scenes. *10th Congr. on Advances in Non-Impact Printing*, New Orleans, 587-591 (1994).
9. V.V. Gavrik. Probability properties and iterative isolation of independent single factor functions of visual image quality from multifactor photometric data. *J. Opt. Technol.* (OSA) **62**: #4, 232-237 (1995).
10. H. Plotkin. *The Nature of Knowledge: Concerning Adaptations, Instinct and the Evolution of Intelligence*. Allen/Lane Harvard Univ. Press, 1994, 269 p.
11. V.V. Gavrik, and L.N. Kirillova. A non-information sharpness influence on the visual quality of a color image with people. *J. Opt. Technol.* (OSA) **64**: #2 (1997).
12. V.V. Gavrik. Non-information aspects of naturalness rendering. *Int. Conf. on Digital Printing Technol.,* Seattle, 1997, 5 p.
13. V.V. Gavrik, and I.M. Sokolova. Effect of changes in screen illuminance on the quality of a color projection image. *Sov. J. Opt. Technol.* (OSA) **57**: # 8, 49-51 (1990).
14. H. de Ridder. Naturalness and Image Quality: Saturation and Lightness in Color Images of Natural Scenes. *J. Imag. Sci. Technol.* **40**: 487 (1996).
15. S.N. Yendrikhovski, F.J.J. Blommaert, H. de Ridder and S.T.J. van Helvoirt. Naturalness Judgements of Object Colours. *Proc. 8th Congr. of Intern. Colour Assoc.,* Kyoto, **1**: 163-166 (1997).
16. H. de Ridder. Naturalness and Image Quality: Hue, Saturation and Lightness Variation in Color Images of Natural Scenes. *Proc. 8th Congr. of Intern. Colour Assoc.,* Kyoto, **2**: 658-660 (1997).
17. E. Land. Experiments in Color Vision. *Scient. Amer.*, **200**, 84-94, 96-99 (1959).