A Review of Image Quality Assessment with application to Computational Photography

Henri Maître
LTCI, Telecom ParisTech, Université Paris-Saclay
46 rue Barrault, 75013-Paris, France

ABSTRACT

Image quality assessment has been of major importance for several domains of the industry of image as for instance restoration or communication and coding. New application fields are opening today with the increase of embedded power in the camera and the emergence of computational photography: automatic tuning, image selection, image fusion, image data-base building, etc.

We review the literature of image quality evaluation. We pay attention to the very different underlying hypotheses and results of the existing methods to approach the problem. We explain why they differ and for which applications they may be beneficial. We also underline their limits, especially for a possible use in the novel domain of computational photography. Being developed to address different objectives, they propose answers on different aspects, which make them sometimes complementary. However, they all remain limited in their capability to challenge the human expert, the said or unsaid ultimate goal.

We consider the methods which are based on retrieving the parameters of a signal, mostly in spectral analysis; then we explore the more global methods to qualify the image quality in terms of noticeable defects or degradation as popular in the compression domain; in a third field the image acquisition process is considered as a channel between the source and the receiver, allowing to use the tools of the information theory and to qualify the system in terms of entropy and information capacity.

However, these different approaches hardly attack the most difficult part of the task which is tomeasure the quality of the photography in terms of aesthetic properties. To help in addressing this problem, in between Philosophy, Biology and Psychology, we propose a brief review of the literature which addresses the problematic of qualifying Beauty, present the attempts to adapt these concepts to visual patterns and initiate a reflection on what could be done in the field of photography.

Keywords: Image quality assessment, IQA, Information Theory, Computational Photography, Aesthetic

1. INTRODUCTION

Image quality has been the object of a great interest and excellent studies may be tracked back as early as in the 70s. It has been explored mostly to optimize the different stages of the image production, processing and communication, so that these stages may be efficiently operational.

In recent days, imbedded facilities on-board the camera, as well as efficient networking solutions allow to consider new possible applications where quality assessment would be involved very early in the production of the image, either by tuning the acquisition parameters or by combining several acquisitions. Moreover, these methods could be used in the image-life and participate to tasks where the human observer is today highly demanded: image ranking, image selection, résumé building, data-base annotating and editing, etc.

When looking back to image quality evaluation, we observe that some very different approaches exist in the literature. Three of them at least are familiar to engineers in the field of image processing. Although they share some elementary tools they have progressively migrated towards quite different solutions and do not maintain strong connections, because of their different objectives. Therefore, it is often difficult to simultaneously benefit from the best results gained in these different fields. The purpose of this presentation is first to clarify the reasons

Further author information: send correspondence to Henri Maître - E-mail: henri.maitre@telecom-paristech.fr, Telephone: +33 145 817 655

of these domains to have been separated, then to go through each field, to resume the most important results it hosts, to discuss on which benefits we may hope for when using it and to review their limits.

We will show that a common limit comes from the fact that they hardly address a dimension of major importance to evaluate an image: its aesthetic content, i.e. the pleasure we have at looking at it, or the strength of the emotions it awakes in the observer's mind. However this domain is not a desert of academic studies. It is paved with at least one century of scientific publications (and even more if we trace it back to Plato and Aristotle). However, as most of the existing literature is issued from the Humanities departments (Psychology, Sociology, Philosophy, Fine Arts), it is quite difficult for the computer scientist to translate it in algorithms. We will briefly offer a review of some important results for our purpose and present the existing extension of these results in the field of automatic quality evaluation.

1.1 Image quality: the Grand Picture

The first domain of interest is the field of experimental physics and instrumentation which abundantly uses signal processing tools as developed in Fourier analysis: spectral density function, point spread function, signal-noise ratio, etc. It nourishes abundantly the specific literature dealing with optimizing the sensor and the optical system, where, for instance, people is interested in compromising between f-number and exposure time, or ISO sensitivity and f-number. It purveys the most used tools for comparing camera performances in specialized journals for photographers or scientists or for choosing an adequate sensor for a given application. It relies on objective and deterministic physical measurements obtained from experimental benches, with normalized targets and controlled physical parameters. It does not pay attention to the image, nor to the observer. Its results are expected to be valid for any image and deterministically established.

The second domain has been developed mostly for broadcasting (either with still images or with videos) when the community is looking for the best visual quality under practical constraints on bandpass, computation power and storage facilities. It may also take into account viewing conditions and economical balance. These methods pay a great attention to the human visual system and consider the image as a whole. They often balance various types of drawbacks to evaluate a global judgment. Based on shared catalogs of image impairments, they propose heuristic formulas (often inspired from biological or psychophisical experiences) which adequately combine pooled local and specific measures. They generally take their value in a statistical sense, when many different experiments are done with various and significant samples of images, and sufficient observer cliques.

The third family of approach is based on Information Theory. Within this framework, it is possible to provide a global modeling of the acquisition process, by considering the scene to be observed as an emitter, the digital photo as the received message and the camera, with its specific set-ups, as a channel possibly impaired with transmission impairments (noise, diffraction, focusing errors, integration on the sensor, etc.). Although rather difficult to instantiate, this approach is prone of fruitful results as shown in the recent literature. It provides bounds on the performances we may expect from a system and allows for a global optimization of the parameters. However it is not yet accepted by the community for well justified reasons.

2. SIGNAL PROCESSING APPROACH

The approach based on the measurement of the signal content of images obtained from reference images or from targets is the most used in the domain of photography to characterize the whole photographic process, i.e. mostly the optical lens, the sensor and the software used to reconstruct the image.[1]

2.1 Signal-noise ratio

More than the Signal-Noise-Ratio (SNR), the Peak-Signal-Noise-Ratio (PSNR) is employed, to take advantage of the full dynamics of the signal. Noise may come from different sources. Quantization noise is often negligible. Alone it would result in a 59 dB PSNR for an 8 bit image. Compression noise (for instance using JPEG or JPEG2000) is usually under the user's control and may even be set to zero by using either loss-less compression or RAW format which is available on many cameras. The main sources of noise are therefore due to the photon flow or to the electronic process to detect and store the image.[2,3]

A useful heuristic result has been provided about photon noise, linking this noise to the user's perception. It states that for images with more than 1 000 photons per pixel, the photon noise is usually not perceived, even in flat areas.[4]

Photon noise and electronic noise are often hidden to the user. He/she may only adjust the camera to the lightness conditions by adapting the camera ISO sensitivity. When the camera complies to the ISO[5] norm and provides a sensitivity $\mathcal{S}_{40}$) (measured at 40 dB SNR) and when it is exposed with an incident energy $\mathcal{E}$ (expressed in lux.second), we may expect an SNR (Figure 1 on the left) equal to:

$$SNR(\mathcal{E}) = \frac{40 \log_{10}(\mathcal{E})}{\log_{10}(\mathcal{S}_{40})} \tag{1}$$

2.2 Resolution and MTF

Resolution is another important parameter of the signal processing description of the camera. Although it is greatly conditioned by the sensor geometry, it cannot be reduced solely to it (whatever the commercial turmoil around high-resolution sensors). Optical aberrations, diffraction by the lens as well as misfocussing (Figure 1 on the right) may reduce the signal bandwidth before detection. These defects may be rather well identified using optical benches and adequate targets. Coupling between neighboring detection sites and cross-correlation between pixels when detecting them or when demosaicking the image will further diminish the resolution. They are more difficult to measure, but may be included in global measurements following well established protocols,[6] and specific targets: either geometrical[7] or random.[8–10]

When the MTF has been measured, a value of resolution may be deduced. It has also been normalized[6] using 5% contrast as a criterion on binary targets. This figure is quite low (when compared for instance with Rayleigh criterion,[11] providing optimistic values which may be unreliable in case of noise. Resolution is often a too drastic reduction of MTF to be usefull for engineers.

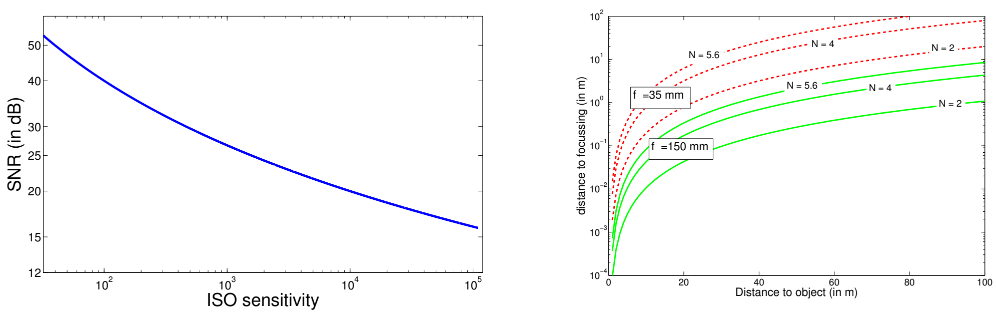

Figure 1. On the left: Signal Noise ratio (in dB) as a function of ISO speed from 32 to 100 000 ISO, and a constant energy (here 100 lux.seconds). On the right, defocusing distance, as a function of the object distance, which induces a smear equal to the diffraction of the lens for 2 focal length (35 MM and 150 MM) and 3 f-numbers (2, 4 and 5.6). Above the curve the focusing distance is dominant, below the curve, the diffraction is dominant.

The Modulation Transfer Function (MTF), $H(u, v)$, function of the two spatial frequencies u and v associated to the space variable x and y, concentrates a great part of the information necessary to the user in a linear signal processing approach. Therefore, it may be found in most of the journals devoted to engineering photographic systems 2. However, it requires an expert knowledge on image processing to be translated into operational image quality. Moreover, MTF, even reduced to 1D by exploiting circular symmetry (not often verified), or to 2 times 1D by reducing it to the two horizontal and vertical resolutions, remains a cumbersome description difficult to exploit. On the contrary, its reduction to MTF50[2] (or MTF50P), as its reduction to the only resolution value are often too crude.

2.3 Acutance

To better adapt the evaluation to the image quality and resume it in a few numbers, human perceptual system has been invoked along with MTF. Under well defined experimental conditions (observation distance, ambient

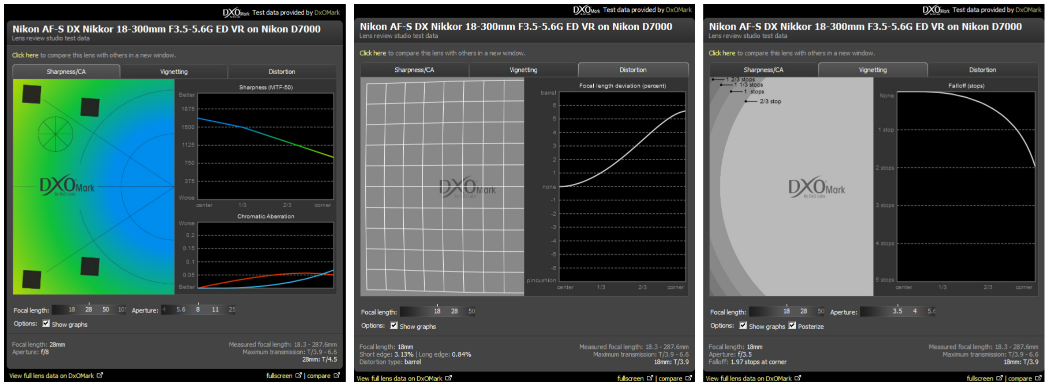

Figure 2. An example of camera analysis as obtained from the DxO website. On the left: MTF, Middle: Geometrical distortion. On the right: Vignetting effect.

lighting, contrast, etc.), one knows the *Spatial Contrast Sensibility Function* (SCSF) which expresses the capacity of human visual system to perceive the a target contrast with given frequency content. Depending on the orientation of the frequency, this SCSF function, as expressed by Mannos & Sakrison,[12] takes the form:

$$\begin{array}{lll} \phi(u_\theta) & = 2,6 \times (0,0192 + 0,144 u_\theta) \exp(-0,114 u_\theta)^{1,1} & \text{iff } u_\theta \geq u_{min} = 8 \\ \phi(u_\theta) & = 1 & \text{else} \end{array} \tag{2}$$

where u_θ is the spatial frequency in direction θ. It depends on θ according to the formula[13] :

$$u_\theta = \frac{u}{0,15\cos(4\theta) + 0,85} \tag{3}$$

Using both the SCSF and the MTF, the *acutance* A computes the average of the MTF weighted by the visibility of the frequency:

$$A = \iint_{-u_{max/2}, -v_{max/2}}^{u_{max/2}, v_{max/2}} \phi(u,v)\; H(u,v)\; du dv \tag{4}$$

This unique figure is a good characteristics of the acquisition system under precise observation conditions. It will allow to compare different cameras. However it does not inform about noise and geometric or colorimetric aberrations. Moreover, it is rather easy to boost the acutance in the post-processing stage of the acquisition system for instance with the well known techniques of *unsharp masking*,[14] where middle range frequencies are amplified at any frequency where SCSF is important. Such a processing is often available on commercial cameras, where, in the best cases, this processing may be tuned by the customer, but some times it is fixed. A high acutance may derive from an amplification of noise and be detrimental to the image quality. Some indicators have been derived both from MTF and from acutance to alleviate the noise problem which often artificially extends the image band-pass. However none of this criteria has been recognized yet as a universal indicator which can be substituted to MTF.

2.4 Signal processing tool: conclusions

Signal processing tools are adequate to characterize the properties of the optical capture system or the properties of a given image. They are mandatory for many image processing operations like filtering and restoration where they allow an excellent control on specific attributes of the quality. Therefore, they have been intensively used inside the camera to provide an image with high contrast, good resolution and low noise. However, even with the help of SCSF, they do not say much about the perceived qualities of the image as seen by a user, because they are only concerned with linear defects (geometrical distortions are therefore not taken into account and requires specific measurement tools as for instance grid targets).

	Quality scale		Impairment scale
5	excellent	5	imperceptible
4	good	4	perceptible, but not annoying
3	fair	3	slightly annoying
2	poor	2	annoying
1	bad	1	very annoying

Table 1. The ITU6R BT 500-13 quality and impairment scale table.

3. GLOBAL QUALITY AND IMAGE QUALITY ASSESSMENT

When developing image communication services (either for still images or for video), the television and telecommunication industry strongly needed global estimators of the image quality, able to replace the human observer in optimizing the processing by a statistical test on large data-bases. Therefore such estimators have been developed since the 1970s. They received improvements in the next 50 years and are still under research in several based, by exploiting new tools issued from machine learning and big-data.

3.1 Global quality reference

At first, a range of quality with 5 levels have been adopted by ITU, from *Excellent* to *Poor* as well as a range of degradations[15] (Table 1). They allow to share between users a common vocabulary to determine the class of degradations which affect the image.

Various strategies have been elaborated to grade the image quality, either with the knowledge of a reference (the ideal image without degradation) or without knowledge of the reference (the so called *blind evaluation*)*. The domain covering these techniques is known as IQA (*image quality assessment*) and reviews may be found in the tutorials from R. Chandler[16] and A. Bovik.[17]

3.2 IQA

3.2.1 With a reference

Methods based on the knowledge of the original image abundantly used the Euclidean distance between images, weighted or not by some visual criteria[18] (for instance using the SCSF). When dealing with color images, a special attention is given to the color distances as determined by the early work of MacAdam.[19] These methods[20–22] are fast and universal, but they do not provide very good results. Recent extensions make use of more complex psycho-visual distances[23] where the defects are weighted, depending on the presence of textures or on their distance to contours to take into consideration the so called "masking effect". More recent techniques benefit from machine learning approaches, using intensive trainings on large labeled data-bases.[24] These improvements resulted in much more reliable evaluations than the previous methods measured as the conformity of the evaluation mark compared to the one attributed by a supervised evaluation.

A typical scheme of reference-based IQA is presented on figure 3, where a psycho-physiological model is used in each of the boxes of the processing.

However, methods with reference are not the most adapted for photography. An exception is when dealing with a set of images of the same scene, but with different tunings (F-number, focusing, exposure time), as for instance using the *burst mode* of the camera. In this case, the reference is often not known. Moreover some parts of image 1 may be a reference for the sequence when another part of image 2 would be the reference somewhere else. Local quality estimators are therefore needed and strategies to combine several estimators have to be found.

*Some methods have also been developed using a partial information on the original image, called *reduced information methods*. They will not be considered here.

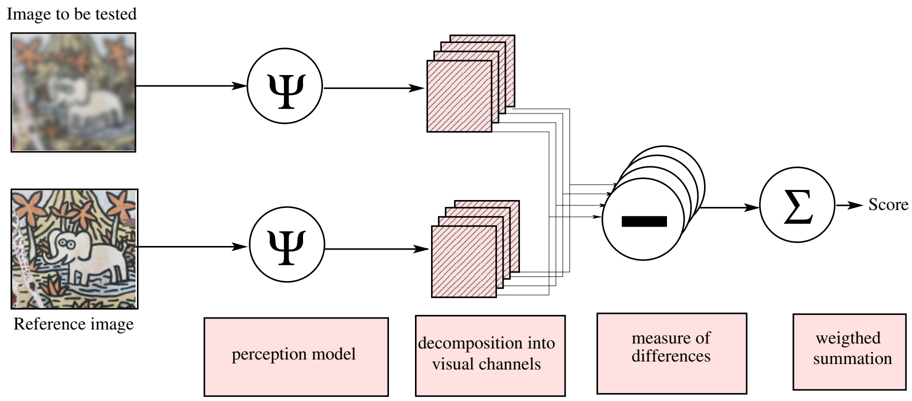

Figure 3. Image quality assessment based on a reference image and a perception model.

3.3 Reference-less methods

Reference-less methods are of greatest interest for computational photography. Two broad families of methods have been developed: feature quality estimation and example-based methods.

3.3.1 Feature estimation

In this approach, a catalog of image quality or, on the contrary, a catalog of the possible defects, is proposed and each element of the catalog is separately evaluated. Then results are *pooled* to provide a unique mark, i.e., adequately combined. Pooling is usually made with psycho-visual criteria.[25] The main attributes of the catalogs concern the contrast (or the dynamic, or the color palette), the noise and the blur.

Noise is often estimated from the auto-correlation function of the image (in most of the image models, it appears as a discontinuity at the spectral power $O(0,0)$ frequency), but it may also be deduced from the internal parameters of the camera (sensitivity adjustment of the camera which is often available in the EXIF file attached to the image (Figure 1 left). Blur is more complex to estimate. It is often measured from the contrast profile of lines perpendicular to edges.[10,26]

Figure 4. Two images with different focusing distances (top row). Local quality estimation based on the NIQE algorithm[20] (bottom row). The darkest values code the best image quality.

Some original techniques have also been developed, based on a completely different approach to estimate the focusing quality, namely the *global phase coherence*.[27,28] Global phase coherence $\mathcal{G}$ measures the probability that the total variation of the image i_ϕ perturbed by a defect ϕ be less than the total variation of the original image $i(x,y)$:

$$\mathcal{G} = -\log_{10} Proba\left[TV(i_\phi) \leq TV(i)\right] \tag{5}$$

where the total variation $TV(i_\phi$ of image $i(x,y)$ is expressed by:

$$TV(i(x,y)) = \iint |\nabla i(x,y)|[dx\ ,\ dy] \tag{6}$$

It may be shown that global phase coherence is very sensitive to blur and provides a good index of quality. However it is rather heavy to accurately compute.[29] A simplified sharpness index was proposed,[30] which benefits from the properties of $\mathcal{G}$ but as the result of a much more tractable computation.

3.3.2 Statistical properties and Example based methods

These methods are exploiting statistical properties[31,32] of natural images[33,34] which have been collected on general or specific data bases. They may also use machine learning approaches, trained with images affected with known defects.[35] They provide results in good agreement with the user's judgement. They may also inform about the impairing defects when it belongs to a known family of defects.[33]

However, in pooling the results collected on the whole image, the global methods are not well adapted to offer a detailed diagnosis on the image quality. They also hardly integrate subtle masking effects between defects which are rather important in the visual process. Great progresses would certainly result from integrating more advanced models of the human perception with some kind of aesthetic evaluation.

4. INFORMATION THEORETIC APPROACHES

4.1 Information capacity, degrees of freedom

This approach was initially developed by physicists in the framework of optical system optimization[36–38] for astronomy or for microscopy. Looking in the camera on both sides of the CCD or CMOS sensor, can estimate two different flows of information:

- just after the measurement by the sensor, a Black & White signal is reduced to $N = n_p \times n_l$ digital values (the number of photo-sites), each coded on n bits affected by a noise which may depend on the photo-site position (because of a possible variable gain or vignetting). The information capacity, number of degrees of freedom or Shannon number is:

$$\nu_1 = \sum_{i \leq n_p, j \leq n_l} 2^n / (max(1, \sigma(i,j))) \tag{7}$$

- just before the sensor, the optical flow is fully determined by the lens aperture N, the field angle and the beam size S. In this case of incoherent optics which concerns photography, the information capacity is given by:

$$\nu_2 = \frac{4SD^2}{(\lambda f)^2} = \frac{4S}{N^2\lambda^2} \tag{8}$$

  where λ is the wavelength.

From these two expressions, we may deduce the configuration which optimizes the configuration by equaling the performances of optical elements and sensor (Figure 5).

In the case of color images, the computation is slightly more complex since we should not only take into account the 3 primary colors ($\lambda = 0.460\ nm$, $0.540\ nm$ and $0.640\ nm$) of the sRGB standard, but also consider the possible correlations between channels. The 17321 ISO norm[39] provides a complete protocol to calibrate a color sensor.

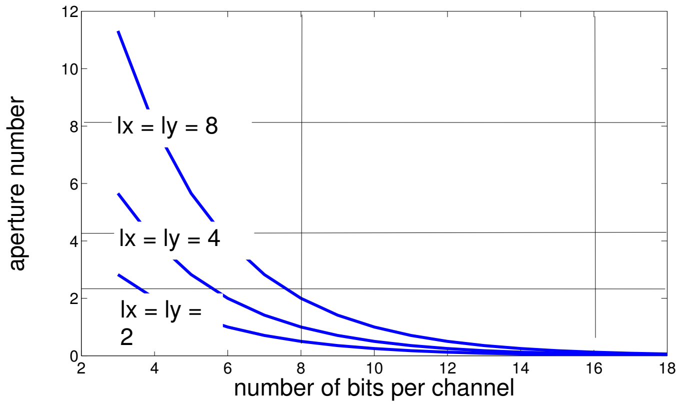

Figure 5. The curves express the relation between the number of bits assigned to each pixel and the f-number of the optical element, in the case where the sensor and lens match perfectly, i.e. the sensor transmits exactly the same number of degrees of freedom recognized by the lens. Above these curves, the sensor is over-dimensioned; below the curves, the sensor is under-dimensioned. Results shown are for a green wavelength (500 nm) and 3 dimensions of photo-sites expressed in micro-meters.

4.1.1 Entropy and Information capacity

If we consider the image as an independent event process[†], its entropy $\mathcal{E}$ (according to Shannon) is given by:

$$\mathcal{E} = -\sum_{k=1}^{K} p(k)\log(p(k)) \tag{9}$$

The information capacity is therefore $\mathcal{C} = N\mathcal{E}$. Due to the different filters which alter the image (at least due to the limited diameter of the optical lens and the integration on the sensor elementary photo-sites), the image suffers a loss of entropy $\Delta\mathcal{E}$ which may be related to the MTF $H(u,v)$ of the process:[40,41]

$$\Delta\mathcal{E} = \frac{1}{B_x B_y}\iint_{-B_x/2,-B_y/2}^{B_x/2,B_y/2} \log(|H(u,v)|^2)dudv \tag{10}$$

When the defect is due to the optical lens and the detection by the photo-site, H is equal to:

$$H(u,v) = H_s(u,v)H_f(u,v) \tag{11}$$

with $H_s = \mathrm{sinc}(2\pi u\delta x).\mathrm{sinc}(2\pi v\delta y)$ is the integration on the photo-site, and:

$$H_f(u,v,\lambda) = \left[\frac{2J_1(\pi\rho D/\lambda f)}{\pi\rho D/\lambda f}\right]^2 \qquad \text{with} \qquad \rho = \sqrt{u^2+v^2} \tag{12}$$

is the optical MTF due to a lens of diameter D and focal length f (J_1 is the modified Bessel function of first kind).[11] Taking into account the detection on a Bayer-like mosaic, with inter-weaved Red, Green and Blue channels, the information capacity for the Green channel is given by:

$$\mathcal{E}_G = 4\sum_{i,j}\iint_{-\frac{1}{8},-\frac{1}{8}}^{\frac{1}{8},\frac{1}{8}} \max[0, n_1(i,j,G) + 2[\log_2(|H_s(u,v)| + \log_2(|H_f(u,v,\lambda_V)|)]]dudv \tag{13}$$

with $n_1(i,j,G)$ the true number of levels in channel G at position i,j as given by equation 7 and similar values (but with different integration limits), for the Red and Blue channels.[42]

[†]This approximation is crude, but finer models exist which consider correlations between pixels which would not provide very different conclusions.

Such calculations are used[43,44] with a much more complete description of the possible degradations coming either from lens or from electronic processing. They allow to accurately evaluate the performances of complete real photographic cameras (optical body + lens), by integrating their measured distortions and limitations. This results in a unique comparison of real cameras potential to acquire information, each of them being used with the best combination of sensitivity and aperture to optimize the information capacity.

From these information theoretic-based methods, we access to a different characterization of the image than with signal processing or global methods. Alike for signal processing approaches, they provide an objective answer which is not limited by heuristic weighting factors and more or less empirical criteria as for IQA techniques. In this case, the "quality" is hidden behind the information content in the same as in signal-processing approaches it is limited to SNR and power spectrum.

5. AESTHETIC POINTS OF VIEW

However, photography is not just the technical acquisition of an image (figure 6). Since the pioneering works of the mid-19th century, a great attention has been paid by many photographers to elevate photography at the level of an Art, as Gustave Le Gray, Eadweard Muybridge, Eugéne Disdéri, Constant Peyo, Oscar Rejlander, Henri Peach Robinson or Karl Blossfeldt.

Their attempt as well as the efforts of many others, fully succeeded, as testified by the popular place occupied today by museums, galleries and exhibitions devoted to photography. It occupies a recognized rank in the cultural life everywhere in the world, aside other fine-arts institutions like painting, drawing or architecture.

Figure 6. G. Garbo, L. Bacall, M. Monroe, E. Taylor. Which image is nicer? In this precise case, photographic beauty is a complex matter which is greatly obscured by personal tastes for one character or another. But it also presents an objective part which is not only a matter of technical quality expressed as band-path and signal noise ratio. It is a complex blend of many different criteria which are usually covered under the global name of aesthetic.

As other fine-arts, photography obeys to more or less known aesthetic rules which make the difference between an amateur's souvenir and a professional photo. Could we take an inventory of these rules? Could we incorporate these rules into the evaluation algorithms presented in previous Sections? Adding an aesthetic dimension to the criteria developed there would probably be quite beneficial in many applications that we have in mind as image selection, automatic labeling, data-base construction, etc.

But the task is not an easy one. At first, we have first to notice that no theory of aesthetic, whatever the domain, is today ready for a translation in engineering words and programs. Then, when adapting some elements of explored approaches, to the photographic domain, we observe that, more than any other fine art, photography is exposed to the fast evolution of the technology which makes it very sensitive to temporal, societal and cultural influences. Therefore, aesthetic criteria are many and far from being universal.

Let us have however a brief review of some tracks which are explored in attempting to provide some elements of rationality in the studies on aesthetic. We will follow several well known texts.[45,46] At first let us note that none of the approaches which have been proposed are dedicated to photography only. Often, they consider the broadest field of application of the term "Beauty" without application domain. In the most favorable situation, they concern the neighboring domain of painting or drawing which are the closest from ours.

5.1 Objectivist vs Subjectivist theories

An exceptional place is occupied by the theory of harmony as developed by Greek philosophers: Aristotle and Plato, mainly with the help of mathematicians like Pythagoras. The theory they developed, named as **objectivist**, considers Beauty as only due to the object itself without any contribution from the observer. Therefore Beauty is universal and intrinsically sharable. In its original form, it granted great importance to the internal construction of the objects (being music, architectural monuments, statues or faces) seen as harmonious when made of sub-parts or components scaling in "harmonious" (i.e. simple) ratios: 1/2, 2/3, 1/4, etc. Under this light photos (as paintings) are concerned not only by the internal organization of their sub-parts, but also by the distribution of colors (although Newton's and Grassman's experiences in color considerably reduced the relevance of this conclusion on this domain). The sequel of the harmony theory is quite important in the history of painting and architecture.[47,48] The only sequel in photo may be traced in the "rule of the third" to locate the object of interest.

During the last century, objectivist approach received a great attention[46,49,50] to define aesthetic. Many other attributes than the internal construction have been collected to participate to the Beauty, like symmetries, balance, contrast, clarity, color, complexity, etc.

Well suited to our purpose, the objectivist theory let aside several problems as for instance the role of the observer in the image evaluation. Moreover objectivist theory requires long experimental statistical studies[51,52] to determine the contribution of each feature in the global evaluation.

Subjectivist approach, on the contrary, in its extreme form, let the only charge of attributing beauty qualifications to the observer; therefore, it is individual, non-sharable and plastic (i.e. exposed to the specific context of the observation and possibly changing in time for an identical stimulus). Well defended in the 19th century in this extreme form under the influence of the Romantic movement, which let no room for our purpose, subjectivist approach have progressively incorporated elements of objectivist considerations to create a modern *interactionist theory* where both object and subject contribute to the emergence of the appraisal. Studies in the domain of neurobiology at the end of last century, fueled the subjectivist approaches with novel ingredients: aesthetic values results from an adequation of the visual stimuli (as proposed by the image) with the cortical structures in charge of their processing,[53,54] i.e. the visual channels, and the visual cortex.

Within these modern approaches, aesthetic judgement is made of two components: pertinent stimuli to create the arousal (proposed by the examined object) and adequate appraisal from the observer.[55] Beauty arrive when the arousal is well adapted to the visual ways.

5.2 Gestalt and Information Theory

The earliest works which addressed explicitly Beauty in Art, in a mathematical framework, mixed arousal and appraisal in a unique approach by means of empirical models which try to mimic the human evaluation, weighting the apparent order of the image by its complexity.[49] Complexity is supposed to express the effort of the observer to interpret the image, while apparent order eases the comprehension. These two ingredients well reflect the idea of richness from one side and clarity on the other. Although premonitory and seminal, these early attempts adopted rather simplistic formulations. For instance, Birkhoff expression of beauty $\mathcal{B}$ for simple shapes was measured by a ratio $\mathcal{B}$:

$$\mathcal{B} = S/C \tag{14}$$

where S expresses the order or regularity in the shape and C its complexity.[56] However, in precise situations, Birkhoff had to derive ad'hoc formulae for every specific case (for polygons or for curved shapes for instance).

Simultaneously developed from strong psychological bases, Gestalt theory offered a more universal framework, with both arousal and appraisal hints.[46,50] The Gestalt toolbox contains a set of criteria to determine what is easily perceived by the visual system: similarity, continuity, symmetry, etc. Admittedly appraisal is not well described and more orientated towards perception than towards aesthetic judgement. Moreover Gestalt, paying a great attention to lines, is better suited to express recognition of drawings than to photographs. In the last decades, the attempts to extend Gestalt theory to image processing by means of *a-contrario theory*[57,58] is likely to provide the wanted framework to derive aesthetic indicators for image processing, but much more has to be made to establish sound formulations in this framework.

Information theory based on Shannon entropy as seen in Section 4 suggests very satisfying appraisal tools by considering the vision (or photographic) process as a channel in between a source (the object) and a receiver. Fed with Gestalt information it provides explicit formulations in terms of information capacity.[59, 60]

For instance Bense[60] proposed to revisit the suggestions of Birkhoff,[49] by using the redundancy or mutual information to express the order measure, and the entropy to express the complexity, in equation 14. Moreover he proposed to adopt a hierarchical approach to express different possible appearances at different scales.

However, Shannon's theory is not well adapted to express the complexity as demonstrated by Kolmogorov, Solomonoff, Levin and Chaitin.[61] The recurs to more adapted formulations of the complexity of images, namely the **calculability theory** and the **algorithmic information theory** seems to be a promising track although it raises many difficulties. Early results in the field of image quality evaluation are mostly methodological.[62] The most advanced works,[63] however provide the *aesthetic ranking* of masterpieces of 19th and 20th century painters, making use of Zurek's physical entropy.[64] They mostly propose a proof of concept which remains to be worked out. Kolmogorov K complexity being not calculable, some approximations are to be done. A rather simple and efficient way is by means of loss less compression schemes (Lampel-Ziv) which may appear as upper bounds of the minimal length algorithm. Let $\tilde{K}$ be the length of the compressed image Equation 14 may be rewritten:

$$\mathcal{B}_{lz} = \frac{N\mathcal{E} - \tilde{K}}{N\mathcal{E}} \tag{15}$$

where N is the number of pixels and $\mathcal{E}$ is given by equation 9. However, we can argue that compressing image with loss (as performed for instance with JPEG coder) are coming closer to the Kolmogorov limit, providing that the user do not notice the difference. let K_{jpeg} be this value. It provides a quality measured by:

$$\mathcal{B}_{jpeg} = \frac{N\mathcal{E} - K_{jpeg}}{N\mathcal{E}} \tag{16}$$

Figure 7. Four images to be compared: Amiens, Bannes, Larmor, Vinzelles.

Although it would have been expected that we test these ideas with the photos 6, we prefer to choose another experimental matter to make it, since we do not control the original technical level of figures 6. On the contrary, photos from figure 7 have been taken with the same camera and under similar conditions (sensibility, aperture, focusing, etc.) so that they behave approximately identically under the criteria of Sections 1 and 2. We present on the plot of figure 8 the quality evaluation of the 4 images of figure 7, following the initial ideas presented in.[63] It would be hard from these simple example to confirm the validity of the approach which needs to be deeply revisited to be adapted to photos when it was derived for paintings.

Some extensions are provided by Rigau et al.[63] where the image is segmented in sub parts to take into account the compositional complexity of the image. Other extensions may relate the image entropy to the human perception by considering color distances and masking and grouping effects.

These early and innovative studies are clearing the ground of adapting aesthetic theory to the domain of image processing. However, much work has still to be done to comfort the results according to the human judgement.

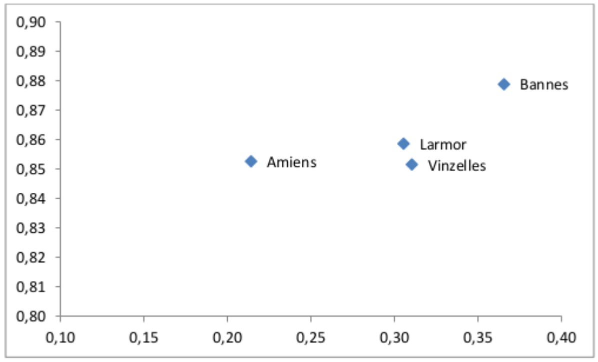

Figure 8. The marks obtained using equation 15 (Kolmogorov complexity evaluated from using Lampel-Ziv compression method) or using equation 16, (Kolmogorov complexity evaluated using a lossy compression) on the four images of Figure 7. $\mathcal{B}_{lz}$ along horizontal axis and $\mathcal{B}_{jpeg}$ along vertical one.

6. CONCLUSION

Evaluation of the quality of photos with an arsenal of criteria similar to what a human expert may propose is an urgent need for the autonomous exploitation of very large collections of images. It is also required in several applications of modern photography, as for instance for extending the focus, for fusing images, for detecting evolutions, etc.

The greatest demand is today to improve the quality of the appraisal in three directions:

- at first it must adhere to the human estimation of the quality, which is quite subtle and context-depending (for instance, a deep range of focusing is mostly a demand in photography, to allow an accurate imaging of the different objects in the field, however, when dealing with portraits, the demand is that the background remains blurred so that the attention is focused on the sole face);
- then it must be reconfigured to adapt to the specific need the user may have;
- at last it must ”explain” the appreciation it gives, so that the different intermediary results be clarified, both from a spatial point of view (where is the problem) and from a identification point of view (which defect).

Image processing is able to propose most of the components of the quality evaluation: criteria are rather well known, measures are reliable, pooling techniques exist to combine these methods. Even the perceptual models exist. However, for the moment, these models are far too close from the signal and do not take into account a more symbolic or semantic representation of the image. The image processing community should pay attention to these aspects. Learning techniques could be the solution, but we may also suggest to take advantage of the many results obtained by the community in psycho-vision, psychology, neurobiology and art-philosophy.

REFERENCES

[1] Maitre, H., [*From Photon to pixel: the digital camera handbook*], ISTE-Wiley, London (Great-Britain) (2015).

[2] Farrell, J., Xiao, F., and Kavusi, S., “Resolution and light sensitivity tradeoff with pixel size,” in [*Proceedings of SPIE Conference on Digital Photography*], **6069**(II), 211–218, SPIE (2006).

[3] Aguerrebere, C., Delon, J., Gousseau, Y., and Musé, P., “Study of the digital camera acquisition process and statistical modeling of the sensor raw data,” HAL-00733538-v3 (2012).

[4] Xiao, F., Farrel, J., and Wandell, B., “Psychophysical thresholds and digital camera sensitivity: The thousand photon limit,” in [*Proc SPIE*], (5678), 75–84, SPIE (2005).

[5] ISO, “Norme ISO 9848 - Photography microfilms - Determination of ISO speed.” norme ISO (2003).

[6] ISO, “Norme ISO 12233-2000(E) - Photography – Electronic still picture cameras – Resolution measurements.” norme ISO (2000).

[7] Joshi, N., Szeliski, R., and Kriegman, D., “PSF estimation using sharp edge prediction,” in [*IEEE conference on Computer Vision & Pattern Recognition (CVPR 2008)*], 1–8 (2008).

[8] Cao, F., Guichard, F., and Hornung, H., “Dead leaves model for measuring texture quality on a digital camera,” in [*Digital Photography VI, Proceedings SPIE, Vol. 7537*], Imai, F., Sampat, N., and Xiao, F., eds., **7537**, 75370E–75370E–8 (2010).

[9] McElvain, J., Campbell, S., and Miller, J., “Texture-based measurement of spatial frequency response using the dead leaves target: extensions, and application to real camera systems,” in [*Electronic Imaging Processing SPIE 7537, 75370D*], (2010).

[10] Delbracio, M., Almansa, A., Musé, P., and Morel, J.-M., “Subpixel point spread function estimation from two photographs at different distances,” *SIAM Journal on Imaging Science* **5 (4)**, 1234–1260 (novembre 2012).

[11] Smith, W. J., [*Modern Optical Engineering*], Mc-Graw-Hill, New-York, Etats-Unis (1990).

[12] Mannos, J. and Sakrison, D., “The effects of visual fidelity criterion on the encoding of images,” *IEEE transactions on Information Theory* **20**(4), 525–536 (1974).

[13] Daly, J., “Application of noise-adaptive contrast sensitivity function to image data compression,” *Optical Engineering* **29**, 977–987 (1990).

[14] Fisher, R., Dawson-Howe, K., Fitzgibbon, A., Robertson, C., and Trucco, E., [*Dictionary of computer vision and image processing*], John Wiley & Sons, Chichester, Angleterre (2005).

[15] ITU-R, “Methodology for the subjective assessment of the quality of television pictures, recommendation,” (2013).

[16] Chandler, R., “Seven challenges in image quality assessment: past, present and future research,” *ISRN Signal Processing* **2013**, 1–53 (2013).

[17] Bovik, A., “Automatic prediction of perceptual image and video quality,” *Proceedings of the IEEE* **101**, 2008–2024 (September 2013).

[18] Wang, Z. and Bovik, A., “Mean squared error: love it or leave it: a new look at signal fidelity measures,” *IEEE Signal Processing Magazine* , 98–117 (January 2009).

[19] MacAdam, D., “Visual sensitivities to color differences in daylight,” *Journal of Optics* **32**(5), 247–273 (1942).

[20] Wang, Z. and Bovik, A., “A universal image quality index,” *IEEE Signal Processing Letters* **9-3**, 81–84 (2002).

[21] Sheikh, H. and Bovik, A., “Image information and visual quality,” *IEEE transactions on Image Processing* **15-2**, 430–444 (2006).

[22] Shnayderman, A., Gusev, A., and Eskicioglu, A., “An SVD-based grayscale image quality measure for local and global assessment,” *IEEE transactions on Image Processing* **15-2**, 422–430 (2006).

[23] Chandler, D. and Hemami, S., “VSNR: a wavelet based visual signal-to-noise ratio for natural images,” *IEEE transactions on Image Processing* **16**(9), 2284–2298 (2007).

[24] Charrier, C., Lezoray, O., and Lebrun, G., “Machine learning to design full-reference image quality assessment algorithm,” *Signal Processing: Image Communications* **27-3**, 209–219 (2012).

[25] Ferzli, R. and Karam, L., “A no-reference objective image sharpness metric based on the notion of Just Noticeable Blur (JNB),” *IEEE transactions on Image Processing* **18-4**, 717–728 (2009).

[26] Ladjal, S., *Flou et quantification dans les images numériques*, PhD thesis, Télécom ParisTech, France (2005).

[27] Zang, W. and Simoncelli, E., “Local phase coherence and the perception of blur,” in [*Proceedings of NIPS’03*], **16**, 786–792 (2003).

[28] Blanchet, G., Moisan, L., and Rougé, B., “Measuring the global phase coherence of an image,” in [*Proc of The IEEE International Conference on Image Processing*], 1176–1179 (2008).

[29] Blanchet, G. and Moisan, L., “An explicit sharpness index related to global phase coherence,” in [*Proc of the IIEEE nternational Conference on Qcoustics, Speech and Signal Processing*], 1065–1068 (2012).

[30] Leclaire, A. and Moisan, L., “No-reference image quality assessment and blind deblurring with sharpness metricsexploiting Fourier phase indformation,” *Journal of Mathematical Imaging and Vision* (2015).

[31] Simoncelli, E. and Olshausen, B., “Natural image statistics and neural representation,” *Annual review of neuroscience* **24 (1)**, 1193–1216 (2001).
[32] Torralba, A. and Oliva, A., “Statistics of natural image categories,” *Network computation in neural systems* **14 (3)**, 391–412 (2003).
[33] Mittal, A., Moorthy, A., and Bovik, A., “No-reference image quality assessment in the spatial domain,” *IEEE transactions on Image Processing* **21 (12)**, 4695–4708 (2012).
[34] Mittal, A., Soundararajan, R., and Bovik, A., “Making a ”complete blind” image quality analyzer,” *IEEE Signal Processing Letters* **20 (3)**, 209–212 (2013).
[35] Moorthy, A. and Bovik, A., “Blind image quality assessment: From natural scene statistics to perceptual quality,” *IEEE transactions on Image Processing* **20 (12)**, 3350–3364 (2011).
[36] Toraldo di Francia, G., “Resolving power and information,” *Journal Optical Society of America* **45**, 497–501 (juillet 1955).
[37] Frieden, B., “How well can a lens system transmit entropy?,” *Journal Optical Society of America* **58**, 1105–1112 (1968).
[38] Toraldo di Francia, G., “Degrees of freedom of an image,” *Journal Optical Society of America* **59**, 799–804 (juillet 1969).
[39] ISO, “Norme ISO 17321 - Graphics technology and photography – Colour characterisation of digital still cameras (dscs). Part 1 : Stimuli, metrology and test procedures.” norme ISO (2006).
[40] Middleton, D., [*An introduction to statistical communications*], McGraw Hill, Columbus, Etats Unis (1960).
[41] Stone, J., [*Information Theory: A Tutorial Introduction*], Sebtel Press (2015).
[42] Maitre, H., [*Du photon au pixel : l’appareil photographique numérique*], ISTE Editions, Londres, Royaume Uni (2015).
[43] Cao, F., Guichard, F., Hornung, H., and Masson, L., “Sensor information capacity and spectral sensitivity,” in [*Electronic Imaging 2009*], **EI 109**, IS&T SPIE (Jan 2009).
[44] Cao, F., Guichard, F., and Hornung, H., “Information capacity: a measure of potential image quality of a digital camera,” in [*Proceedings SPIE Vol 7537 - Digital Potography VI*], Imai, F., Sampat, N., and Xiao, F., eds., **7537**, 75370F–1, SPIE-IS&T Electronic Imaging (2010).
[45] Tatarkiewicz, W., [*History of aesthetics*], Mouton ed. (1970).
[46] Gombrich, E., [*Art and Illusion: a study in the psychology of pictorial representations.*], Princeton University Press, Wahington D.C. (1960).
[47] Kemp, M., [*The Science of Art, Optical themes in western art from Brunelleschi to Seurat*], Yale university Press (1990).
[48] Bouleau, C., [*The Painter’s Secret Geometry: A Study of Composition in Art*], Allegro editions (2014).
[49] Birkhoff, G., [*Aesthetic Measure*], Harvard University Press (1933).
[50] Arnheim, R., [*Art and Visual Perception: A Psychology of the Creative Eye*], University of California Press, 50th Anniversary Edition, (revised version from the 1954 book) ed. (2010).
[51] Berlyne, D., “Novelty, complexity and hedonic values,” *Perception and Psychophysics* **8**, 279–290 (1970).
[52] Reber, R., Schwarz, N., and Winkielman, P., “Processing fluency and aesthetic pleasure: Is beauty in the perceiver’s processing experience?,” *Personality and Social Psychology Review* **8**, 364–382 (2004).
[53] Zeki, S., [*Inner vision: An exploration of art and the brain*], Oxford University Press (1999).
[54] Livingstone, M., [*Vision and Art: the biology of Seeing*], Harry N. Abrams. Inc., Publishers. (2002).
[55] Hsu, L., *Le visible et l’expression. Etude sur la relation intersubjective entre perception visuelle,sentiment esthétique et forme picturale*, PhD thesis, Ecole des Hautes Etudes en Sciences Sociales, Paris (2009). HAL id: tel00401739.
[56] Delahaye, J., “La beauté mise en formules,” *Pour la Science* **45**, 78–83 (2015).
[57] Desolneux, A., L., M., and Morel, J., [*From Gestalt Theory to Image Analysis: A Probabilistic Approach,*], vol. 34, Springer-Verlag, Interdisciplinary Applied Mathematics (2008).
[58] Musé, P., Sur, F., Cao, F., Gousseau, Y., and Morel, J., [*Shape recognition based on an a contrario methodology. In Statistics and analysis of shapes*], In Statistics and analysis of shapes, Birkhauser, (2006).

[59] Moles, A. A., "Théorie de l'information et perception esthétique," *Revue Philosophique de la France et de l'Etranger* , 233–242 (1957).

[60] Bense, M., [*Einfürung in die informationstheoretische Ästhetik,. Grundlegung und Andwendung in der Texttheorie*], Rowoldt Taschenbuch Verlag (1969).

[61] Chaitin, G., [*Algorithmic Information Theory*], Cambridge University Press (1987).

[62] Zenil, H., Delahaye, J., and Caucherel, C., "Image characterization and classification by physical complexity," *Complexity* **17**(3), 26–42 (2012).

[63] Rigau, J., Feixas, M., and Sbert, M., "Information aesthetic measures," *IEEE Computer Graphics and Applications* **2**, 24–34 (2008).

[64] Zurek, W., "Algoritmic randomness and physical entropy," *Physical review A* **40**(8), 4731–4751 (1989).