

The Role of a Virtual Observer in an Automatic, Image Processing System

Bror O. Hultgren
Polaroid Corporation, Image Science Laboratory
Cambridge, MA

Abstract

IQA, Image Quality Assured, is a unified image processing platform designed to enable the automatic enhancement of images from a wide variety of image sources to be displayed on a wide variety of hard and soft copy formats. IQA is designed to replace the menu driven image processing platform to enable the choice and degree of image processing steps to achieve visually optimal images. An essential feature of IQA is the use of an virtual observer as part of the optimization of the image processing chain to produce visually optimal images. This virtual observer represents a codification of extensive psychovisual testing and is implemented in conjunction with a model of the image transformations introduced by the imaging chain and an optimization algorithm to form an intelligent control system for an open image processing architecture that produces psychovisually optimal images.

Imaging System Optimization

The design of any complex imaging system is a study in optimization. A complex imaging system consists of hardware and software to capture an image, hardware and software to transform the image and hardware, software and media to display the image. All of the elements are capable of transforming the content of the image in conscious and unconscious ways. The design of the system is a process in making choices of the transformations that the elements of the imaging system impose upon the image.

Complex imaging systems can be

1. optical and digital devices.
2. hybrid combinations of optical chemical and digital devices.
3. dedicated pairs of acquisition and display devices.
4. open systems of user chosen acquisition and display devices.

For every system the optimization process is different in detail but is directed toward a common goal.

System Design in Dedicated Device Pairs.

If all of the system elements are fixed, the designer can attempt to specify algorithms with fixed parameters that correct for the distortions introduced into the image by the elements of the imaging chain, thereby producing a faithful rendition of the acquired image. For example, a sharpening kernel may be chosen that corrects the blur introduced by the optical elements of the image capture device AND the blur of the final display device. In effect the designer is minimizing the mean square pixel error in the processing chain. In practice, image distortions cannot be completely eliminated. In the example of removing the system blur the inverse filter may not be able to be specified exactly over the entire spatial frequency bandwidth of the system. Further distortions in the imaging chain, such as the noise or granularity, can be introduced in the display device that cannot be corrected. The designer is then faced with the question of determining what distortions and what degree of distortions are acceptable.

As a consequence the system designer is forced to rely on market tests of differing system realizations to determine the customer's acceptance of the inevitable distortions in the final displayed images produced.

System Design in an Open System

Many real systems are not dedicated pairs of devices; the imaging customer may have several acquisition devices-e.g. an electronic still camera, a flat bed scanner for reflection media and a transmission scanner for digitizing color negatives or transparencies- and a number of alternative display rendering devices—e.g. a color monitor for soft display or color printers for hard copy display, each of which may use multiple media. Each combination of devices will introduce its own unique distortions of the image. Thus the imaging system optimization must adapt to the customer's choice of devices.

A number of image manipulation programs exist that provide the customer with an open framework into which he can import images, manipulate them through filters and export them to a display device. The images can be imported either directly through the acquisition device driver or indirectly by opening a previously acquired image. Likewise the images can be exported directly through the display device driver or indirectly by saving the image to a file and then displaying the image as a separate operation.

The manipulation of the image is typically affected through the use of a number of filters—e.g. color matching transformations, lightness and darkness offsets, blur/sharpening transformations, despeckle transformations and image resizing. The customer typically will try a number of the filters applying them in various degrees until a pleasing image is presented. If the color printer imposes additional color transformations, the customer may find the final determination of the choice of filters must be made from a large number of iterative prints produced on the display device. More importantly the filters may interact with one another—e.g. sharpening will increase the apparent grain in the image while the despeckle filter will blur the image. Similarly, increased contrast will increase the apparent sharpness as well as the apparent granularity. These interactions will increase the number of iterative attempts the user will attempt before finding a pleasing result. For a fixed pair of devices the customer will probably determine a fixed series of operations that will yield pleasing results, in effect caching the operation set as an optimization of that imaging chain. However the solution may be neither optimal or unique, being suitable only for that particular user.

Task Oriented Goal

In the above situations the system is optimized with a particular goal: either maximizing the fidelity of the displayed image to the original, or maximizing the customers preference of displayed image. In most cases the observer or customer is an essential element in the design optimization process. Thus we would like to determine how the observer responds to the task of image preference as a means to building a predictive model of the observer.

Psychovisual Description of the Observer

The observer is considered to be representative of a chosen population. We are interested in quantifying the observer's preference of images. Determination of the patterns of observer image preference is effectively a problem of scaling or 'measuring' the psychological space of the observer: The characterization of this psychological space is well established in the literature [1] and is routinely performed by a number of businesses not limited to the imaging community. The process yields an operational definition of the standard observer: Within a well defined population the standard observer is defined by the population response to a well defined task specified by well defined range of stimuli. Within these stipulations it has been routinely demonstrated that various scalings of the standard observer's psychological space are related one to another by well defined transformations [2].

We have chosen the standard observer to represent the response of the average photographic customer to home imaging of the type that one would put in the family scrapbook. There are other populations and tasks that can be characterized, e.g. insurance adjusters, and radiological and aerial reconnaissance interpreters, whose responses will differ.

Observer Population	Task	Stimuli Range
Amateur Photographer	Assessment of Image Quality for Home Photo Album	4X5 Contact ↓ ↓ Poor Exposure Blurred Grainy

Some people choose to define the task to be the determination of the 'naturalness' of the image. Alternatively the task could be redefined to represent some particular attribute of an image such as sharpness, colorfulness, or graininess, but we have found it convenient to focus on the all-in psychological response to an image and map the multi-dimensional aspects of the integrated response by testing the observer's response to series of images that vary in some single attribute.

A number of techniques exist for 'measuring' the psychological response of the observer: Ratio Scaling, Forced Pair Comparisons, Acceptance Scaling, and Category Scaling[1]. We have chosen to use Category scaling to measure the observer's psychovisual response. We have shown that well defined transformations exist between all of the aforementioned techniques [2]. Further we have found that the scale provides an easily communicated metric when explaining results to project management and product development teams. Category scaling rests upon determining a word scale that represents uniformly spaced psychological responses to stimuli. Based upon our own work and that of Zwick [3], we have found that the following scale represents a sufficiently delimited, uniform scale that is easily understood by the test population:

Excellent
Very Good
Good
Fair
Poor
Very Poor

Additional categories merely serve to complicate the observer's response.

The observer is presented with the full range of images and queried for acceptance; the purpose of the pre-testing is to acquaint the observer with the range of stimuli. Further, we routinely provide anchor images to establish the range of stimuli. These images are drawn from high quality 35 mm prints or contact prints from 4X5 color negatives to establish the high quality anchor. The low quality anchor images are chosen from a fixed set of badly exposed, blurred and grainy images. We have determined that changes in the anchor images results in a simple linear transformation of the category scale [i.e. the scale is an interval scale];

$$Q' = a * Q + b \qquad 1.$$

Within the linear transformation established by the anchor images, we have found that the resulting Image Quality [Category] Scale has remained fixed over the 12 years of testing.

Objective Model of the Observer

The Category Scale defined above, represents the psychovisual response of an observer to images produced by an imaging system. It is thus very useful for testing a fully realized imaging system or a simulation of an imaging system in development. Such a scale does NOT, by itself, provide the predictive power to permit determination of design optimization without building the system. For system quality prediction, one must develop a model of the psychovisual response of the observer based upon OBJECTIVE attributes of an image.

Based upon extensive work in the literature[4] and observer interviews, we have established a sufficient set of objective correlates of psychovisual Image Quality:

- **Color and Tone Reproduction Errors:** color balance, contrast, under and over exposure. These errors are quantified by determining the magnitude of these errors in a visually uniform color space ΔE^* [2].
- **Spatial Errors:** sharpness errors [quantified in terms of SQF] [5,6]; graininess [quantified in terms of granularity G] [7].
- **Artifacts:** streaks, compression artifacts.

The observer's Image Quality metric is then constructed from these objective image quality correlates:

$$Q = f\left[Q_s(SQF), Q_g(G)\right] * g(\Delta E^*) \qquad 2.$$

where

$$Q_s = \alpha_s * SQF + \beta_s$$

$$Q_g = \alpha_g * \log(G) + \beta_g$$

The resulting combined metrics have been shown to yield a robust, predictive metric for perceived image quality based upon objective measures of the system performance.

Design of an OPEN Real Time System Optimization Algorithm

IQA is designed to be an integrated image processing system capable of importing images from a wide variety of devices and displaying the processed images on a wide variety of monitors and printers. IQA is designed to function in an OPEN system environment. An IQA open environment is construed to be not only an environment where devices can be specified by the user, subject to the constraint that they possess IQA compliant profiles, the image processing chain is also user chosen from a set of IQA supported image processing algorithms. IQA is further designed to produce visually optimal images from arbitrary pairs of supported devices.

As shown in Figure 1, IQA is designed with an architecture [8] that consists of:

1. An IMAGE PROCESSING CHAIN of image processing elements $\{A_i\}$ defined by a user chosen instruction set. This instruction set specifies the image processing filters from a set of algorithms including exposure correction, highlight and shadow correction, sharpening, wiener noise reduction, and image resize. Several of these image processing algorithms require parametric input.
2. An INTELLIGENT CONTROL SYSTEM that determines parameters for the image processing algorithms from a parallel processing chain that models the predicted image system performance. This in turn consists of :
   - An ALGEBRAIC SYSTEM MODEL consisting of a chain of algebraic transformations $\{\alpha_i\}$ that is in one-one correspondence with the image processing chain $\{A_i\}$.
   - IQA compliant PROFILES. These profiles provide a description of the transfer properties of both the source and display devices. These descriptions function in the system model as do the algebraic models of the image processing elements.
   - A VIRTUAL OBSERVER, codified into the control system, that functions to provide an assessment of the perceived quality produced by the imagining chain for a given set of parameters. IQA has codified the observer, as described above, to return a value of the system image quality as a function of system color and spatial domain errors.
   - An OPTIMIZATION ALGORITHM that searches for a set of image processing parameters that will yield maximal values of image quality as calculated by the virtual observer.

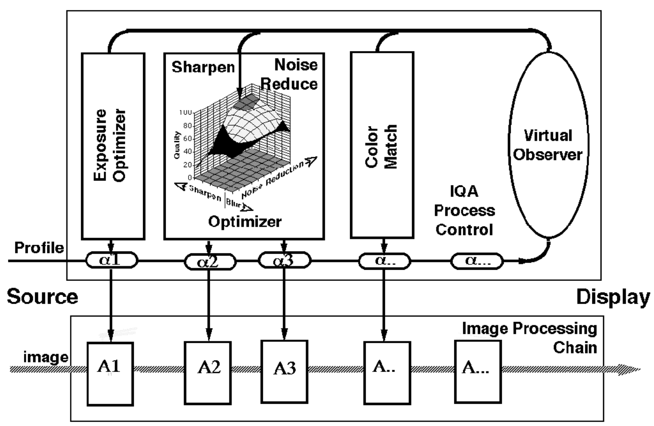

Figure 1. IQA architecture: A_i represent Image processing filters in the image processing chain; α_i represent algebraic models of the Image processing filters in the image processing chain.

ALGEBRAIC SYSTEM MODEL: The system is modeled by the chain of algebraic representations of each filter in the image processing chain, including the source and display devices. The algebraic representation of each filter is a characterization of the transfer function **K(u,v)** associated with the filter, see Figure 2. [Note: for some elements the transfer function for noise- $\mathbf{K_n(u,v)}$ - may differ from the transfer function for signal- $\mathbf{K_s(u,v)}$.] The source and display device transfer functions are given by their respective color profiles and in the spatial domain by the device MTF and Wiener Noise Power Spectra[NPS]. By storing the device MTF and NPS as registered private tags in an ICC compliant profile the IQA control system can determine the imaging system response for an open system.

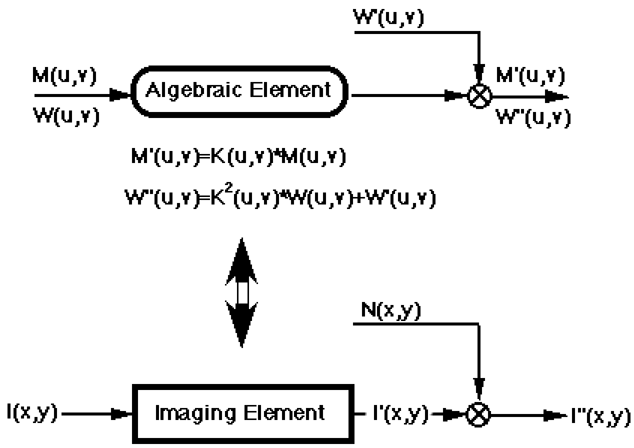

Figure 2. Linear stationary model of Imaging element as implemented in IQA Process Control .

The algebraic model shown in Figure 2 represents linear stationary systems, which describe many image processing elements. For non linear and non stationary image processing elements approximate representations are found that satisfy the linear stationary criteria of the model. Because many of the image processing filters are specified by parameters, the calculation of the system response is dependent upon the choice of those parameters.

$$SQF = \frac{\int E(f) * M(f) * {}^{df}\!/_{f}}{\int E(f) * {}^{df}\!/_{f}}$$

$$G = \log\left[\frac{\int E^2(f) * W(f) * {}^{df}\!/_{f^2}}{\int E^2(f) * {}^{df}\!/_{f^2}}\right] \qquad 3.$$

VIRTUAL OBSERVER: The essential feature of the control system is the inclusion of a VIRTUAL OBSERVER, codified by the function Q = f(Q_S(SQF),Q_g(G)) *g(ΔE*) given in eq. 1., capable of predicting the perceived image quality for any set of processing parameters. The quantities SQF[6] and Granularity G are determined from the system MTF and Wiener Noise Power Spectrum. SQF and log[G] are measures of objective image properties that provide visually uniform measures of the psychovisual attributes of sharpness and graininess quality. Both of these measures are measures based upon the contrast sensitivity function of the human visual system E(f)[6,8].

The function f(Q_S(SQF),Q_g(G)) is a Minkowski type function which serves as a masking function limiting the perceived quality at high sharpness and low granularity

$$f(Q_s, Q_g) = \left[(Q_s)^{-2.5} + (Q_g)^{-2.5}\right]^{-0.4} \qquad 4.$$

The function(ΔE*) which represents the degradation in quality from color and tone reproduction errors is given by:

$$g(\Delta E^*) = \exp\left(-\left({}^{\Delta E^*}\!/_{\Delta_0}\right)^2\right) \qquad 5.$$

Based upon the predictions of the system model, the control system determines an optimal set of processing parameters by employing an OPTIMIZATION of the virtual observer's perceived image quality. It should be noted that while the optimization for the spatial attributes of the image is determined without reference to the image the corrections for exposure and color balance are determined in part from knowledge of the image content.

Conclusion

By utilizing the concept of virtual observer we have been able to implement an intelligent control system for an open image processing architecture that produces psychovisually optimal images.

References

1. C.J. Bartleson and F. Grum, *Optical Radiation Measurements*, Vol. **5**, Visual Measurements, New York, Academic Press (1984).
2. J.Bowman, J. Bullitt F.R. Cottrell, and B.O. Hultgren, *Proceedings of the IS&T 42nd Annual Conf.,* 465 (1989)
3. D.F. Zwick, *Proc. Symposium on Photographic and Electronic Imaging*, Cambridge, 46 (1984).
4. C. J. Bartleson, *J. Photogr. Sci.*, **30**, 33 (1982).
5. E M. Granger and K.N. Cupery, *Photogr. Sci. & Eng*., **3**, 53 (1972)
6. B. O. Hultgren, *Proc. S.P.I.E.*, **1249**,12 (1990)
7. C. J. Bartleson *J. Photogr. Sci*., **33**, 117 (1985).
8. F.R. Cottrell and B.O. Hultgren, *System and Method for Automatically Processing Image Data to Provide Images of Optimal Perceptual Quality*, U.S. Patent 5,694,484