Challenges in International Standardization of Image Quality Evaluation

Toshihiko Inagaki
Research & Development Center, Document Products Company, Fuji Xerox Co. Ltd.
Nakai-machi, Ashigarakami-gun, Kanagawa, Japan

Abstract

The international standardization activities of an image quality evaluation were described from both sides of subjective evaluation and objective evaluation. The standardization activities in ITU-R are described as an example of the subjective evaluation method and SCID (Standard Color Image Data) of ISO/TC130 as an example of a standard image. The standardization activities of the density and the color measurement method in ISO/TC42 and TC130, and the image quality attribute measurement in ISO/IEC JTC1/SC28 are described.

Introduction

Image quality evaluation is classified into subjective evaluation and objective evaluation. Subjective evaluation is desirable because image quality is usually evaluated by subjective judgment of the persons who use the image. However, the subjective evaluation fluctuates by the viewing condition, the past experience of observer, the type of image, the fatigue of observer or the purpose to use the image. On the other hand, the objective evaluation by physical measurements has good reproducibility, and decision or judgment is easier because of the numerical expression. And many researches to find a physical descriptor that indicates high correlation with subjective evaluation have been performed.

Since the purpose of an image quality evaluation is in an image quality improvement or image quality design of an imaging system, the correlation study between the physical characteristic of a system and a subjective evaluation has been performed widely. The image quality evaluation technology is to construct a objective evaluation method which correlate highly with the subjective evaluation which is a result of advanced information processing of a human being as shown in Figure 1.

Image quality evaluation technology has been advanced with image quality improvement activity in various field of imaging technology such as graphic arts, silver halide photography, television and electrophotography.

Although many image quality evaluation technologies that correlate physical characteristics and subjective evaluation has been developed in the various field, the single scale is not in the state where it is accepted universally.[1-4]

In the field of a subjective evaluation, since an evaluation changes with an evaluated image, observation conditions, subjective evaluation methods, etc., the international standardization activities for specifying common evaluation conditions have mainly been advanced.

In the field of an objective evaluation, the physical measurement method for quantifying an image reproduction system with a common scale has mainly been advanced.

In recent years, standardization of the measurement method of the image quality attribute, which constitutes a subjective evaluation, has been advanced.

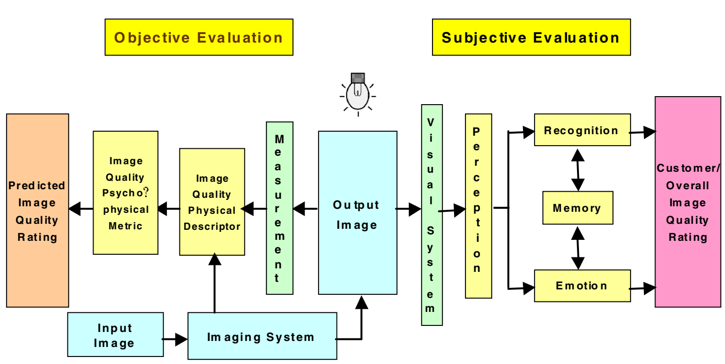


Figure 1. Image quality evaluation system

Some typical standards relevant to the image quality evaluation established by international standardization organizations, such as ISO (International Standardization Organization), ITU (International Telecommunications Union), IEC (International Electrotechnical Commission) and CIE (Commission Internationale de I'Eclariage) are shown in Table 1.

Table 1. Standards Relevant to Image Quality Evaluation

Year	Standard: Colorimetry, Density measurement, Test chart, Evaluation condition, Psychological Scaling, Image Quality measurement
1931	CIE 1931 Standard colorimetric system (XYZ)
1964	CIE 1964 Supplementary colorimetric standard system ($X_{10}Y_{10}Z_{10}$)
1974	ＩＴＵ-R BT.500 Methodology for the subjective assessment of the quality of television pictures
1975	ISO 2846 Set of printing inks for offset printing -- Colorimetric characteristics ISO 3664 Photography -- Illumination conditions for viewing colour transparencies and their reproductions
1976	CIE 1976 L*a*b* color space, L*u*v* color space
1977	ANSI PH2.17 Photography - Density measurements
1982	ISO 6328 Photography – Photographic mat. – Determination of ISO resolving power
1983	ISO 5-4 Photography - Density - Part 4: Geometric conditions for reflection density
1984	ISO 5-1:Photography - Density measurements - Part 1: Terms, symbols and notations
1986	CIE Standard Illuminants for Colorimetry
1988	CIE 2 degree Spectral Luminous Efficiency Function for Photopic Vision
1989	ANSI PH2.30 Photography - Color prints, transparencies, and photomechanical reproductions - Viewing Conditions
1991	ISO/CIE 10526 CIE Standard Colorimetric Illuminants （CIE S 001） ISO/CIE 10527 CIE Standard Colorimetric Observers （CIE S 002） ISO 5-2:1991Photography -- Density measurements - Part 2: Geometric conditions for transmission density
1993	ANSI IT8.7/1 Graphic Technology - Color Transmission Target for Input Scanner Calibration ANSI IT8.7/2 Graphic Technology - Color Reflection Target for Input Scanner Calibration ANSI IT8.7/3 Graphic Technology - Input Data for Characterization of 4-Color process printing
1994	ANSI/CGATS9 Graphic technology -- Graphic arts transmission densitometry measurements -- Terms, equations, image elements and procedures
1995	ANSI/CGATS TR001 Graphic technology -- Color characterization date for Type 1 printing (SWOP) ISO 5-3:1995 Photography - Density measurements - Part 3: Spectral conditions ISO 5-4:1995 Photography - Density measurements - Part 4: Geometric conditions
1996	CIE S003 Spatial Distribution of Daylight - CIE Standard Overcast Sky and Clear Sky ISO 13655 Graphic technology -- Spectral measurement and colorimetric computation for graphic arts images
1997	ISO 12641 Graphic technology - Prepress digital data exchange -- Colour targets for input scanner calibration ISO 12640 Graphic technology - Prepress digital data exchange -- CMYK standard colour image data (CMYK/SCID)
1998	ISO 12645 Graphic technology - Process control -- Certified reference material for opaque area calibration of transmission densitometers ISO 12639 Graphic technology -- Prepress digital data exchange -- Tag image file format for image technology (TIFF/IT)
1999	ISO 10526 CIE Standard Illuminants for Colorimetry（CIE S 005） IEC 61966-2-1: Multimedia Systems and Equipment - Colour Measurement and Management - Part 2-1: Colour Management - Default RGB Colour Space - sRGB
2000	ISO 13656 Graphic technology -- Application of reflection densitometry and colorimetry to process control or evaluation of prints and proofs ISO 14981 Graphic technology -- Process control -- Optical, geometrical and metrological requirements for reflection
2001	ISO 13660 Information Technology -- Office equipment -- Measurement of image quality attributes for hardcopy output -- Binary monochrome text and graphic images

In this paper the international standardization activities relevant to image quality evaluations, such as a standard image, the subjective evaluation method, physical measurement metod, and an image quality attribute measurement method, are described.

International Standardization related to Subjective Image Quality Evaluation

As an element of a subjective evaluation, a subjective evaluation method, and a standard image for subjective evaluation, a viewing condition, etc. are mentioned. Standardization has been performed in the various field, standard images for a subjective evaluation and viewing conditions. Although many subjective evaluation methods called as psychological scaling has been developed, standardization activities are not active other than ITU-R (radio-communications section). In the following, the standardization activities in ITU-R are described as an example of subjective evaluation method and SCID (Standard Color Image Data) of ISO/TC130 (Graphic Technology) as an example of standard image.

Standard Image for Subjective Evaluation[5-8]

In ISO/TC130, the high precision color digital image set by cyan, magenta, yellow and black (CMYK) dot values (8 bits per color) for printing called CMYK/SCID was established as ISO 12640 in 1997,[5] and is widely used as a standard image of color printing. The CMYK/SCID was prepared by IPTS (image processing technology standard) committee in Japanese Standards Association, and was established as JIS X 9201 in 1995.[6]

CMYK/SCID consists of eight natural images and ten synthetic images that contain five resolution charts and five color charts, and consists of the primary and the alternate set with which the dot value range and resolution differs. The output image size of natural images is intended as 128mm by 160mm. As for the primary set, the size of a natural image is 2560 x 2048 pixels, and the printing machine whose resolution is 16 pixels/mm is assumed. The data range of 28 to 228 representing dot percent area coverage values of 0% to 100%. As for the alternate set, the size of a natural image is 1920 x 1536 pixels, and the printing machine whose resolution is 12 pixels/mm is assumed. The data range of 0 to 255 representing dot percent area coverage values of 0% to 100%. In order to distinguish both, the character "ISO 400" for primary set and the character "ISO 300" for alternate set are inserted in the upper part of each image.

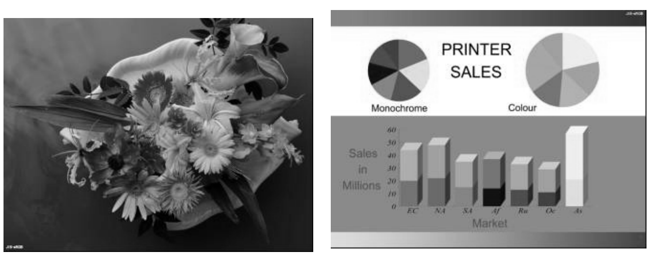
CMYK/SCID is not clear in relationship with chromaticity, because image data is expressed at the rate of CMYK dot area for printing. And, because each pixel is expressed by 8 bits, if an image is processed, precision will fall.

Then, the needs for a new standard image applicable not only to a printing but a monitor display were studied, and new image data set called as XYZ(sRGB)/SCID with higher bits and expressed in the device-independent form was developed. XYZ(sRGB)/SCID was established as JIS X 9204 in 2000,[8] and is proposed also as an international standard ISO12640-2.[7]

XYZ(sRGB)/SCID consists of three sorts of image set which consist of 16 bits XYZ, 8-bit RGB and CIELAB data, and each set consists of five natural images and seven synthetic images (computer graphics, a business graph, a color chart, and a color gradation charts) generated by computers.

For the size of the image for a subjective evaluation, a natural image is 4096 x 3072 pixels (about 12 million pixels), computer graphics, and a business graph are 1536 x 2048 pixels. Four computer graphics and the business graph are created for the color monitor output use. Color chart (1332 x 2736) and color gradation (2608 x 4256) is for a color reproducibility evaluation.

As SCID image family, CIELAB/SCID which uses a device-independent color signal (CIELAB) has been developing as a next version of CMYK/SCID which uses the signal for printing and XYZ(sRGB)/SCID which uses the signal for monitors.



(Exhibition JIS X 9204 XYZ/SCID)

Figure 2. Examples of XYZ/SCID images

A joint study group was organized in 1992, when CMYK/SCID development activities was in the final stage, to investigate a high precision color image data for evaluating the performance of a imaging system in which the data expressed in the form of device-independent to exchange image data freely between different equipments by the Institute of Image Electronics Engineers of Japan (IIEEJ), the Japanese Society of Printing Science and Technology, the Institute of Electronics, Information and Communication Engineers, the Society of Photographic Science and Technology of Japan, and the Imaging Society of Japan. The joint study group published SHIPP (Standard HIgh Precision Picture data) from IIEEJ. The SHIPP is origin of XYZ(sRGB)/SCID. From the argument on the joint study group, the final version of SCID will be the image data expressed with spectral distribution in each pixel.

Subjective Evaluation Method[10-11]

In ITU-R, the subjective evaluation method of a television image quality is standardized as recommendation ITU-R BT.500. Based on BT.500, the evaluation methods, such as conventional television systems (BT.1128), high-definition television (BT.710), stereoscopic television pictures (BT. 1438) are recommended in ITU-R, and digital cable television (J.140) and multimedia (P.911) are recommended in ITU-T (telecommunication section).

The subjective evaluation method is divided into two classes, the quality evaluation for establishing the performance of the system under optimum conditions, and the impairment evaluation for establishing the ability of a system to retain quality under non-optimum conditions, in ITU-R BT.500. The single stimulus method, the stimulus comparison method, the double stimulus method, and the continuous evaluation method are specified as the subjective evaluation method to each class. And, the viewing condition, number of observers, and analysis method of evaluation result etc. are specified. Because each evaluation method has the merit and demerit, to select the method suitable for the purpose of an evaluation is needed.

The single stimulus method has three kinds, the categorical judgment method, the non-categorical judgment method, and the performance method. The categorical judgment method has two kinds of five-grade category scales shown in Table 2, and non-categorical judgment method has two kinds, a continuous sca1ing and a numerical scaling. In stimulus comparison method, a comparison scale by seven categories is recommended. The performance method is the method of measuring the speed and correctness in the case of reading a document, when searching for the target information.

Table 2. ITU-R Quality and Impairment Scale

Rating	Quality Scale	Impairment Scale
5	Excellent	Imperceptib1e
4	Good	Perceptible but not annoying
3	Fair	S1igthly annoying
2	Poor	Annoying
1	Bad	Very annoying

In the double-stimulus method where a standard image and an evaluation image are shown, the double-stimulus continuous quality scale method and the double-stimulus impairment scale method are specified.

The observers are specified that they must be the non-experts who is inexperienced in an image quality evaluation, and has not been engaged in the work in connection with an image quality. At least 15 or more observers are needed.

Since ITU-R BT.500 were established in 1974, it has been revised 10 times and improvement has been continued. Ratio-scaling method has been also studied at ITU-R BT.1082.

Because ITU-R BT.500 are a subjective evaluation method for a display image, the subjective evaluation method for hard copy or the method of evaluating hardcopy and a display image in common is expected as future standardization.

International Standardization of Objective Image Quality Evaluation

As an element of objective evaluation, a measurement method of physical measure such as density or colorimetric value, and a standard image for objective evaluation, a definition of image quality attribute and the method of measurement, etc. are mentioned. In the following, the standardization activities of the density and the color measurement method in ISO/TC42 (Photography) and TC130, and the image quality attribute measurement in ISO/IEC JTC1/SC28 (Office Equipment) are described.

Image Quality Metrics Measurement

In ISO 5, the definition of an optical density, geometric conditions, the spectral characteristic, etc. are specified, and it has been used widely. On the other hand, in CIE, the color measurement method based on human color sensitivities has been developed, and the basic physical measure has been moving to colorimetric value from optical density.

Although only the spectral sensitivity products of a measurement system determine a density, a colorimetric value cannot be determined without the illuminant and the viewing angle (2 degrees or 10 degrees) of observing an image. It is supposed that a density is the physical measure for density control of coloring material rather than the measure for visual evaluation. ISO 5 has been used as density measurement of photography, revision of ISO 5 has been advancing by Joint Working Group of TC42 and TC130 from a viewpoint of density measurement for the process control of printing.

And, geometric conditions, and illumination conditions, backing material, etc. for reflection density measurement specified by ISO 5 are used by ISO 13655 as color measurement conditions for graphic arts printing, and make smooth the transition to colorimetric value from optical density.

Image Quality Attribute Measurement[12]

In ISO/IEC JTC1 SC28, the device-independent image quality attribute of monochrome hardcopy is systematized, and the international standardization activities that can measure the image quality attribute automatically with easy measurement equipment have been advanced from 1992. As the first step, the image quality attribute measurement method of binary monochrome text and graphic images was standardized as ISO 13660 in 2001. The definition of eight items of the character and line image quality attribute and six items of large area image quality attribute are specified in ISO 13660 as shown in Table 3.

In ISO 13660, the method of measuring the characteristics of the image itself directly is specified without specifying a test chart or a reference image. The measurement system needs to be calibrated and test objects and the target values are specified to prove that the calibration of the measurement system is acceptably good. However, the targets values of large area image attributes are not specified in ISO 13660 and a system compliance test method for large area image attribute has been developing to be added to ISO 13660. Test object for system compliance test proposed at the SC28 plenary meeting held in 2001 is shown in Fig. 3.

Measurement in ISO 13660 is to be performed by the equipment that has 8 bits per pixel (256 gradation) and the dynamic range of 0.1 to 1.5 in density on the basis of 600dpi.

Table 3. Image Quality Attributes Specified in ISO 13660

	Large area attributes	Character and line attributes
1	darkness, large area	blurriness
2	background haze	raggedness
3	graininess	line width
4	mottle	darkness, character
5	extraneous marks, background	contrast
6	voids	fill
7		extraneous marks, character field
8		background haze, character field

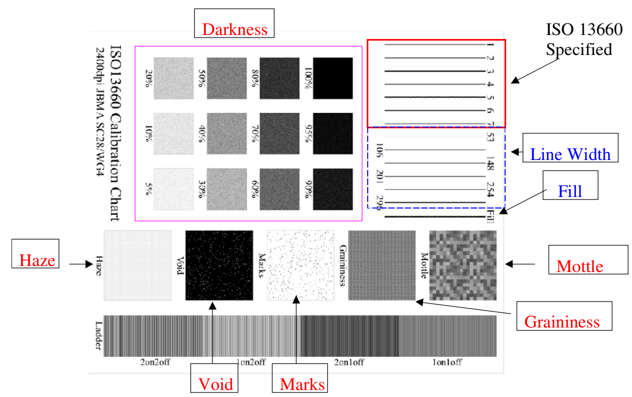


Figure 3. Test images for system compliance

For example, in a graininess and mottle measurement, a region of interest (ROI) is at least 300 x 300 pixels (12.7 x 12.7 = 161mm^2) is divided into 100 (10 x 10 tiles) evenly-spaced, non-overlapping square tiles with area at least 1.61 mm^2.

Mottle is defined as density fluctuation at a spatial frequency less than 0.4 cycles/mm. The measure of a mottle is the standard deviation of the average density of each tile in ROI.

Graininess is defined as density fluctuation at a spatial frequency greater than 0.4 cycles/mm. The measure of graininess is the root mean square of standard deviation σ_i of all the tiles in ROI.

$$graininess = \sqrt{\left(\sum_{i=1}^{n} \sigma_i^2\right) \Big/ n} \quad (1)$$

The algorithm proposed by R. Shaw and R. Dooley is shown below as an example of the psychophysical measure of graininess.[13]

$$graininess = u_0 \, e^{-1.8D} \sum_{j=1}^{N} \sqrt{WS(ju_0)} \quad VTF(ju_0) \qquad (2)$$

where u_0 is a fundamental spatial frequency, WS is a Wiener spectrum, and VTF is the spatial frequency characteristics of the visual system

Although Eq. (1) says that physical measure can express the psychological quantity, Eq. (2) says that physical measure needs to correct it with visual characteristics. All the measure of ISO 13660 makes the correction by visual characteristics unnecessary. And the question whether the measured value according to ISO 13660 shows the value that suited the subjective evaluation really has come.

And, it is under verification whether there is sufficient relationship between the value of each image quality attribute measured according to ISO 13660 and the rating by subjective evaluation. Revision of ISO13660 and reconsideration of a large area color image quality attribute measurement proposal will be achieved based on these verification results.

Discussions on Findings

The international standardization activities of image quality evaluation were reviewed from both sides of subjective evaluation and objective evaluation. In the review, it was assumed that international standardization of the image quality evaluation method independent of the equipment that enables the image quality management that unifies the image reproduction between different imaging equipment by the popularization of digital imaging equipment was asked.

From a viewpoint of image quality management, SCID is considered to be an image set for a color reproduction evaluation. As image for subjective evaluation, not only natural image but also character/text image and computer graphic image should be considered.

The standard subjective evaluation method is indefinite, although paired comparison, the category method, etc. are used in interval scale construction and the magnitude estimating method is used in ratio scale construction. It cannot be said that the standardization activities of the subjective evaluation method are active. In the subjective evaluation of hard copy, because it is not necessary to take into consideration the number of the order of presenting a stimulus, and the stimuli presented simultaneously etc. unlike a display, since the number of stimuli can be increased easily, evaluation conditions and the psychological scaling method are adequately arranged in efficiency and precision. The psychological scaling method that the ordinary engineers in the world can use in common needs to be studied before the discussion on interval scale or ratio scale. And the viewing condition that can compare the displayed image (soft copy) with hard copy simultaneously for a subjective evaluation should be studied.

Image quality measurement technology has been advanced with image quality improvement activity in various field of imaging technology, and it is the technology depending on the company who led each imaging technology or its field. The versatile measurement method should be standardized by industry-university cooperation.

From a viewpoint of image quality measurement, microscopic image structure measurement is needed and establishment of the standard method of micro-colorimetry is expected to be a subject.

Conclusion

The international standardization activities of an image quality evaluation were reviewed on the assumption that the international standardization of the image quality evaluation method that enables same image reproduction between different imaging equipment would be asked. And psychological scaling method, the micro-colorimetry method, etc. were extracted as a subject of international standardization activity.

References

1. R. Shaw, A Century of Image Quality, Proc. of PICS Conference, 221-224 (1999).
2. Edited by R. Shaw, Selected readings in image evaluation, SPSE (1976)
3. R. E. Jacobson, An Evaluation of Image Quality Metrics, J. Photogr. Sci., 43, 7-16 (1995).
4. Joint Publication Committee of Japanese Photography Society and Imaging Society of Japan, Fine imaging and hardcopy, p.477, Corona Publishing Co., Ltd. (1999)
5. ISO 12640:1997, Graphic technology - Prepress digital data exchange CMYK standard colour image data (CMYK/SCID)
6. JIS X 9201:1995, Standard colour image data (SCID)
7. ISO/WD 12640-2, Graphic technology -- Prepress digital data exchange - XYZ(sRGB) standard colour image data (XYZ(sRGB) /SCID)
8. JIS X 9204:2000, XYZ/sRGB standard colour image data (XYZ/SCID)
9. Rec. ITU-R BT.500-10, Methodology for the subjective assessment of the quality of television pictures
10. J. Kumata, Y. Nishida, K. Watanabe, H. Ikegami, M. Matsuki, F. Ono, The international standardization trend relevant to the image evaluation, The Institute of Image Information and Television Engineers, 54 No.7 (2000)
11. T. Inagaki, The international standardization activities of image quality evaluation, O plus E, 23 No.10 (2001)
12. ISO 13660: 2001, Information Technology - Office Equipment -Measurement of image quality attributes for hardcopy output - Binary monochrome text and graphic images
13. R. Doolley, R. Shaw, Noise perception in electrophotography, J. Appl. Photogr. Eng., 5, pp.190-196 (1979)

Biography

Toshihiko Inagaki is image quality research group manager of Research & Development Center of Document Product Company in Fuji Xerox Co. Ltd. He has been developing image quality evaluation methods and systems for 25 years in Fuji Xerox. He has been in charge of extraction of psychological attribute in image quality evaluation, development of the objective measurement method of psychological attribute, development of image quality prediction model, development of image quality meter, and so on. He is project editor of ISO 13660 addendum on system compliance test and ISO 13660-2 measurement of image quality attributes for hardcopy output – large area color images. And he is chief examiner of Working Group 1 of technology committee in the Imaging Society of Japan since 1990.