INTERNATIONAL STANDARD

ISO 20462-2

First edition
2005-11-01

Photography — Psychophysical experimental methods for estimating image quality —

Part 2: Triplet comparison method

Photographie — Méthodes psychophysiques expérimentales pour estimer la qualité d'image —

Partie 2: Méthode comparative du triplet

Reference number
ISO 20462-2:2005(E)

ISO copyright office
Case postale 56 • CH-1211 Geneva 20
Tel. + 41 22 749 01 11
Fax + 41 22 749 09 47
E-mail copyright@iso.org
Web www.iso.org

Published in Switzerland

Contents

Page

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 2.

The main task of technical committees is to prepare International Standards. Draft International Standards adopted by the technical committees are circulated to the member bodies for voting. Publication as an International Standard requires approval by at least 75 % of the member bodies casting a vote.

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights.

ISO 20462-2 was prepared by Technical Committee ISO/TC 42, *Photography*.

ISO 20462 consists of the following parts, under the general title *Photography — Psychophysical experimental method for estimating image quality*:

— *Part 1: Overview of psychophysical elements*

— *Part 2: Triplet comparison method*

— *Part 3: Quality ruler method*

Introduction

This part of ISO 20462 is necessary to provide a basis for visually assessing photographic image quality in a precise, repeatable and efficient manner. This part of ISO 20462 is needed in order to evaluate various test methods or image processing algorithms that may be used in other international and industry standards. For example, it should be used to perform subjective evaluation of exposure series images from digital cameras as part of the work needed for future revisions of ISO 12232.

The opportunities to create and observe images using different types of hard copy media and soft copy displays have increased significantly with advances in computer-based digital imaging technology. As a result, there is a need to develop requirements for obtaining colour-appearance matches between images produced using various media and display technologies under a variety of viewing conditions. To develop the necessary requirements, organizations, including the CIE and the ICC, are developing methods to compensate for the effect of different viewing conditions, and to map colours optimally across disparate media having different colour gamuts.

Such technical activities are often faced with the need to evaluate proposed methods or algorithms by visual assessment based on psychophysical experiments. K.M. Braun *et al.*[1] examined five viewing techniques for cross-media image comparisons in terms of sensitivity of scaling, and mental and physical stress for the observers. CIE TC1-27 “Specification of Colour Appearance for Reflective Media and Self-Luminous Display Comparisons” proposed guidelines for conducting psychophysical experiments for the evaluation of colorimetric and colour-appearance models[6]. Accordingly, for the design and evaluation of digital imaging systems, it is of great importance to develop a methodology for subjective visual assessment, so that reliable and stable results can be derived with minimum observer stress.

When performing a psychophysical experiment, it is highly desirable to obtain results that are precise and reproducible. In order to derive statistically reliable results, large numbers of observers are required and careful attention should be paid to the experimental setup. Multiple (repeated) assessments are also useful. Observer stress during the visual assessment process can adversely affect the results. The order of image presentation, and the types of questions or questionnaires addressed by the observers, can also affect the results.

Table 1 gives a comparison of three visual assessment techniques commonly used for image quality evaluation. The advantages of the category methods include low stress and high stability, since the observer’s task is to rank each image using typically five or seven categories. However, its scalability within a category is less precise. One of the most common techniques for image quality assessment is the paired comparison method. This method is particularly suited to assessing image quality when precise scalability is required. However, a serious problem with the paired comparison method is that the number of samples to be examined is to be relatively limited. As the number of the samples increases, the number of combinations becomes extensive. This causes excessive observer stress, which can affect the accuracy and repeatability of the results. The third method, commonly known as magnitude scaling, is magnitude estimation. This method is extremely difficult when the psychophysical experiments are conducted using ordinary (non-expert) observers to perform the image quality assessment.

Table 1 — Comparison of typical psychophysical experimental methods

Name of method	Scalability	Stability	Stress
Category	Low	High	Low
Magnitude estimation	Medium	Low	Medium
Paired comparison	High	High	High

G. Johnson *et al.*[3] have proposed "A sharpness rule", where the magnitude of sharpness was analyzed in terms of resolution, contrast, noise and degree of sharpness-enhancement. Likewise, preferred skin colour may be considered not only from the viewpoint of chromaticity, but also with respect to the lightness, background and white point of the display media[4]. These examples show that image quality is not always evaluated by a single attribute, but may vary in combination with multiple attributes. In cases where a psychophysical experiment is designed for a new application, the experimenter may need to vary many attributes simultaneously during the course of the experiment. In these situations, the number of the samples to be examined becomes excessively large, making it difficult to employ the paired comparison technique.

