International Standard 1831

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION•МЕЖДУНАРОДНАЯ ОРГАНИЗАЦИЯ ПО СТАНДАРТИЗАЦИИ•ORGANISATION INTERNATIONALE DE NORMALISATION

Printing specifications for optical character recognition

Spécifications d'impression des caractères pour reconnaissance optique

First edition — 1980-10-15

UDC 681.327.5.048 **Ref. No. ISO 1831-1980 (E)**

Descriptors : data processing, printing, optical characters, character recognition, optical recognition, printing papers, specifications, dimensions, legibility, measurement.

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards institutes (ISO member bodies). The work of developing International Standards is carried out through ISO technical committees. Every member body interested in a subject for which a technical committee has been set up 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.

Draft International Standards adopted by the technical committees are circulated to the member bodies for approval before their acceptance as International Standards by the ISO Council.

International Standard ISO 1831 was developed by Technical Committee ISO/TC 97, *Computers and information processing*, and was circulated to the member bodies in January 1979.

It has been approved by the member bodies of the following countries :

Australia	Ireland	Romania
Belgium	Italy	South Africa, Rep. of
Czechoslovakia	Japan	Spain
Finland	Mexico	Sweden
France	Netherlands	Switzerland
Germany, F. R.	Poland	USSR

The member bodies of the following countries expressed disapproval of the document on technical grounds :

Canada
United Kingdom
USA

This International Standard cancels and replaces ISO Recommendation R 1831-1971, of which it constitutes a technical revision.

Printed in Switzerland

Contents

Printing specifications for optical character recognition

0 Introduction

The purpose of this International Standard is to establish the basis for industry standards for paper and printing to be used in Optical Character Recognition (OCR) systems, particularly for document interchange, and to aid in the implementation and use of such systems.

It provides for the identification and measurement of, and establishes specifications for, the relevant parameters and gives guidance for their use.

0.1 Interpretation of the International Standard

A printing system is defined as a single unit comprising a printing machine, paper and inked ribbon (the latter only if required by the printing process). A printing system which produces printed material for OCR applications is called an OCR printing system.

The values in this International Standard shall apply to OCR printed material regardless of the printing system, font (OCR-A, OCR-B) and the specific application. The dimensional and optical characteristics of the printed image are given for three quality ranges.

Tolerance limits are specified for each parameter. These limits at least shall be achieved, but all parameters are expected to be kept well within them. If some of these parameters subject to variations of a statistical nature deviate from the specified limits, then the number and magnitude of these deviations can be reduced by using special precautions, such as a more accurate choice of the OCR printing system components, more frequent maintenance of the printing machine, a reduction of the printing speed, a shortening of the ribbon life, etc.

If the performance of an optical character recognition system is subject to variations of a statistical nature and if rejections or substitutions of characters within the tolerance limits occur then, again, the number and magnitude of these deviations can be reduced by using special precautions, such as a more frequent maintenance of the recognition system, etc.

0.2 Use of the International Standard

The measurement methods and the values of the parameters given in this standard are intended for use in OCR applications.

As a continuous, complete fulfilment of these values cannot be achieved because of the deviations of a statistical nature to which both printing and recognition systems are liable, some rejection and substitution of characters may occur. The number of rejections and substitutions which are allowed depends on the specific OCR application and shall be agreed upon, in statistical terms, by the user, the supplier(s) of the printing system and the supplier(s) of the recognition system.

In the guarantee of printing systems, the manufacturer of the printing system is given the right to specify the maintenance rate for the printing system and the supplies to be used (for example paper and ribbon).

In the guarantee of the recognition system, the supplier of the recognition system is given the right to specify the environmental conditions (temperature, humidity, illumination, maximum amount of mechanical vibrations and electromagnetic noise, etc.) and to establish the level of maintenance for the reader.

Statistical sampling plans by inspection of attributes can be used to check whether these guarantees are being observed, provided that these plans are coherent with those normally used in quality control.

Once a sampling plan has been agreed upon, the sample size (i.e. the number of characters or documents to be examined) is established by the plan.

To allow the printing system to be checked, the parameters of the printed material to be measured and the measurement methods are given in this International Standard.

When the recognition system is checked, only printed material meeting the specifications given in this International Standard shall be used, or — by agreement — representative samples of current material may be used. In the latter case the rejects must be evaluated according to their compliance with this International Standard.

0.3 Annexes

The annexes are not an integral part of this International Standard but give additional information.

1 Scope and field of application

This International Standard contains the basic definitions, measurement requirements, specifications and recommendations for OCR paper and print.

Three major parameters of a printed document for OCR media are covered. These are :

— the optical properties of the paper to be used;

— the optical and dimensional properties of the ink patterns forming OCR characters;

— the basic requirements related to the position of OCR characters on the paper.

The major factors of each of these areas pertinent to OCR are identified. Definitions of these items are given and bases for measurements are established.

Basic specifications applicable to all OCR materials are imposed and recommendations for the implementation of an OCR system are made.

2 References

ISO 216, *Writing paper and certain classes of printed matter — Trimmed sizes — A and B series.*

ISO 1073/1, *Alphanumeric character sets for optical recognition — Part 1 : Character set OCR-A — Shapes and dimensions of the printed image.*

ISO 1073/2, *Alphanumeric character sets for optical recognition — Part 2 : Character set OCR-B — Shapes and dimensions of the printed image.*

ISO 2469, *Paper, board and pulps — Measurement of diffuse reflectance factor.*

ISO 2471, *Paper and board — Determination of opacity (paper backing) — Diffuse reflectance method.*

CIE Publication 15 (E 1.3.1) 1971 — *Colorimetry — Official recommendation.*

3 Spectral requirements

3.1 General

This clause defines spectral bands of interest for OCR applications.

They shall be defined since character readers operate in specific spectral regions and paper and ink characteristics change with the wavelength considered.

3.2 Spectral bands

In this clause, a set of bands is defined as reference for the paper and printed image specification. Their use and the measuring procedures are specified in the clauses on paper reflectance, paper opacity and PCS measurement.

Table 1

Band	Peak nm	Bandwidth nm, 50 % level
B 425	425 ± 5	50 or less
B 460	460 ± 5	60 or less
B 490	490 ± 5	60 or less
B 530	530 ± 5	60 or less
B 570	570 ± 10	100 or less
B 620	620 ± 10	100 or less
B 680	680 ± 10	120 or less
B 900	900 ± 50	400 or less

The bands B 425 up to B 900 represent the spectral responses required from the complete measuring instrument (light source, filter, detector). These responses shall be smooth curves without secondary peaks and with no major parts of the response curves beyond the specified 50 % points. The energy content of the illumination at wavelengths shorter than 400 nm should not exceed 5 % of that in the particular band under consideration.

4 Paper specifications for OCR

4.1 General

The papers to be used in OCR applications should be white (see annex A), have low gloss, and be of high opacity (see annex A). Factors causing variation in reflectance (such as dirt, uneven formation, watermarks and fluorescent additives) should be avoided.

In particular OCR applications, some mechanical properties of paper (such as stiffness, porosity, tear resistance and smoothness, etc.) may be important. For both optical and mechanical properties, agreement between users and manufacturers of OCR systems on the specific papers to be used is advisable.

4.2 Luminous reflectance factor R_o of paper

Reflectance measurements shall be carried out using a reflectometer as described in ISO 2469, or an instrument calibrated against such a reflectometer.

Reflectance measurements shall be referred to the perfect reflecting diffuser (100 % reflectance). However, in practice