

**Paper Forms Design Optimization For
Electronic Image Management (EIM)**

Technical Report

ANSI/AIIM TR32-1994



AIIM

Association for Information and Image Management

1100 Wayne Avenue, Suite 1100
Silver Spring, Maryland 20910
Telephone 301/587-8202

Technical Report for Information and Image Management —

Paper Forms Design Optimization For Electronic Image Management (EIM)

An ANSI Technical Report
Prepared by the Association for Information and Image Management

Association for Information and Image Management

Abstract

This technical report provides guidelines for the design of printed forms that are used in EIM systems. The primary audience for this document is individuals involved with designing or specifying requirements for printed paper forms.

Contents

Foreword	i
1 Scope and purpose	1
2 References	1
3 Definitions	2
4 Layout and design	3
5 Providing user instructions	9
6 Physical considerations	11
7 Forms testing	12
8 New technology	13

Figures

1a Sample constraint boxes	4
1b Sample constraint boxes	5
2 Sample optical character recognition (OCR) code	6
3 Sample Magnetic Ink Character Recognition (MICR) code	6
4 Sample bar code	7
5 Samples of Kermit code	8
6 Handprint example	10

Annexes

A Suggested readings	13
B Examples of forms designs guides	14
C Generation degradation table	14

Foreword Forms may be used to collect information and convert it into computer-processable data. Therefore, it is necessary to make a form usable, marketable, and scanable. Making a form appealing with color or graphics may increase the probability of it being filled out, but may also decrease its scanability and increase the amount of data storage required for its scanned image. This document provides guidance for optimizing the design of printed forms used with electronic image management (EIM) systems.

Data captured from forms represents a significant market, tens of billions of dollars in size, and establishing general methods for forms design will greatly assist insurance, financial, and other large organizations that depend on data capture and manipulation from scanned forms on a regular basis. Today's banking industry could come to a halt if the relatively simple transfer authorization, the personal check, did not easily convert to hundreds of bits of electronically-readable data. Most business forms carry little or no such functional data. In a specific, but familiar case, the United States Internal Revenue Service (IRS) is responsible for processing data from billions of pieces of paper each year, yet a very small percentage of that paper contains electronically-readable data for easy processing by an EIM system. The majority of documents received by the IRS contain handwritten or handprinted information that is manually keyed.

Many government agencies and large organizations are setting goals for optimizing the design of their forms using electronic imaging, increasing data capture using electronic recognition technology, and reducing the emphasis on manual keying.

This technical report covers different forms design characteristics including the following:

- a) type fonts;
- b) handprint characters;
- c) weight, quality, and size of paper;
- d) layout; and
- e) print clarity and quality.

References to additional reading material can be found in Annex A, Suggested Readings.

Publication of this ANSI Technical Report has been approved by the Accredited Standards Developer, Association for Information and Image Management (AIIM). This document is registered as a Technical Report series of publications according to the *Procedures for the Registration of ANSI Technical Reports*. This document is not an American National Standard and the material contained herein is not normative in nature. Comments on the content of this document should be sent to Association for Information and Image Management (AIIM), 1100 Wayne Avenue, Suite 1100, Silver Spring, Maryland, 20910.

At the time it approved this technical report, the AIIM Standards Board had the following members:

Marilyn Courtot, <i>Chair</i>	Association for Information and Image Management
Thomas C. Bagg	National Institute of Standards and Technology
Thomas E. Berney	Consultant
Loretta D'Agnolo	American Express Company
Bruce Evans	3M Company
Bruce A. Holroyd	Eastman Kodak Company
Don Klosterboer	Anacomp, Inc.
E. Brien Lewis	I-NET, Inc.
Alan S. Linden	Wang Laboratories
Charles A. Plesums	USAA
George Thoma	National Library of Medicine
Charles F. Touchton	IBM Corporation
Herbert J. White II	Genealogical Society of Utah

The AIIM Electronic Imaging/Input Committee, C13, approved this technical report. The committee had the following members at the time this technical report was approved:

<u>Name of Representative</u>	<u>Organization Represented</u>
Linda Wallace, <i>Chair</i>	FileNet Corporation
Larry Albertson	Applied Image, Inc.
Tom Atwood	IMNET Corporation
Thomas C. Bagg	National Institute of Standards and Technology
Alan Bain	Smithsonian Institution
Gerald Bensi	Bell & Howell Company
Chuck Biss	Photographic Sciences Corporation
Bob Blackwelder	Florida Sheriff's Department
Sylvie Bokshorn	Polaroid Corporation
Robert W. Bristol	Xerox Corporation
Bill Cox	Qualis International
Wayne Doran	NCR Canada Ltd.
Jack Eisen	MIT Microreproduction Laboratories
Eric Erickson	Genealogical Society of Utah
Jon M. Fech	First National Bank of Chicago
Tom Fine	Nevada Micrographics
Richard Gershbock	Information International
Scot Gilkenson	Cognetics Corporation
Steven Gilheany	International Imaging, Inc.
Ralph Grant	3M Company
Pawan Gupta	Kofax Image Products
Joseph G. Hardy	US Department of the Army
William Hooton	I-NET, Inc.
Paul Horowitz	Xerox Corporation
Dean Hough	Kofax Image Products, Inc.
Jonah Howells	Bell & Howell Company
John Jamieson	Blueridge Technologies, Inc.
Clara F. Jehle	State of Alabama
Don S. Kyser	Computer Microfilm Corp.
Richard Leslie	Polaroid Corporation
E. Brien Lewis	Consultant
Basil Manns	Library of Congress

Bob Matthews
Wanda McCraw
Rodney Miller
Paul G. Montgomery
James J. Mott
Hidetoshi Nakamura
John Nelson
Michael O'Donnell
Walter A. Orlof
Shahzad S. Qazi
Bruce Radl
Fred Ratterman
Brent M. Reber

Don Shulsinger
Herman R. Silbiger
Tom Stack
David Stephens
Douglas O. Stewart
Hiraoki Sugihara
A. J. Wicks

Canon USA, Inc.
Duke Power Company
Eastman Kodak Company
A & P International
Polaroid Corporation
Canon USA, Inc.
Image Technologies
Wang Laboratories
US Department of Army
Eastman Kodak Company
Ektron Applied Imaging
Access Corporation
Genealogical Society of Utah
PIXEL Magic
Applicom
L.P. Associates
AT&T Corporation
FileNet Corporation
Canon USA, Inc.
Wicks & Wilson Limited

The AIIM Electronic Imaging/Forms committee, C13.8, developed this technical report. The committee had the following members at the time this technical report was developed:

Linda Kempster, <i>Chair</i>	Strategic Management Resources, Ltd.
Betty Burton	Department of the Army, Image Systems Support Directorate
Larry Dambrose	Department of the Treasury, Internal Revenue Service
Lei Ge	NCR Corporate Imaging Systems Division (ISD)
Margaret Hobson	Lincoln National Corporation
Basil Manns	Library of Congress
Dean Perry	Department of the Treasury, Internal Revenue Service (IRS)
Marty Waxman	Marada
Kermit Way	Department of the Treasury, Internal Revenue Service (IRS)
James Weller	International Business Forms Association (IBFI)

Technical Report for Information and Image Management — Paper Forms Design Optimization For Electronic Image Management (EIM), ANSI/AIIM TR32-1994

1 Scope and purpose

1.1 Scope

This document provides guidelines for the design and creation of printed paper forms to be filled out and returned by users and scanned for processing by EIM systems.

1.2 Purpose

The purpose of this technical report is to provide information on characteristics of printed forms that will make them readily accepted in various EIM applications. This document covers forms characteristics that effect scanning. It also addresses forms layout, recognition technology, scanner performance and data processing, and the effect on data capture and data storage. This ANSI technical report is not intended to address forms removal technologies or the design of electronic forms.

2 References

2.1 Referenced American national standards

ANSI/AIIM MS44-1988 (R1993), Standard Recommended Practice for Quality Control of Image Scanners.

ANSI/AIIM MS52-1991, Standard Recommended Practice for the Requirements and Characteristics of Original Documents Intended for Optical Scanning.

ANSI X3.17-1981 (R1989), Character Set for Optical Character Recognition (OCR-A).

ANSI X3.182-1990, Guideline for Bar Code Print Quality.

ANSI X3.45-1982 (R1989), Character Set for Hand-printing.

ANSI X3.49-1975 (R1989), Character Set for Optical Character Recognition (OCR-B).

ANSI X9.27-1988, Print Specifications for Magnetic Ink Character Recognition (MICR).

2.2 Referenced international standards

ISO 7811/1-1985, Identification Cards-Recording Technique-Part 1: Embossing.

European Computer Manufacturers Association (ECMA) Standard Number Three, CMC7 Printed Image Specifications, Second Edition, 1966.

2.3 Referenced publications

AIIM TR2-1992, Technical Report for Information and Image Management — Glossary of Imaging Technology.

Automatic Identification Manufacturers, Uniform Symbology Specification, 1993, A Subset of Code 39.

Federal Information Processing Standards Publication (FIPS PUB) 40, 1976, Guideline for Optical Character Recognition Forms.

2.4 Related standards

ANSI X3.62-1987 (R1990), Information Systems — Optical Character Recognition (OCR) — Paper Used in OCR Systems.

ANSI X3.86-1980 (R1987), Inks, Optical Character Recognition (OCR).

ANSI X9.13-1990, Placement and Location of MICR Printing, Specifications for.

ANSI X9.27-1988, Print Specifications for Magnetic Ink Character Recognition.

ANSI X3.93M-1981 (R1989), Optical Character Recognition (OCR) Positioning.

ANSI X3.96-1983 (R1991), Forms, Paper Sizes for Single-Part Continuous Business.

ANSI X3.99-1983 (R1991), Guideline for Optical Character Recognition (OCR) Print Quality.

ANSI X3.111-1986 (R1992), Optical Character Recognition (OCR) Matrix Character Sets for OCR-M.

ANSI X3.151-1987, Bond Papers and Index Bristols — Common Sheet Sizes.

ANSI X3.209-1992, Information Systems — Optical Character Recognition (OCR) — Matrix Character Sets for OCR-MB.

ANSI/ASME Y14.2M-1992, Engineering Drawing and Related Documentation Practices — Line Conventions and Lettering.

ANSI Z39.48-1984, Permanence of Paper for Printed Library Materials.

TAPPI T452 OM-1992, Technical Association of the Pulp and Paper Industry — Standard For Brightness of Pulp, Paper, and Paperboard (Directional Reflectance at 457 nm).

TAPPI T519 OM-1991, Technical Association of the Pulp and Paper Industry — Diffuse Opacity of Paper (d/0° paper backing).

TAPPI T527 CM-1992, Technical Association of the Pulp and Paper Industry — Color of Paper and Paperboard in CIE Y; X, Y, or Y, Dominant Wavelength and Excitation Purity.

TAPPI T538 OM-1988, Technical Association of the Pulp and Paper Industry — Standard For Smoothness of Paper and Paperboard (Sheffield Method).

ISO 216-1975, International Standard for Writing Paper and Certain Classes of Printed Matter — Trimmed Sizes — A and B Series.

3 Definitions

The following definitions apply to terms that appear in this technical report. Other terms are defined in AIIM TR2. In addition, relevant definitions may be found in the documents listed in Annex A, Suggested Readings.

3.1 Alphanumeric

Pertaining to a character set that contains letters, numbers, and usually other characters, such as punctuation marks and symbols.

3.2 Bar code

Array of vertical rectangular marks and spaces in a predetermined pattern.

3.3 Code density

Number of code elements that can appear per unit of length.

3.4 Density

Measurement of an image (printed using lithographic ink or toner) by determining the amount of incident radiant energy (light) reflected by a test sample.

3.5 Dropout ink

Ink that cannot be detected by a scanner. A scanner's light source (used to illuminate a form) may be "blind" to this ink color, or the ink color may be filtered out of the light source. There are no "typical" dropout ink colors and those used should be specified by an imaging equipment manufacturer.

3.6 Font

Complete family of a given size of type, including capitals, small capitals, and lower-case, together with figures, punctuation marks, ligatures, etc. Italics are spoken of as a separate font.

3.7 Generation

One of the successive stages of photographic reproduction of an original or master. Copies made from a first generation are second generation, etc. For example, the first generation is either a printed or electronic version of a general purpose business form. See generation, Nth.

3.8 Generation, Nth

(of an electronic image) The number of generations from the original. Example: the 2nd generation (A-2, B-2, etc.) is a copy of the printed or electronic form. See generation.

3.9 Generation, printing

Use of copies from an original or master to make additional copies. See Generation, Nth.

3.10 Generation test

Means of determining the number of times usable copies may be reproduced from succeeding generations. In this test, copies are successively reproduced until a print has been generated that is unusable. This indicates the anticipated range of generation copies that may be reasonably expected. For example, the generation test covers converting a form through printed or electronic versions, i.e., creating subsequent generations by printing, scanning, and re-scanning.

3.11 Landscape orientation

Mode of rendering an image in which the vertical dimension of the presentation is smaller than the horizontal dimension. Contrast with portrait orientation.

3.12 Magnetic Ink Character Recognition (MICR)

Machine recognition of digits printed with ink that can be magnetized.

3.13 Optical Character Recognition (OCR)

An optical technique by which characters can be machine-identified, then converted into computer processable codes, e.g., American Standard Code for Information Interchange (ASCII), Extended Binary Coded Decimal Interchange Code (EBCDIC).

3.14 Pitch (character)

The number of characters per inch measured horizontally. Fixed spacing printers have the same pitch for every letter, regardless of the letters' widths. Proportional spacing has varying pitch, depending on the letter.

3.15 Portrait orientation

Mode of rendering an image in which the vertical dimension of the presentation is greater than the horizontal dimension. Contrast with landscape orientation.

3.16 Print Contrast Signal (PCS)

A measurement of the print contrast where the $R(b)$ is the average reflectance of the background measured in white light. $R(p)$ is the average reflectance of the print measured in white light. PCS is the result of subtracting $R(p)$ from $R(b)$ and dividing the difference by $R(b)$.

3.17 Recognition zone

The area around a recognition data field that is free of other data.

3.18 Sans serif

A typeface without a serif.