Document Indexing

using the

SCAMAX® Document Scanner

Version 10 / 2004

InoTec GmbH . Biedrichstraße 11 . 61200 Wölfersheim
Tel.: +49 (0) 6036 9708-0 . Fax: +49 (0) 6036 9708-15
e-mail: inotec@scamax.com . http://www.scamax.com

Table of Contents

1. Introduction

Dear Reader,

this manual is intended to give you information about the capabilities of the SCAMAX® scanners to index documents during the scan process. It also attempts to show you how to use these features in everyday applications.

To fully understand the various terms used in this manual please read the chapter „Configuration“ in the user manual. It explains the methodology how to set the indexing parameters.

This manual only describes the use and configuration of the indexing parameters for the SCAMAX® document scanners. Please refer to the respective manuals for the operation of the Scan-Software and Scan-PC.

1.1 Necessity for Document Indexing

Indexing is a methodology that archives a document (be it in analogue or digital form) in such a way that it can be retrieved easily and with little effort.

One has to be acutely aware that a scanner does nothing more than produce a digital copy of each document. It is not capable of identifying what kind or type of document it is, neither does it extract the contents of the document.

To be able to retrieve a specific image out of countless others, it has to be given a unique identifier. The identifier either be meaningful in itself or must be linked to meaningful information.

1.2 Index Types

When archiving documents in digital form we distinguish between two types of indexing:

1.2.1 Descriptive Indexing

Descriptive indexing is the modern type of archiving documents. However, this method requires relatively high effort, both technically and in personel. It also requires the employment of appropriate management software (DMS). Data are extracted from the documents either via data entry or software to be stored in conjunction with the images.

If we use incoming mail archiving as an example, appropriate keywords would be **document type**, **document date**, **sender** and possibly **document number**. If this information, together with image filename and storage location, is passed to a document management system, subsequent retrieval can be fast and accurate by simply entering the desired keyword/s.

1.2.2 Structured Indexing

This is the traditional type of indexing based on logic. It is widely used in both analogue and digital archiving of documents. The creation of structured indexes requires considerably less effort than descriptive indexing, but retrieval takes much longer since all searches must be performed sequentially.

Paper Archive

It is the best known form of structured indexing with nearly every office having one in the form of filing cabinets or ring binders. The most common structure is binder number or name, register (A-Z or 1-31) and document date or number.

Microfilm Archive

Structured indexing is a standard procedure in microfilming. Apart from roll number a three-level blipcode is commonly used. These are defined as short, medium or long blips. Blips are bars of three different width and are exposed onto the bottom edge of the film. Each width defines a different level in the structure.

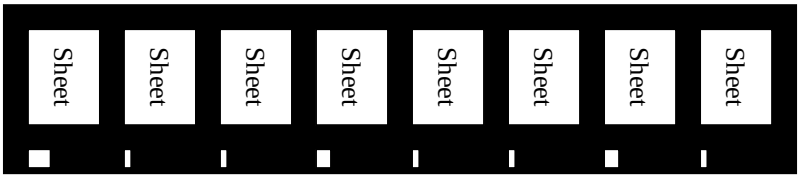


Illustration Microfilm

Image Archive

When structured indexing is used for digital archiving, a pre-defined index structure is used to store the scanned images. It is not uncommon to use a name structure for this purpose.

Since current information technology provides for file and folder names to be up to 255 characters long, the generation of so-called „talking“ structures or names is almost limitless.

Examples:

D:\Images\INV2003\04\Doc0387\0006.tif
D:\Images\INV_2003_04_0387_0006.tif

Without much prior training one can conclude that the example points to page 6 of document 387 for the invoices of April, 2003.

It is not absolutely necessary to use a management system for a digital archive of this type. The image or document can always be roughly identified by its name or storage location.

2. Index Functions of the SCAMAX® Scanners

SCAMAX® document scanners allow you to generate a multi-level index structure during the scan process. You don't need expensive capturing software to generate the data required for a structured index.

These data are passed to any subsequent document management or workflow system as image address within the image header. Detailed information about the structure of the image header can be found in chapter 4.1

2.1 Scanner Mode

Various fields are available within the scanner configuration for the generation of a definitive image address. These are, together with some other parameter settings, stored in the so called mode memory . A total of sixteen (16) modes are available in SCAMAX® scanners, which allows up to 16 jobs to be configured and saved.

As you have already read in chapter 5 of the SCAMAX® user manual the 16 modes are configured and set directly on the scanner using the so called TSCP (TouchScreen-Communication-Panel). No special administrator program is required for this task.

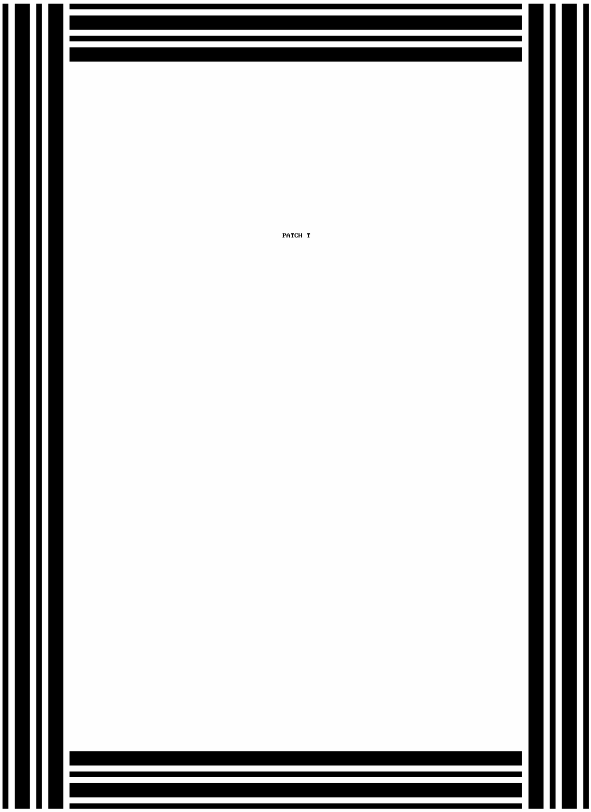
When the SCAMAX® scanner is switched on it will always initialise mode zero (0), regardless of which mode was active prior to switch off. A specific mode can be loaded using either the TSCP (see User Manual) or via the driver interface in your scan client.

The following explanations will be limited to the index settings in the menu area. For a complete description of items in the menu structure and the operation of the onscreen keyboard please refer to chapters 5-7 in the user manual of your SCAMAX® scanner. A diagram of the menu structure (in A5 format) was packaged with the user manual and is part of the standard items supplied with your SCAMAX® scanner.

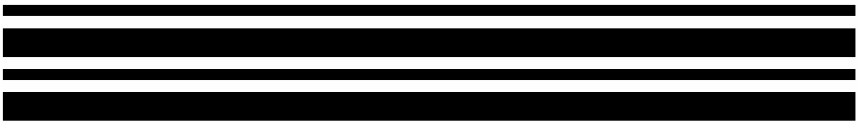
2.2 Patchcodes

The term „Patchcode“ will pop up again and again in the following chapters. A patchcode is a “Control Code”, which originated in the microfilm era. This code is still used today despite the myriad of barcode standards. The reason is its simplicity and clarity, which ensures easy and safe detection by the scanner.

A patchcode is made up of four bars with two different thicknesses. This allows six different patchcodes. A patchcode is almost always printed on a separate patchcode sheet. The most common form of printing a patchcode is as shown in the illustration, which ensures the patchcode is always recognized, regardless of the orientation in which the sheet is fed into the scanner.



2.2.1 Available Patchcodes



PATCH T

PATCH 2

PATCH 3

PATCH 1

PATCH 4

PATCH 6

Print-ready copies of the patchcode sheets in various formats can be found on the „Driver & Utilities“ CD which was supplied with your SCAMAX® scanner.

2.2.2 Patchcode Parameters

Patch-T is a transfer patch, it was used to terminate a document (Start/Stop-Patch). Patchcodes 2,3,1,4 and 6 are socalled comand patches.

Under job configuration you are able to define the patchcode read area (refer also to point 7.7 of the user manual).

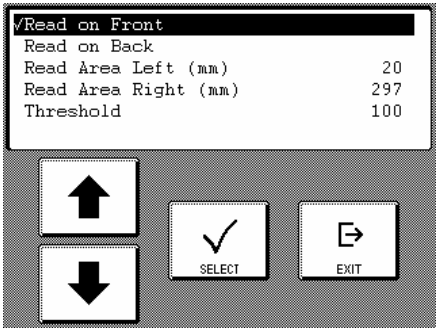


Illustration PatchDecoder Screen

The patchcode detection parameters of front and back page can be set independently. Either one or both parameters must be activated depending on whether patchcodes must be read on the front only, the back only or on both sides.

The parameter ‚read area' limits the horizontal recognition field. Values entered are always relevant for the entire scan area, i.e. **left edge of scan area = 0**. If, for example, the left read area is set to 50mm and the right one to 250mm a patchcode will only be recognised if it occurs within those 200mm. To recognise patchcodes across the entire scan area set the left margin to **0** and the right one to **316**.

Vertically a patchcode is only detected within the first 50mm (2 inches) of the document. Thereafter patchcode detection is inactive.

2.3 Index Settings

You will find the definitions for the image address content in the menu item „Index Settings“. Up to 50 positions are available in each mode (job). They are filled from the four different counters and the fixed text (Fixtext). The counters and the Fixtext can have a maximum of 10 characters each.

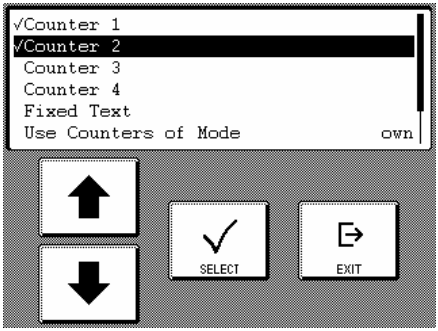


Illustration Index Settings Screen

2.3.1 Counter (1-4)

Configuration of the 4 definable counters is identical and covered as a common item in the following. Before starting definition, it must be determined whether counters are dependent on each other and if so, in which way. A mode configuration sheet containing all counter control parameters can be downloaded from the „Driver & Utilities“ CD supplied with your SCAMAX® scanner.

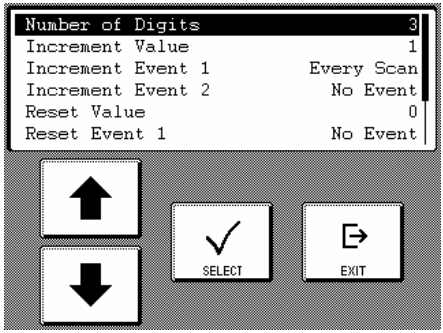


Illustration counter parameters Screen

Number of Digits

The default setting is „inactive“, indicating the counter is not used. If the counter is to be used you must select a value from „1 character“ to „10 characters“. Depending on the number of characters defined, leading zeroes are inserted. Once the counter reaches its maximum value (depends on the number of characters defined), it is reset to zero (0).

Increment Value

This parameter defines how much the counter is incremented for each event. The default setting is „1“ for normal decimal counting (1,2,3,4,5,6,7,8.....). Any increment is possible, which makes it easy to comply with requirements of post-processing programs. In praxis what happens is that the value entered under increment is added to the previous value for each counter event. (Increment=10 | Counter=10,20,30,40,50,60,and so on).

Increment Event

The counter vent defines the timing for increasing the counter by the increment value (see above). The following events can be chosen:

- **No Event –** Counter event inactive.
- **Patch T (2,3,1,4,6) –** Counter is incremented when the respective patchcode is detected.
- **Any Patch –** Any patchcode will trigger a counter increment.
- **Counter X triggered –** The counter is incremented when counter „X“ has reached its trigger value (refer to Trigger Value). Only the other three counters can be selected, i.e. for counter 1 the counters 2, 3 and 4 are available.
- **Every Scan –** counter is incremented for each scan.
  Please note that with a duplex scanner images for the front and back receive the same counter value! (For further details refer to Image Header)

Two independent counter events can be defined for each counter. They are shown separately as increment event 1 and 2 in the parameter list.

Reset Value

Defines the value the counter is reset to when a reset event occurs (see next point). The default value is zero (0).

Reset Event

A reset event determines the point in time when a counter is reset to the specified value (see point above). The following events are selectable:

- **No Event –** Reset event is inactive.
- **Patch T (2,3,1,4,6) –** When the patchcode specified is recognised, the counter is reset.
- **Any Patch –** Any patchcode recognition will trigger a counter reset.
- **Counter X triggered -** The counter is reset, when counter „X" has reached its trigger value (refer to Trigger Value). X can be any of the four counters, i.e. a counter can trigger a reset event when it reaches its own trigger value.
- **Every Scan –** The counter is reset after each scan.

Each counter can have up to three independent reset events defined. They are shown as reset event 1, 2 and 3 in the parameter list.

Trigger Value

When the value defined for this parameter is reached the counter passes a trigger signal (see counter and reset events).

Please Note: *the counter must reach the value specified* ***exactly.*** *Exceeding the value specified does not trigger anything. Example: a counter has been specified to have a reset value of "0", an increment of "2" and an odd trigger value = this would have no effect whatsoever!*

If the trigger value of a counter is used as reset event, the value should always be specified as ***1 higher*** *than the actual value needed! Example: it is determined that 100 documents are to be scanned per batch. This requires that document counter 1, which triggers a counter event for batch counter 2, must be set to 101!*

2.3.2 Fixed Text

In addition to the four counters, static information by way of an alphanumeric definition can be defined. It can have a maximum of ten characters. The information is input via the onscreen keypad (refer to user manual). Any combination of upper and lower case as well as special characters is allowed. This socalled „Fixtext“ is part of the image address and is transmitted in the image header.

2.3.3 Use Counters of Mode

This parameter allows using the counters of another mode. That means, several modes can utilise counter values from one particular mode. In real life, this is used for processes that require definitions across several modes.

Within this parameter you can choose between „Use Own Counters“ and „Use Counters of Mode X“ (X = 0-15). The first choice only allows using mode internal counters. The second choice allows use of counters from any other mode selected. *Please note that this only applies to the **Counter Values** and not the events!*

2.3.4 Ignore First Increment Event

If this option is activated, the first increment event after a counter value modification will be ignored. This secures the validity of this counter value for the first sheet if it starts an increment event. (each scan).

If i.e. a document counter must start with ’0’ and will be incremented by each patchT. All documents begin with a patchT-page for increasing this counter. You have to use this option, because otherwise the first page of the first document after a counter reset will increase this counter to ’1’.

2.3.5 Default Settings

This selection serves to initialize the counter with standard settings. The counter values remain unchanged. In default mode a counter is **not active** and the following settings have been stored as default:

- *Number of Digits* *0*
- *Increment Value* *1*
- *Increment Event 1* ..... *No Event*
- *Increment Event 2* ..... *No Event*
- *Reset Value* *0*
- *Reset Event 1* ..... *No Event*
- *Reset Event 2* ..... *No Event*
- *Reset Event 3* ..... *No Event*
- *Trigger Value* *0*

2.4 Flags

An additional index function apart from the four counters and fixed text, is available, it is a so-called Flag. Flags are „Markers“ that can be activated either by pressing a key on the display panel (refer chapter 5 of the user manual) or via the footswitch (see next point) as required. The image header has the capacity to store two different flag types.

The **Momentary Flag** must activated for each scan. That means it must be set **prior** to the document being pulled in. It is automatically cancelled for subsequent scans.

The **Latched Flag** is activated like a flip-flop switch. This means as long as the option is active, the latched flag is set for every scan. To deactivate the option the switch must be pressed again.

2.5 Footswitch

As detailed in the user manual, up to two footswitches can be connected to a SCAMAX® scanner. The footswitches can perform various control functions which would otherwise have to be done via the display.

On the one hand this may be functions used to control the scanner, like Scan-Stop or Double Feed Detector.

On the other hand the footswitches can also be used to control the index values. This can mean that setting of a flag (see above) or a patchcode function may be executed by pressing the footswitch.

3. Practical Examples of using Counters

Having shown you the varying indexing possibilites available with SCAMAX® scanners and their parameters, we would like to demonstrate their use by way of some practical examples.

From the previous chapter you know that counters in a scan job can be incremented by linear events (a scan or another counter) or when a particular patchcode is detected.

3.1 Counter Control by Linear Events

As the word „linear“ implies, this method is only suitable for the generation of synchronous, continuous structures. In real life this is used when data is passed to a post processing system or for the controlled document flow through the scanner.

Example 1:
Management gives instructions to the records manager to digitise all documents kept in the various departments. The documents are delivered in ring binders, boxes and manila files. After scanning, they are to be bundled and placed into boxes to be stored in the archive. The instructions are to place 5,000 documents in each box, divided into batches of 250. For audit purposes documents must be able to be retrieved using the image address/header.

To process the documents a SCAMAX® scanner fitted with an endorser is available. The scan client used is capable of converting the index data supplied by the scanner into file names.

The administrator defines a scan job with the following parameters:

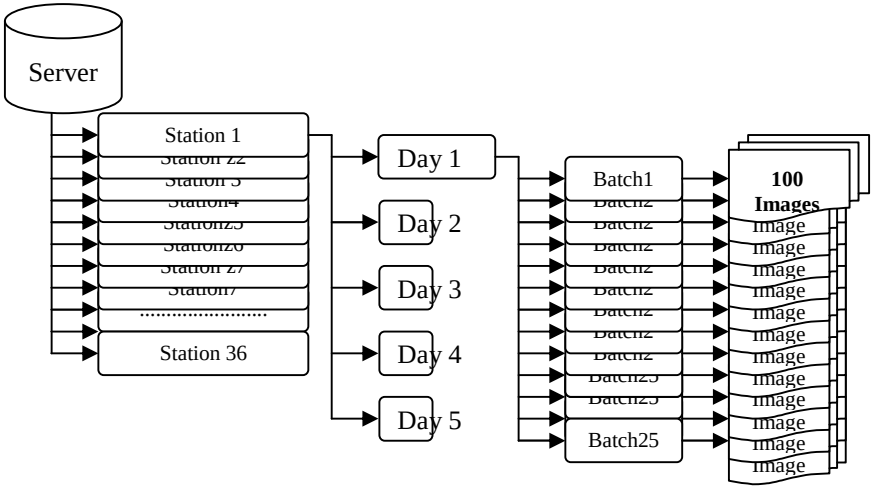
Parameter	Counter 1	Counter 2	Counter 3	Counter 4
Number of Digits	3	2	4	0
Increment	1	1	1	0
Increment Event 1	Every Scan	Count.1 triggered	Count.2 triggered	No Event
Increment Event 2	No Event	No Event	No Event	No Event
Reset Value	1	1	1	0
Reset Event 1	Patch T	Patch T	No Event	No Event
Reset Event 2	Count.1 triggered	Count.2 triggered	No Event	No Event
Reset Event 3	No Event	No Event	No Event	No Event
Trigger Value	251	21	0	0

- *Endorser String.....Cntr03_Cntr02_Cntr01_CCYYMMDD*
- *Print Control.....Print on (Default)*
- *Scan Stopp Event .....Counter 1 triggered*

Definition:
Counter 1 is the document counter. It only needs three digits and is incremented by one for each scan. When the 251st document is processed, counter 1 reaches its trigger value, resets itself to 1, increments counter 2 by 1 (its defined increment) and triggers the scan stop event. This allows the operator to bundle the batch. Counter 2, which is the batch counter, only needs two digits since its maximum value will be 20 (20x250 documents = 5000 documents). Counter 3 is the box counter and is triggered by counter 2. It does not have or need a trigger value or reset value. The number of digits per counter is variable. Counters 1 and 2 are reset upon reaching their respective trigger values or when a Patchcode T is detected. This allows the operator to force the start of a new box by inserting a Patchcode T sheet. The endorser prints, on the back of each document, a string consisting of: counter3_counter2_counter1_EuroDate (example: 0002_08_123_20040421). The counters contain variables with leading zeroes. (Detailed information about definitions, print position, print format and so on can be found in the scanner's user manual). The scan software uses the string printed to construct the file name for each image. Thus, the file name defines precisely the box number, batch number and document number.

Example 2:
A service bureau scans documents that are then passed on to the in-house data entry section. The production manager is instructed to organise the image data in such a way that all data entry stations receive an equal amount of work for five days. He knows there are 36 data entry stations, each one can process 25 batches per day. Batches are to be limited to 100 documents each. Data are to be stored on a server in the following structure:

Schematic Illustration:



A SCAMAX® scanner is used for imaging. The scan software is capable of using the index data supplied by the scanner for further processing tasks.

The production manager creates a scan job with these parameters:

Parameter	Counter 1	Counter 2	Counter 3	Counter 4
Number of Digits	8	8	2	1
Increment	1	1	1	1
Increment Event 1	Every Scan	Count. 1 triggered	Count. 1 triggered	Count. 3 triggered
Increment Event 2	No Event	No Event	No Event	No Event
Reset Value	1	1	1	1
Reset Event 1	Count. 1 triggered	No Event	Count. 3 triggered	Count. 4 triggered
Reset Event 2	Patch T	No Event	Patch T	Patch T
Reset Event 3	No Event	No Event	No Event	No Event
Trigger Value	101	0	37	5

Definition:
What initially looks like a definition error is, on closer inspection, a really nifty solution. Four counters are used in this scan job. It is quite obvious that counter 1 is used as document counter since it is incremented on each scan and has a trigger value of 101. Counter 1 also serves as counter event for counters 2 and 3. Counter 2 is the batch counter. Every batch must be indentifiable by a single number. For that reason, the counter is not reset. Counter 3 is used to track the 36 data entry stations, thus it resets itself when the rest value of 37 is reached. Counter 3 also increases counter 4 upon reaching its reset value. Counter 4 tracks the working days. Since every data entry station receives an identical volume of data the value of 25 batches per day is of no consequence in the scan process. A detected PatchT will reset the counters 1, 3 and 4.

The following structure can be generated using the counter data:

//Server\Station<Counter3>\Day<Counter4>\Batch<Counter2>\Counter1.tif

In real life it would look something like this:

F:\Station01\Day1\Batch00000001\00000001.tif
F:\Station01\Day1\Batch00000001\00000002.tif
F:\Station01\Day1\Batch00000001\00000003.tif
……
F:\Station23\Day4\Batch00000131\00000099.tif
F:\Station23\Day4\Batch00000131\00000100.tif
F:\Station24\Day4\Batch00000132\00000001.tif
……
F:\Station36\Day5\Batch00000180\000000099.tif
F:\Station36\Day5\Batch00000180\00000100.tif
F:\Station01\Day1\Batch00000181\00000001.tif
and so on

3.2 Counter Control via Patchcodes

In contrast to the regularity in linear processing, as discussed previously, counter control via patchcodes is quite chaotic. When patchcodes occur, and in which sequence, is unpredictable and entirely dependent on circumstances.

This mode is used when volumes in structures and increments are changing all the time. A typical example would be accounts payable, where the originator (supplier) as well as amount and types of documents received per day varies widely.

Example 1:
The accounts payable department of a large industrial corporation decides to bring all supplier invoices from the last ten years into a digital workflow system. Departmental staff will perform any subsequent processing required. Currently the invoices are stored in ringbinders. One binder can contain invoices from one or more suppliers, filed by supplier number, fincancial year and date.

Document preparation consists of removing invoices from binders and sorted into the following hierarchy: batch (a financial year), bundle (one supplier) and parcel (an invoice with several pages). A Patch 3 sheet follows each parcel.

A SCAMAX® scanner fitted with an optional footswitch is used for processing. The scan software gathers image and header data to generate image files with appropriate index files to pass to the workflow system.

The Project Manager defines the following mode configuration:

Parameter	Counter 1	Counter 2	Counter 3	Counter 4
Number of Digits	4	8	8	4
Increment	1	1	1	1
Increment Event 1	Every Scan	Patch 3	Patch 2	Patch T
Increment Event 2	No Event	No Event	No Event	No Event
Reset Value	1	1	1	1
Reset Event 1	Every Patch	Patch 2	Patch T	No Event
Reset Event 2	No Event	Patch T	No Event	No Event
Reset Event 3	No Event	No Event	No Event	No Event
Trigger Value	0	0	0	0

- *Fixed Text* ..... *AP-INV*
- *Footswitch 1* ..... *Patch 2*
- *Menu Key* ..... *Patch T*
- *Patch-Decoder Setting* ..... *Read on Front Page*

Definition:
For this project all four counters are used. As always counter 1 is used as document counter. In this case it is reset by every patchcode read. Counter 2 is incremented every time a patch 3 is encountered, signaling the completion of a parcel. A reset for this counter is only triggered by a patch 2 or patch T. Patch 2, indicating a new supplier, causes counter 3 to be incremented. Patch 2 is not physically read in this case. It has been assigned to the footswitch, since it is quite easy for the operator to press this switch each time a new supplier is encountered. Patch T resets the supplier counter (counter 3) AND triggers counter 4 (the year counter) to be incremented. Since patch T is not required very often, it has been assigned to a key on the display panel.

The scan software can create any index file desired using the information created by the above definition.

Example 2:
The following solution shows a continuous workflow often employed by health insurance companies. SCAMAX® scanners are used to digitise all documents received by the mailroom. The resulting images are sent to a forms/character recognition system. This system automatically recognises invoices, extracts the respective ‚value' fields and transfers the numbers to accounts payable. Unrecognised documents (i.e. everything but invoices) are sent to the general workflow system for allocation to the respective departments/persons for further action. Since it is necessary to reference back to the original documents during the processing period, the endorser is used during the scan process to print the scan date (YYYYMMDD), scanner number and contents of counters 1 to 3) on the back of each document.

The project uses the following mode configuration:

Parameter	Counter 1	Counter 2	Counter 3	Counter 4
Number of Digits	3	2	4	0
Increment	1	1	1	0
Increment Event 1	Every Scan	Patch 2	Patch 3	No Event
Increment Event 2	No Event	No Event	No Event	No Event
Reset Value	0	0	0	0
Reset Event 1	Patch T	Patch T	Patch T	No Event
Reset Event 2	Patch 2	Patch 3	No Event	No Event
Reset Event 3	Patch 3	No Event	No Event	No Event
Trigger Value	0	0	0	0

- *Endorser-String ..... CCYYMMDD_ScannerID_ Cntr03_Cntr02_Cntr01*

Definition:
Three different patchcodes are used to control the flow. Patch T is the so-called „Start of Day" code and resets all counters. Patch 2 is an envelope separator, since an envelope can contain multiple sheets/pages. Patch 2 increments counter 2 (no. of envelopes in batch) and resets counter 1 (no. of sheets per envelope). As always counter 1 is the document counter and is incremented on each scan. Patch 3 increments counter 3, which is the batch counter, and resets counters 1 and 2.

Illustration of Definition:

Start of Day sheet to initialise all counters
Counter1: 000 Counter2: 00 Counter3: 0000

Patch3 – batch separator
Counter1: 000 Counter2: 00 Counter3: 0001

Patch2 – envelope separator
Counter1: 000 Counter2: 01 Counter3: 0001

Counter1: 001 Counter2: 01 Counter3: 0001
Counter1: 002 Counter2: 01 Counter3: 0001
Counter1: 003 Counter2: 01 Counter3: 0001
Counter1: 004 Counter2: 01 Counter3: 0001
etc.....

Patch2 – envelope separator
Counter1: 000 Counter2: 02 Counter3: 0001

Counter1: 001 Counter2: 02 Counter3: 0001
Counter1: 002 Counter2: 02 Counter3: 0001
Counter1: 003 Counter2: 02 Counter3: 0001
Counter 1: 004 Counter2: 02 Counter3: 0001
etc.....

Patch3 – batch separator
Counter 1: 000 Counter2: 00 Counter3: 0002

Patch2 – envelope separator
Counter1: 000 Counter2: 01 Counter3: 0002

Counter1: 001 Counter2: 01 Counter3: 0002
Counter1: 002 Counter2: 01 Counter3: 0002
Counter1: 003 Counter2: 01 Counter3: 0002
Counter1: 004 Counter2: 01 Counter3: 0002
etc.....

4. Image Address

After generating a useable image address using the various configurations and settings, it will be of interest to learn how to use these data.

The scanner transmits, for each image, a so-called image header. The header has a length of 512 bytes (characters) and contains all data relevant to the image. Important: this image header is NOT part of the TIFF header and must not be confused with it.

If the SCAMAX® scanner is controlled by professional capturing software, the image address can, in most cases, be accessed via parameters. If it is desired to drive the scanner via the ISIS driver and 'home grown' scan client, or to integrate it into a workflow environment, the image address can be obtained via so-called ISIS tags. A listing of relevant ISIS tag commands and further information is available upon request.

4.1 Image Header Structure

To gain an overview over the image header structure please refer to the matrix on the following page.

The numbers 0-511 in the grey shaded lines indicate the position within the header. Position jumps within a field are sometimes shown like this: „...".Information relevant for the respective positions is displayed below them.

4.1.1 Offset Positions of Important Image Information

Information	Position
Page Type (Front/Rear)	000-004
Unique Scan-ID	007-016
Scanmode Number	054-055
Fixed Text	109-118
Counter 4	205-214
Counter 3	120-129
Counter 2	131-140
Counter 1	142-151
Momentary Flag (0/1)	155
Latched Flag (0/1)	157
Scan Date (mmddyy)	175-180
Scan Time (hhmmss)	189-194
Scan Resolution in dpi	220-222
Patchcode recognised (0-5 see below)	336
Endorsed (0/1)	349

Patchcode (Offset 336) returns a value as follow:

0=Patch T, 1=Patch II, 2=Patch III, 3=Patch I, 4=Patch IV, 5=Patch VI,

4.1.2 Image Header Matrix

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17						
F	r	o	n	t		#	-	-	-	-	-	-	-	-	-	0							
18	19	20	21	22	23	24	25	26	27	88	29	30	31	32	33	34	35	36					
L	e	n	g	t	h		=		-	-	-	-	-	-	-	0							
37	38	39	40	41	42	43	44	45	46	47													
L	e	v	e	l		=		-	0														
48	49	50	51	52	53	54	55	56															
M	o	d	e	=		-	0	lf															
57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80
L	i	n	e		L	e	n	g	t	h		=		-	-	-	-	-	-	-	0		
81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	
P	a	g	e		L	e	n	g	t	h		=		-	-	-	-	-	-	-	0	lf	
104	105	106	107	108	109	...	118	119	120	...	129	130	131	...	140	141	142	...	151	152	153		
I	A		=		-	-	0	.	-	-	0	.	-	-	0	.	-	-	0				
154	155	156	157	158																			
-	0	-	0	lf																			
159	160	161	162	163	164	165	166	167															
C	m	p		=		-	0																
168	169	170	171	172	173	174	175	176	177	178	179	180	181										
D	a	t	e		=		m	m	d	d	y	y											
182	183	184	184	186	187	188	189	190	191	192	193	194	195										
T	i	m	e		=		h	h	m	m	s	s											
196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215				
R	o	l	l		#		=		-	-	-	-	-	-	-	-	-	0					
216	217	218	219	220	221	222	223																
R	e	s	=	-	-	0																	
224	225	226	227	228	229																		
B	o	=	-	0																			
230	231	232	233	234	235	236	237	238															
S	k	=	-	-	-	0		lf															
239	...	255	256	257	258	...	287																
nu	nu	nu	:	lf	nu	nu	nu																
288	289	290	291	292	293	294	295																
R	e	v	=		-	0																	
296	297	298	299	300	301	302	303	304	305	306	307	308	...	322	323								
I	A	_	F	i	x	_	e	x	t	=		-	-	-									
324	325	326	327	328	329	330	331	332	333	334	335	336	337										
P	a	t	c	h	c	o	d	e	=		-	0											
338	339	340	341	342	343	344	345	346	347	348	349	350	351	...	511								
E	n	d	o	r	s	e	d	=		-	0		nu	nu	nu								

4.2 Use of Index Information

Once the image index has been created during the scan process by using counters, text and other information, this information can be used in various forms.

Two possibilities were touched in chapter 1. The use of index informations to create folder structures or file name structures. Both possibilities are employed in circumstances where the data are archived in either structured or ‚talking' form without employing a retrieval function or program.

Another possibility is to save the information, including the relation to the image, in an index file. The storage location of the image and its naming are of secondary importance in this case. However, what is required is the processing of the data by a retrieval or document management system. These systems ensure a meaningful relation between image and index data. When creating index files we generally distinguish between the itemised and collective form.

The itemised form generates a separate index file for each image. This file is usually saved using the same file name and path as the image. Only the file extension differs (Example: Image=xxxxx.TIF / Index=xxxxx.TXT).

The collective form collects the data for several images, and their relationship to the image, in one file. The size of this collective file is entirely dependent on the respective processing. The index file does not necessarily have to be saved in the same path as the images.

In the following we show some index processing possibilities using the example of the account payable application discussed in chapter 3.2:

- Fixed Text..... AP-INVOICE (10 chars)
- Counter 1 ..... Document counter (4 chars)
- Counter 2 ..... Parcel counter (8 chars)
- Counter 3 ..... Supplier counter (8 chars)
- Counter 4 ..... Year counter (4 chars)

4.2.1 Example – File Name Structure

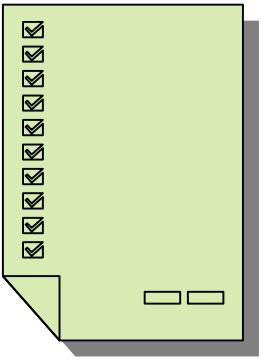
AP-INVOICE_Year1994_SNo00000001_Doc00000001_Page0001.TIF
AP-INVOICE_Year1994_SNo00000001_Doc00000001_Page0002.TIF
AP-INVOICE_Year1994_SNo00000001_Doc00000001_Page0003.TIF
.....
AP-INVOICE_Year1997_SNo00000815_Doc00004711_Page0012.TIF
AP-INVOICE_Year1997_SNo00000815_Doc00004711_Page0013.TIF
AP-INVOICE_Year1997_SNo00000816_Doc00000001_Page0001.TIF
AP-INVOICE_Year1997_SNo00000816_Doc00000001_Page0002.TIF
.....
AP-INVOICE_Year1999_SNo00002456_Doc00000676_Page0027.TIF
AP-INVOICE_Year1999_SNo00002456_Doc00000676_Page0028.TIF
AP-INVOICE_Year2000_SNo00000001_Doc00000001_Page0001.TIF
AP-INVOICE_Year2000_SNo00000001_Doc00000001_Page0002.TIF

4.2.2 Example – Folder Structure

Format:
Drive:\Document Type\Year\Supplier\Document\Page.tif

F:\AP-INVOICE\1994\00000001\00000001\0001.TIF
F:\AP-INVOICE\1994\00000001\00000001\0002.TIF
F:\AP-INVOICE\1994\00000001\00000001\0003.TIF
.....
F:\AP-INVOICE\1997\00000815\00004711\0012.TIF
F:\AP-INVOICE\1997\00000815\00004711\0013.TIF
F:\AP-INVOICE\1997\00000816\00000001\0001.TIF
F:\AP-INVOICE\1997\00000816\00000001\0002.TIF
.....
F:\AP-INVOICE\1999\00002456\00000676\0027.TIF
F:\AP-INVOICE\1999\00002456\00000676\0028.TIF
F:\AP-INVOICE\2000\00000001\00000001\0001.TIF
F:\AP-INVOICE\2000\00000001\00000001\0002.TIF

4.2.3 Example – Index File (Itemised)

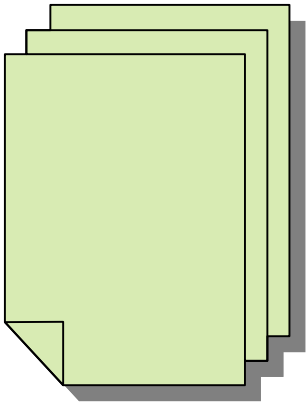


C:\Images\47110815.TIF

```
Document type:   AP-Invoice
   Fin. Year:    1998
    Supplier:    00002456
    Document:    00000676
        Page:    0028
   Scan date:    20040412
     Scanner:    Scamax410
```

C:\Images\47110815.TXT

4.2.4 Example – Index File (Collective)



C:\Images*.TIF

```
Year;Supplier;Doc-No  ;Page;  Date    ;Image
1998;00002456;00000676;0028;20040412;004711.tif
1998;00002456;00000676;0029;20040412;004712.tif
1998;00002456;00000676;0030;20040412;004713.tif
1998;00002456;00000676;0031;20040412;004714.tif
1998;00002456;00000676;0032;20040412;004715.tif
1998;00002456;00000676;0033;20040412;004716.tif
1998;00002456;00000676;0034;20040412;004717.tif
1998;00002457;00000001;0001;20040412;004718.tif
1998;00002457;00000001;0002;20040412;004719.tif
1998;00002457;00000001;0003;20040412;004720.tif
1998;00002457;00000001;0004;20040412;004721.tif
```

C:\Images\AP-INVOICE.TXT

Index