E. FILM CLEANING

Because specks of dirt on film can cause the Retrieval Unit to misread the coded information in the data blocks on the film, the film must be kept clean.

Before film copy is made, either in the expand or copy modes, the film from which the copy is being made will be cleaned. All working reels will routinely be cleaned each month; reels with high usage will be cleaned each week. New film will be cleaned immediately after developing.

A special film cleaning table has been prepared to simplify the process. The reel of film is placed on the lower rewind reel, threaded through the cleaning station, past the blower, and onto the takeup reel. Either Naccolene or Tuff Coat may be used as a cleaning agent, although Tuff Coat is preferred for working reels because of its protective qualities. After the takeup motor has been started, a turn of the reel in the correct direction will start the operation.

CHAPTER VI

RETRIEVAL TECHNIQUES

A. INTRODUCTION

The FileSearch Retrieval Unit searches the coded film and retrieves requested documents. When used to advantage, the system can provide a requester with relevant, complete and timely information in a matter of minutes.

The form of the retrieved documents is determined by the operating mode of the Retrieval Unit. By pushing BROWSE, matching documents can be viewed and desired pages recovered as hard copies. By pushing RETRIEVE, hard copies of all pages of matching documents are obtained automatically. The EXPAND mode results in film copies of selected documents; this may either be 16mm film with images of the documents for viewing on an off-line reader-printer, or machine-readable 35mm film for use on a Retrieval Unit. Using the COUNT HITS mode, retrieval takes the form of a displayed total. The operating procedures are given in FileSearch Retrieval Unit Instructions.

Retrieval techniques are the same for all modes of operation. These techniques, based on the request logic of the Retrieval Unit, comprise the search strategy of the system. If the indexer, retrieval analyst and equipment operator are aware of retrieval techniques, the FileSearch system can perform to optimum advantage. The use of

proper search strategy makes the system more responsive to user requirements by providing relevant documents in less time. Proper search strategy also results in economy of operation by reducing search time and effecting savings in retrieval.

This chapter is concerned with retrieval techniques. The request logic of the Retrieval Unit and console switches, which can be used to modify a request, are discussed. Some of the more common types of requests are illustrated, and the role of the retrieval analyst examined. Additional information concerning retrieval strategy can be found in the DIA guide (Ref. 2).

B. REQUEST LOGIC

1. The Request Form

A request is stated in terms of the parameters of the system using a request form such as shown in Fig. 6-1. The information on the form is punched into a request card (Fig. 6-2) for input to the Retrieval Unit, or input directly using Manual Retrieval Console.

The Retrieval Unit has six registers, labeled A through F. They contain the descriptors of the request which are to be matched with the index terms read from film. The coding of the request descriptors corresponds to the coding of index descriptors.

The function and association determine the subsequent role played by each register during search. The function specifies how each register is to compare descriptors and store hits. The association specifies how the hits recorded by the registers are to be related for recovery of desired documents.

The setting of each register's phrased/factored switch indicates the field assignment of the corresponding descriptor. The switches control the Retrieval Unit to recognize or ignore phrase boundaries during search.

FMA FILESEARCH
REQUEST FORM

REGISTER	ASSOCIATION	DESCRIPTOR TAG	DESCRIPTOR	FUNCTION	SWITCH PHRASED	SWITCH FACTORED
A	X					
B						
C						
D						
E						
F						

Figure 6-1. REQUEST FORM

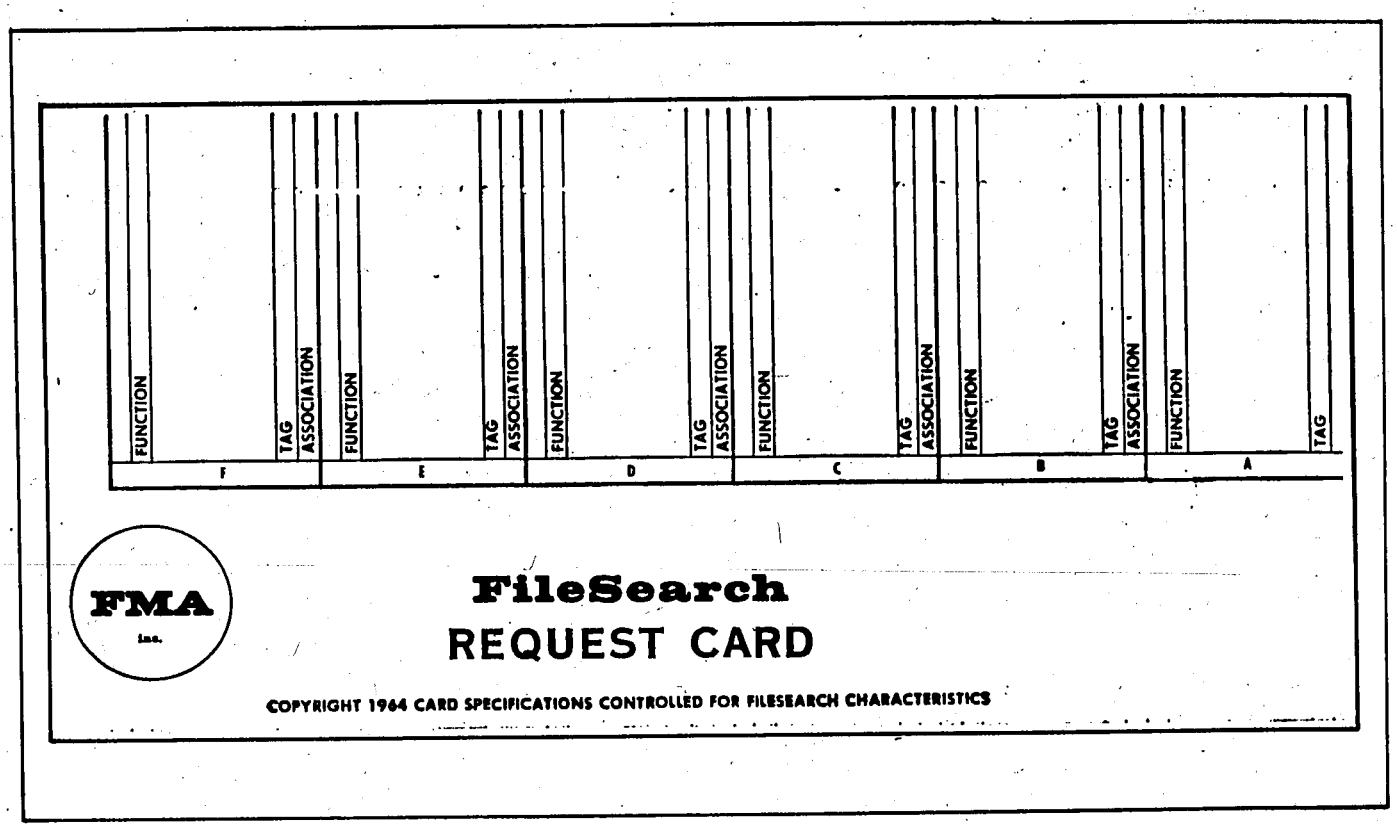


Figure 6-2. REQUEST CARD

2. The Search Sequence

When the Retrieval Unit is searching, a sequence of events occurs during the scanning of the index code from one document.

a. Each character of a descriptor scanned from film is compared with the corresponding character of every descriptor in the request. If there are six descriptors in the request and a tag is scanned from film, it is compared with each of the six tags of the request terms simultaneously. A register can compare on the basis of equality, e.g., documents produced in March 1965; or as a limit, e.g., documents produced between March 1965 (a lower limit) and October 1965 (an upper limit). The normal function (N) and continued function (C) correspond to the equal comparison and the limit function (L) corresponds to the limit comparison.

b. When the last character of a descriptor on film has been scanned (as determined by reading a tab or carriage return) the results of the comparisons are stored in flip-flop memories. Each register has such a memory, and when a hit is recorded, the corresponding indicator light on the Retrieval Unit's console is turned on.

c. The flip-flop memories continue to store the hits until either a phrase boundary (Q-tag) or the beginning of the next document (an arrow) is sensed. At this time, the hits, associated according to the AND, OR, AND NOT of the request, are examined to determine if the document should be retrieved.

3. The Superdigit

Each descriptor in the request always has seven characters. The first is the tag, unused positions to the right are filled out with superdigits. The superdigit,

entered as an asterisk, causes the Retrieval Unit to consider the corresponding character read from film as matching. It can be used in any position of any register, except the tag position, and with any function.

The superdigit is extremely useful in forming requests. It allows documents, indexed with very specific terms, to be recovered for more general (generic) requests. For example, the request term X65Ø3** matches all documents written in March 1965 (e.g., indexed with X65Ø3Ø1, X65Ø3Ø5, X65Ø316). Similarly, the request term X65**** matches all documents written in 1965. A request for the ISC subject health, hygiene and sanitation, coded as A271***, matches documents concerning disease incidence (A271Ø7Ø), medical facilities (A2711 7Ø), water control and supply (A271330), etc.

The superdigit alleviates the problem of spelling clear-text index terms. For example, consider the request for documents concerning a Mr. Clawson; the spelling is uncertain. Reference to a phone book yields several variations in spelling:

Possible Index Terms	Matching Request Terms
CCLAWSO -N.	
CCLAUSE -N.	CCL**S* -N.
CCLAUSO -N.	
CCLOSSO -N.	
CKLASSE -N.	CKLASS* -N.
CKLASSO -N.	
CCLAUSS -ON.	CCLAUSS -*N.
CCLAUSS -EN.	

REGISTER	ASSOCIATION	DESCRIPTOR							FUNCTION	SWITCH	
		TAG								PHRASED	FACTORED
A		C	C	L	*	*	$	*	C	✓	
B	AND	-	N	.	*	*	*	*	C	✓	
C	OR	C	K	L	A	$	$	*	C	✓	
D	AND	-	N	.	*	*	*	*	C	✓	
E	OR	C	C	L	A	U	$	$	C	✓	
F	AND	-	*	N	.	*	*	*	C	✓	

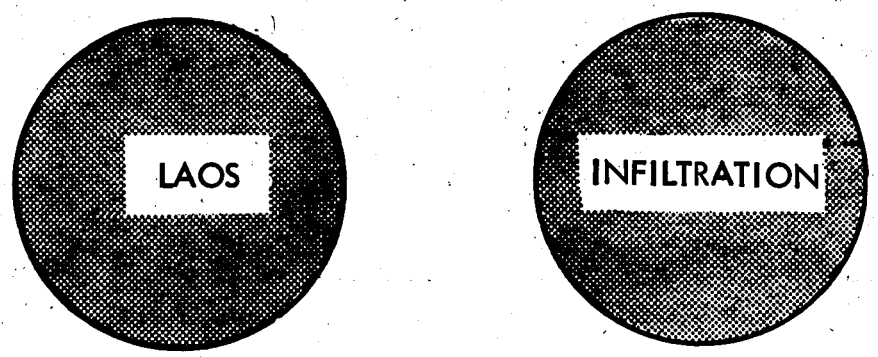
4. The Associations

Figure 6-3 illustrates the three associations AND, OR, AND NOT, for the two request terms Laos and infiltration. If the request asks for Laos or infiltration, documents indexed with either term, or both terms, are recovered. If the request asks for Laos AND infiltration, the retrieved documents must be indexed by both terms. If the request asks for Laos AND NOT infiltration, all documents concerning Laos except those indexed with infiltration are retrieved. The AND NOT association is frequently referred to as EXCEPT.

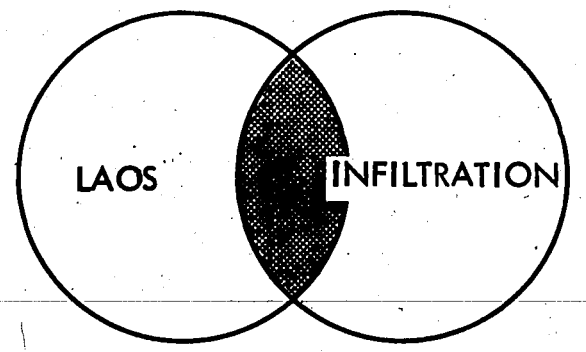
a. AND

The association between the descriptors is dictated by the request, and is usually simple and straightforward. The AND association is probably used most frequently. For example, the request for all documents concerning disease incidence (AND) in Southeast Asia (AND) reported since 1960 involves three ANDed terms. The request for IR's (AND) concerning all medical facilities (AND NOT) except for health resorts (AND) in Addis Ababa, (AND) Ethiopia involves five terms, four of which are ANDed.

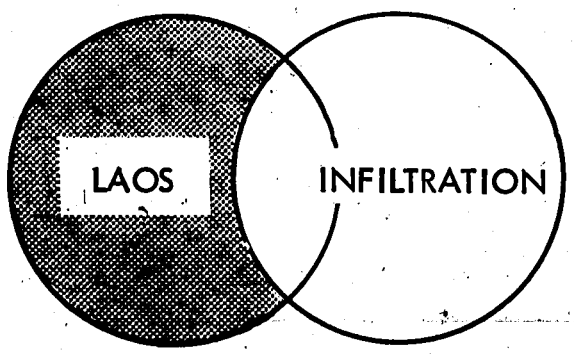
THE SHADED AREA INDICATES DOCUMENTS WHICH ARE RETRIEVED



LAOS OR INFILTRATION



LAOS AND INFILTRATION



LAOS AND NOT INFILTRATION

Figure 6-3. ASSOCIATION LOGIC - TWO TERMS

PAGE(S) 6-9 to 6-32 (IS) (ARE) MISSING FROM ORIGINAL TEXT AND ARE UNAVAILABLE FOR FILMING.

	REQUEST TERM	DESCRIPTOR
10.	Area of South Việt Nam	B A V S * * *
11.	The Việt Cộng	B A V S V C C
12.	North Vietnamese units in South Việt Nam	B A V S V N C
13.	Việt Cộng or North Vietnamese units in South Việt Nam	B A V S V * C
14.	North Việt Nam or its units in South	B A V * V N C

EXAMPLES:

	Switch	Register	Association	Descriptor	Function
1. A request for Typhoid incidence information for Asia in general:					
	Phrased	A		A271Ø75	N
	Phrased	B	AND	BAASØØØ	N
2. A request for Typhoid incidence information in any area in Asia:					
	Phrased	A		A271Ø75	N
	Phrased	B	AND	BA*****	N
3. A request for Typhoid incidence information in South Việt Nam					
	Phrased	A		A271Ø75	N
	Phrased	B	AND	BAVS***	N
4. A request for Typhoid incidence information in the Việt Cộng and North Vietnamese units in the South					
	Phrased	A		A271Ø75	N
	Phrased	B	AND	BAVSV*C	N

	REQUEST TERM	DESCRIPTOR
10.	Area of South Việt Nam	B A V S * * *
11.	The Việt Cộng	B A V S V C C
12.	North Vietnamese units in South Việt Nam	B A V S V N C
13.	Việt Cộng or North Vietnamese units in South Việt Nam	B A V S V * C
14.	North Việt Nam or its units in South	B A V * V N C

EXAMPLES:

	Switch	Register	Association	Descriptor	Function
1. A request for Typhoid incidence information for Asia in general:					
	Phrased	A		A271Ø75	N
	Phrased	B	AND	BAASØØØ	N
2. A request for Typhoid incidence information in any area in Asia:					
	Phrased	A		A271Ø75	N
	Phrased	B	AND	BA*****	N
3. A request for Typhoid incidence information in South Việt Nam					
	Phrased	A		A271Ø75	N
	Phrased	B	AND	BAVS***	N
4. A request for Typhoid incidence information in the Việt Cộng and North Vietnamese units in the South					
	Phrased	A		A271Ø75	N
	Phrased	B	AND	BAVSV*C	N

2. Place Names Searches

Place names are indexed in clear text; the names of South Vietnamese provinces are preceeded by the number of the Corps Tactical Zone (CTZ) in which they are located.

Because of the similarity of Vietnamese place names they should normally be continued in the request to a maximum of twelve characters. If they were not continued many false drops might occur, as with such names Bình Đồng and Bình Đon or Phung Thành, Phung Tường, Phung Tân or Phung Trang.

Requests for all material on a given CTZ may be made by superdigiting all positions after the number indicating the zone. Thus, information on the Việt Cộng provincial party structure in IV corps would be requested:

REGISTER	ASSOCIATION	DESCRIPTOR TAG							DESCRIPTOR FUNCTION	SWITCH PHRASED	SWITCH FACTORED
A	X	A	1	3	5	2	7	5	N	✓	
B	AND	D	4	*	*	*	*	*	N	✓	

3. Searches using UTM Coordinates

a. The Normal Function

Using a normal function and the superdigit, documents for various size square areas may be recovered:

- A 1-Kilometer square B R 2 5 3 6
- A 10-Kilometer square B R 2 * 3 *
- A 100-Kilometer square B R * * * *

The request terms may be combined with subjects or other terms to further define a request, as, rice caches in square BR.

b. The Limit Function

Because UTM coordinates may only be indexed with a single specific area in each indexing phrase, limit logic may be used without problems of false drops. To request all information on the area in square BR, from 23 to 54 east-west and 41 to 57 north-south the query would be written:

REGISTER	ASSOCIATION	DESCRIPTOR							FUNCTION	SWITCH	
		TAG								PHRASED	FACTORED
A		L	B	R	2	3	*	*	L	✓	
B	AND	L	B	R	5	4	*	*	L	✓	
C	AND	L	B	R	*	*	4	1	L	✓	
D	AND	L	B	R	*	*	5	7	L	✓	

This query may be further refined by added subjects or unit designations in registers E and F.

If the area desired extends across two grid squares, two separate requests will need to be made.

4. Retrieving the Việt Cộng Military Regions

Retrieving by the province numbers it is possible to approximate the Việt Cộng military regions. For example, MR5 includes provinces 3 through 11 (Quảng Nam to Khánh Hòa) and 13 (Đarlac). A request for information on medical units in MR5 would be written:

REGISTER	ASSOCIATION	DESCRIPTOR TAG							FUNCTION	SWITCH PHRASED	SWITCH FACTORED
A		A	7	3	5	3	5	Ø	N	✓	
B	AND	D	2	1	3	*	*	*	N	✓	
C	OR	D	*	Ø	3	*	*	*	L	✓	
D	AND	D	*	1	2	*	*	*	L	✓	
E	AND	A	7	3	5	3	5	Ø	N	✓	

Note that when using limits it is wise to superdigit the CTZ position. Also, the upper limit must be one number greater than the last province desired. Figure 6-6 shows the approximate correspondence between Việt Cộng military regions and the province designations.

Region	*Province Numbers*
MR 1	22 - 27 (*Hậu Nghĩa is split between MR-1 & MR-2.*)
MR-2	27 - 28 (*Kiên Giang is split between MR-2 & MR-3.*) 30-- 37
MR-3	37 - 46
MR-4	29, Saigon, Cholon
MR-5	3 - 11, 13
MR-6	12, 15 - 18
MR-10	14, 19, 21
Trị-Thiên-Huế	1 - 2

Figure 6-6. Approximate Correspondence Between Provinces and Military Regions

E. REQUESTS USING NAMES OF PEOPLE

Vietnamese names are entered in full, and the first name of an individual is then reentered in a separate C-tag. This procedure assures maximum flexibility during retrieval with a minimum of false drops. Depending on the amount of information known about the person, the following procedures apply.

1. Full Name

When the full name of an individual is known and it does not exceed twelve characters, the request should be made considering it as a single term. Thus, to retrieve information on Nguyễn văn Dầu the query would be written:

REGISTER	ASSOCIATION	DESCRIPTOR TAG	DESCRIPTOR	FUNCTION	SWITCH PHRASED	SWITCH FACTORED
A	X	C	NGUYEN	C	✓	
B	AND	-	VAN DAU	C	✓	

If the name exceeds twelve characters the first name (tên) can be repeated to avoid the possibility of a false drop, as, for example, between Nguyễn văn Thành and Nguyễn văn Thạng

REGISTER	ASSOCIATION	DESCRIPTOR TAG	DESCRIPTOR	FUNCTION	SWITCH PHRASED	SWITCH FACTORED
A	X	C	NGUYEN	C	✓	
B	AND	-	VAN THAC	C	✓	
C	AND	C	THANH.	N	✓	

If, after the request has been run an insufficient amount of information has been located, the request may be rerun using just the first name (tên). But to avoid a large number of false drops caused by the limited number of names, especially since the diacritical markings are dropped, some qualifier should be used, as Dầu in Bình Định province, or intelligence agent Thanh in II Corps:

REGISTER	ASSOCIATION	DESCRIPTOR TAG							FUNCTION	SWITCH PHRASED	SWITCH FACTORED
A	X	C	D	A	U	.	*	*	N	✓	
B	AND	D	2	B	I	N	H	D	N	✓	
C	OR	C	T	H	A	N	H	.	N	✓	
D	AND	A	1	5	9	*	*	*	N	✓	
E	AND	D	2	*	*	*	*	*	N	✓	

2. Family (Họ) and First Name (Tên)

If only the last and first names are known, the request should be written to retrieve all documents with these two names, whether or not the middle name (đệm) was indexed. Thus, to query for Phan Duy, Phan would be requested as a họ with the rest of the name superdigited, and Duy as a tên entered independently, ended with a period:

REGISTER	ASSOCIATION	DESCRIPTOR TAG							FUNCTION	SWITCH PHRASED	SWITCH FACTORED
A	X	C	P	H	A	N	*	*	N	✓	
B	AND	C	D	U	Y	.	*	*	N	✓	

3. Cover Names

Cover names will usually be treated as complete names. A request for information on Chín Nam should usually be written:

REGISTER	ASSOCIATION	DESCRIPTOR TAG	DESCRIPTOR	FUNCTION	SWITCH PHRASED	SWITCH FACTORED
A	~~X~~	C	CHINNAC	C	✓	
B	AND	-	M.*****	C	✓	

When both the real name and one or more cover names are known, all should be queried for independently. Thus, for Nguyễn chi Thành, with cover names Sáu (6) Di, Sáu Ri (pronounced the same in North Vietnamese) and Trường Sơn, the request would be written:

REGISTER	ASSOCIATION	DESCRIPTOR TAG	DESCRIPTOR	FUNCTION	SWITCH PHRASED	SWITCH FACTORED
A	X	C	NGUYENC	C	✓	
B	AND	-	CHITHAC	C	✓	
C	OR	C	$AURI.N	N	✓	
D	OR	C	$AUDI.N	N	✓	
E	OR	C	TRUONGC	C	✓	
F	AND	-	$ON.**	C	✓	

Should an additional cover name be required, Sáu Ri and Sáu Di might be requested in a single register:

REGISTER	ASSOCIATION	DESCRIPTOR							FUNCTION	SWITCH	
		TAG								PHRASED	FACTORED
A	X	C	$	A	U	*	I	:	N	✓	

F. DATE DESCRIPTORS

1. Requesting

Four kinds of dates can be used in indexing and requesting: the date of keypunching the document index (date of accession, /-tag), the date the document was prepared (X-tag), the date of an event described in the document (J-tag), or the date of capture (Y-tag).

EXAMPLE:

Request	Switch	Register	Association	Descriptor	Function
1. All documents input since 1 June, 1965	Factored	A		65Ø6**	L
2. Documents written since 1 June 1965	Factored	A		X65Ø6**	L
3. Documents written in 1965	Factored	A		X65****	N
4. An event that happened in 1965	Phrased	A		J65****	N
5. An event occurring in Spring 1965	Phrased	A		J65Ø4**	L
	Phrased	B	AND	J65Ø7**	L

2. Date of Information

Usually the date of an event is indexed into one J-tagged term. In unusual circumstances however, it may be desirable to describe an event occuring over a period of time. Significant time periods can be indexed by using two J-tagged terms in one phrase, corresponding to the first and last dates of the period. Events covering a time period can be recovered by requests which use limits, even when the indexed period of time is greater than that of the request.

EXAMPLE:

Consider two documents. The first concerns an event occuring on 13 July 1965; the second concerns a similar event which took place from 15 June 1965 to 30 September 1965. They are indexed:

Document 1 | J | 6 | 5 | Ø | 7 | 1 | 3 |

Document 2 | J | 6 | 5 | Ø | 6 | 1 | 5 | | J | 6 | 5 | Ø | 9 | 3 | Ø |

A request is made for a corresponding event which occurred in July or August, 1965. It can be coded using limits so that both documents are recovered:

REGISTER	ASSOCIATION	DESCRIPTOR							FUNCTION	SWITCH	
		TAG								PHRASED	FACTORED
A	X	A							N	✓	
B	AND	J	6	5	Ø	8	3	2	L	✓	
C	AND	J	6	5	Ø	7	Ø	1	L	✓	

In this case, documents are desired which concern one date or a period of time. Any date prior to the upper limit causes a hit in register B (e.g., J65Ø713, J65Ø614); register C records a hit if a date is on or after the lower limit (e.g., J65Ø713, J65Ø93Ø). Both of the above documents cause hits in registers B and C. Document 2 is recovered even though it covers a longer period of time than the request. Notice that if the lower limit were coded J65Ø7** and the upper limit J65Ø9**, the same documents would be retrieved.

3. Date of Keypunching

The date of keypunching is used in retrieval to select documents input to the system since the last time the query was run for the same person. This is particularly important in monitoring the response to collection requirements or automatic film dissemination for example, at a given time each month a query might be run to produce a microfilm of all documents added to the data base during the preceeding month on military aid granted by Communist China. To avoid duplication of documents, each month's query would specify documents keypunched during the preceeding month (in the following example December 1966):

REGISTER	ASSOCIATION	DESCRIPTOR							FUNCTION	SWITCH	
		TAG								PHRASED	FACTORED
A		A	7	1	7	*	*	*	N	✓	
B	AND	B	A	C	H	*	*	C	N	✓	
C	AND	Y	6	6	1	2	*	*	N		✓

G. PROGRAMMING REQUESTS

There are many ways to state a request and retrieve the same documents. Some requests, even though programmed differently, require the same number of descriptors. In other cases, the number of descriptors can be reduced. The different methods of programming involve the request logic or take advantage of the coding of the descriptors.

The following examples illustrate some of the techniques available.

1. If a descriptor can appear only once per document (e.g., a factored descriptor) and it is a limit, then:

AND + lower limit = AND NOT + upper limit
AND + upper limit = AND NOT + lower limit

EXAMPLE:

AND	X64****	lower limit is the same as
AND NOT	X64****	upper limit

2. Two limits can sometimes be replaced with one term containing super-digits and the normal function.

EXAMPLE:

The area from 21°30'N up to (but not including) 21°40'N and 15°50'W up to 16°00'W:

	LN0213*	N
AND	LW0155*	N

3. An OR association can sometimes be replaced by an AND or AND NOT association, thus economizing on registers. This technique makes use of the coding system.

EXAMPLE:

All raw intelligence concerning sightings made since June 1965 in the China or Yellow Sea. Sightings are indexed as a modifier (M***079). The area codes are arranged as follows:

BPCØØØ	Pacific Ocean	(not wanted)
BPCBRØ	Bering Sea	(not wanted)
BPCCHØ	China Sea	(wanted)
BPCTMØ	Tasman Sea	(not wanted)
BPCYLØ	Yellow Sea	(wanted)
BXPØØØ	Persian Gulf	(not wanted)

If the OR association is used, eight request terms (two request cards and two searches) are required:

	Card 1				Card 2		
A		J65Ø6**	L (phrased)	A		J65Ø6**	L
B	AND	M***Ø79	N (phrased)	B	AND	M***Ø79	N
C	AND	V***1**	N (Factored)	C	AND	V***1**	N
D	AND	BBPCCHØ	N (phrased)	D	AND	BBPCYLØ	N

However, the request can be programmed into six registers:

REGISTER	ASSOCIATION	DESCRIPTOR TAG							FUNCTION	SWITCH PHRASED	SWITCH FACTORED
A		J	6	5	Ø	6	*	*	L	✓	
B	AND	M	*	*	*	Ø	7	9	N	✓	
C	AND	V	*	*	*	1	*	*	N		✓
D	AND NOT	B	B	P	C	T	M	*	N	✓	
E	AND	B	B	P	C	C	H	*	L	✓	
F	AND	B	B	P	D	*	*	*	L	✓	

The limits in E and F select the codes for China Sea to Yellow Sea; the term in D excludes Tasman Sea.

H. THE HIT COUNTER

The hit counter is associated with the Output Selector, an optional feature of the FileSearch Retrieval Unit. When the Count Hits mode is selected, the number of documents matching the request is counted and displayed. The film is not stopped when a matching document is sensed, as in the other modes of operation, but the total is augmented by one.

The hit counter can be used to respond to a request, rather than documents. For example, the query "how many documents from Detachment 1 concerning communication systems, have been received in the past two months?" can be answered by a document count.

The hit counter can be used to audit the files so that collection activities can be monitored. A typical request might involve a subject and area code; the total number of matching documents would indicate whether the field reporting is adequate. Trends can also be identified by comparing the number of matching documents received during various periods of time.

The expected response to a query can be monitored with respect to the volume of output it would produce. The total may result in selecting film copy, rather than hard copy, for off-line viewing. The total may also indicate that the request should be made more specific or more general.

I. THE ROLE OF THE RETRIEVAL ANALYST

The person who is responsible for analyzing and coding requests frequently deals with a requester who is not familiar with the FileSearch system. Typically, this requester is accustomed to manual systems, to requesting a file folder and subsequently looking for more specific information. In this case, the retrieval analyst should try to help the requester to be as specific as possible. For example, he may ask for material on Albania when his real interest is in intelligence briefs concerning refineries in Albania.

When the analyst assists the requester, two benefits accrue. First, the response to the query is made more satisfactory. Second, the requester is learning about the system. As he becomes more familiar with the full potential of the system, the communication between analyst and requester improves.

Several techniques are available for refining a request. For example, by using the hit counter, the requester can be informed as to the number of documents which match his request. He can modify the request accordingly. The Browse mode is also useful in this respect. By viewing retrieved documents and perhaps their index forms, the request can be modified by adding terms. This is specifically advantageous in querying for military units. A request can initially be run for division 341. The first document hit reveals that this unit is made up of the 31st and 32nd regiments, the later with cover designations K101 and Song Hong. The query can then be written to retrieve all of these designations:

REGISTER	ASSOCIATION	DESCRIPTOR TAG	DESCRIPTOR						FUNCTION	SWITCH PHRASED	SWITCH FACTORED
A		F	D	N	3	4	1	.	N	✓	
B	OR	F	R	T	3	1	.	*	N	✓	
C	OR	F	R	T	3	2	.	*	N	✓	
D	OR	F	*	*	K	1	Ø	1	N	✓	
E	OR	F	*	*	$	O	N	G	C	✓	
F	AND	-	H	O	N	G	.	*	C	✓	

Better responses can be provided if the requester cooperates by evaluating the retrieved documents. This feed-back can be the basis for improvements in any phase of the automated document storage and retrieval system, e.g., collection, indexing procedures, indexing vocabularies, distribution and request processing. The system can dynamically progress, adapting to the changing needs of the users.

J. THE RETRIEVAL CHECK LIST

If the CONTINUED FUNCTION is being used:

1. Only registers A and B, or C and D, or E and F can be paired. Registers B and C, or D and E cannot.
2. Both registers must have the "C" function.
3. If the first register of the pair is C or E and its association is AND NOT, the second register must also be associated with AND NOT. In all other cases, the association between a pair of continued registers must be AND.

If the LIMIT FUNCTION is being used:

1. Any number of limits (one, two, three, etc.) can be used in a request.
2. Lower limits must be in registers A, C, or E; upper limits in registers B, D, or F.
3. The upper limit should be "one" greater than the largest descriptor desired.

   EXAMPLES:

   a. Names from Haack to Joyce inclusive:
      CHAACK* (lower limit) AND CJOYCF* (upper limit)

   b. Dates from July 1965 to September 1965 inclusive:
      X65Ø7** (lower limit) AND X651Ø** (upper limit) or --
      X65Ø7Ø1 (lower limit) AND X65Ø932 (upper limit).

A register's PHRASED/FACTORED SWITCH should be set according to the tag of the descriptor in that register.

1. A register's switch must never be set initially for phrased descriptors if it contains a factored descriptor and vice versa.
2. Either setting of the switches of unused registers can be chosen.

CHAPTER VII

MISCELLANEOUS PROCEDURES

A. CLEANLINESS OF FILESEARCH AREAS

The Retrieval Unit utilizes optical systems in reading and copying film. Dust in the optics is copied onto new generations of film, and can cause the film to misread index codes. Cleanliness is therefore of primary importance. The following procedures will be followed in equipment areas:

1. The machine room will be vacuumed daily. A broom will not be used.

2. Top surfaces will be wiped down daily.

3. No smoking is permitted.

4. The door will be kept closed.

5. The trash can will be emptied daily.

6. The entire area will be mopped weekly.

7. The film processor will be cleaned each time chemicals are changed.

8. The darkroom will be kept clean. Spilled chemicals will be cleaned up immediately.

B. RECORDS AND LOGS

1. Document Logs

Document logs will be maintained by the Indexing Section to keep track of the documents being input to the system. Four basic log books will be kept using the following formats:

Information Reports

Log No.	Received	Indexed	Indexer	To Mach.Rm.	Batch	Remarks
6026002766	21 JUN 67	21 JUN 67	EJC	26 JUN 67	1095	

Summaries and Translations

Log No.	Type	Received	Indexer	Indexed	To Mach.Rm	Batch	Admin.
02-2457	FT	10 MAY 67	LJS	10 MAY 67	13 MAY 67	755	23 JUN 67
06-1603	S	6 JUN 67	DVC	13 JUN 67	16 JUN 67	968	29 JUN 67

Interrogations

Number	Received	Indexer	Indexed	To Mach.Rm	Batch
492	20 MAY 67	RLP	27 MAY 67	2 JUN 67	892

Miscellaneous

Log No.	Received	Indexed	Indexer	To Mach. Rm.	Batch	Unit
541	21 JUN 67	22 JUN 67	LJS	24 JUN 67	1011	1st DIV

2. Statistical Log

A statistical log will be kept by the Indexing Section. This log will be used in compiling the monthly statistical summary. Format for the summary is shown in Figure 7-1.

3. Batch Log

The Machine Section will maintain a log of document batches. Batch numbers will be assigned by the Indexing Section. Each batch will be logged by the Machine Section when it is received, the book will be kept in batch number order. The dates when the batches are flexed and microfilmed will also be entered. The following format will be used:

Batch No.	Received	No. Doc.	Type Doc.	Flex. Op.	Date Flexed	Accession Numbers
67-1101	26 JUN 67	50	IR's	Nguyệt	28 Jun 67	6•4513 - 6•4562

Roll No.	Date Microfilmed	Date Returned	Paper Tape
127	29 JUN 67	3 JUL 67	√

DATE		SUMMARIES	TRANSLATIONS	INFORMATION REPORTS	INTERROGATION REPORTS	MISCELLANEOUS INTSUMS MSG'S	TOTALS	
	BALANCE	269	55	35	77	193	629	
	RECEIVED	325	58	26	55	110	574	574
1 - 6	TOTAL	594	113	61	132	303	1,203	
	INDEXED	403	82	46	58	0	589	589
	BALANCE	191	31	15	74	303	614	
	BALANCE	191	31	15	74	303	614	
	RECEIVED	388	279	164	229	52	1,112	1,686
7 - 13	TOTAL	579	310	179	303	355	1,726	
	INDEXED	495	80	163	133	74	945	1,534
	BALANCE	84	230	16	170	281	781	
	BALANCE	84	230	16	170	281	781	
	RECEIVED	306	459	45	136	92	1,038	2,724
14 - 20	TOTAL	390	689	61	306	373	1,819	
	INDEXED	327	57	0	148	0	532	2,066
	BALANCE	63	632	61	158	373	1,287	
	BALANCE	63	632	61	158	373	1,287	
	RECEIVED	372	49	35	121	33	610	3,334
21 - 27	TOTAL	435	681	96	279	406	1,897	
	INDEXED	365	347	18	132	74	936	3,002
	BALANCE	70	334	78	147	332	961	
	BALANCE	70	334	78	147	332	961	
	RECEIVED	129	43	0	79	16	267	3,601
28 - 31	TOTAL	199	377	78	226	348	1,228	
	INDEXED	89	40	0	60	28	217	3,219
	BALANCE	110	337	78	166	320	1,011	

Figure 7-1. Statistical Log

4. Machine Room Daily Activity Logs

The Machine Section will maintain two logs of the amount of material processed each day. They will be kept in the following formats:

Flexing Log

Date	1 JUL 1967	2 JUL 1967	3 JUL 1967	4 JUL 1967
On Hand	445	455	476	252
Docs Rec'd	313	337	38	---
Total	758	792	514	252
Docs Flexed	303	316	262	208
BALANCE	455	476	252	44

Record Unit Log

Date	1 JUL 1967	2 JUL 1967	3 JUL 1967	4 JUL 1967
Backlog	228	303	267	373
Docs Flexed	303	316	262	208
Total	531	619	529	581
Docs Recorded	228	352	156	118
	303	267	373	463

5. Queries Log

The Machine Section will maintain a log of the queries levied against the system, to include the requestor's name and organization. The following format will be used:

Query Nº	Organization	Date	Nº Queries	Nº Ans.	16mm	Nº Cartridges	Date To Ops
974	500th MI GP	29-1-68	1	1	✓	2	29-1-68
975	1st Inf.Div	30-1-68	4	3			30-1-68

The columns marked "Nº Q" and "Nº R" will be used for the number of queries and the number of those queries which produced responses. The number of queries will be considered to be the number of independent questions asked by the requestor.

6. Monthly IDHS Statistical Worksheet

At the end of each week, the statistics for all operations in the system will be entered onto the Monthly Statistical Worksheet. These data will be extracted from the logs listed earlier in this section. The first and fifth weeks for the month may each be partial weeks, with correspondingly smaller figures. A sample Monthly Statistical Worksheet is shown in Figure 7-2. The top figures given in the large blocks that have been subdivided are the running totals.

1-6 7-13 14-20 21-27 28-31

FMA STATISTICAL WORK SHEET FOR THE MONTH OF: January 68

INPUT — Week	1st	2nd	3rd	4th	5th	Total
Number of Documents Received:	574	1686 / 1112	2724 / 1038	3334 / 610	3601 / 267	3601
Number of Documents Indexed:	589	1534 / 945	2066 / 532	3002 / 936	3219 / 217	3219
Number of Documents Backlogged:	614	781	1287	961	1011	1011

Number of Documents Received:	397	1485 / 1088	2129 / 644	2981 / 852	3303 / 322	3303
Number of Documents Keypunched:	602	1927 / 1325	3129 / 1202	3953 / 824	4037 / 84	4037
Number of Documents Backlogged:	1469	1232	674	702	940	940

Number of Documents Received:	602	1927 / 1325	3129 / 1202	3953 / 824	4037 / 84	4037
Number of Documents Recorded:	809	1770 / 961	3886 / 2116	5009 / 1123	5551 / 542	5551
Number of Documents Backlogged:	1308	1672	757	458	0	0

OUTPUT

Number of Queries Received:	11	18 / 7	19 / 1	23 / 4	28 / 5	28
Number of Queries Answered:	11	18 / 7	19 / 1	23 / 4	28 / 5	28

Figure 7-2. Monthly Statistical Worksheet

APPENDIX A

GLOSSARY

1:1	A ratio of unity; used in optics to indicate the same size rather than a reduced or magnified size.
Association	One of three logical connections (AND, OR and NOT) which can be made between the request terms, relating the hits recorded by the individual registers in the Retrieval Unit; punched respectively as &, Ⓡ, - into the request card.
Beginning-of-document Character	A control character (punched as an arrow) which appears as the first character of a document's indexing; controls the Retrieval Unit to reset hits, page counts, etc. during search; not used in a request.
Bit	A Binary digit - that is, a digit from a numbering scheme which has only two: 1 and 0. The two digits correspond to machine-readable states such as hole/no hole; opaque area/translucent area; on/off.
Binary Representation	The configuration of bits (holes, opaque areas) which represent a character to the FileSearch equipment. Seven bits are used per character; the first six represent data and the seventh, right-hand bit is used for checking. The representations have been chosen to maintain numeric and alphabetic order:

FILESEARCH BINARY REPRESENTATION

DATA BITS*

higher order / lower order	00	01	10	11
0000	0	space		
0001	1	A	J	
0010	2	B	K	S
0011	3	C	L	T
0100	4	D	M	U
0101	5	E	N	V
0110	6	F	O	W
0111	7	G	P	X
1000	8	H	Q	Y
1001	9	I	R	Z
1010			→	
1011	-	OR		stop
1100				*
1101	/			
1110			CR	tab
1111	.		AND	delete

*The check (parity) bit is chosen so that each binary representation has an odd number of one bits.

Carriage Return Character — A control character punched by the Flexowriter when the carriage return key is depressed, thereby returning the carriage of the Flexowriter to the left-hand margin stop position; an end-of-word character; CR.

Characters — The letters, numbers or symbols used in the FileSearch system. Forty-eight characters are available on the Flexowriter, including the period, dash, etc. Each is represented by a unique 7-bit code on both the punched card and film.

Code Block	On one frame, the area designated by the frame marker which contains the binary index code (up to 56 characters or 392 bits).
Control Character	A character which is used to control an operation of the Retrieval Unit.
Continued Word	The second of two words used to index or request a term which has more than six characters (excluding the tag).
Descriptor	A descriptor is a descriptive term; an index or request term. On the film, a descriptor can have seven or less characters, and as many descriptors as needed can be used to describe a document. In the request, a descriptor has seven characters and is held in one of the six registers of the Retrieval Unit.
Document	Any recorded information, regardless of its physical form, is a document. The FileSearch system excludes sound recordings, but handles documents which are written, typed, printed, drawn; graphics; photographs, etc.; and which can have one or more pages.
End-of-word Character	A control character at the end of a descriptor on the indexed film which signals the Retrieval Unit to store any hits and begin comparing a new descriptor; both the tab and carriage return are end-of-word characters.
Extended Word	See Continued Word.
Factored Descriptor	A descriptor assigned to the factored field because it can be combined with any of the other terms in a document's index.
Factored Field	The first of two fields of variable length in a document's coded index; contains the factored, rather than phrased, descriptors.
False Drop	An irrelevant document retrieved in response to a request.
FETM	An abbreviation for the File Expansion Transport Module which contains unprocessed film and is part of the 1:1 camera unit; sometimes referred to as FEM.

File	This term is used in the FileSearch system to denote any collection of documents (library) stored on coded film.
File Expansion	The automatic structuring of the film files into categories to reduce access time to requested documents; analogous to manually cross-filing copies of documents.
File Item	A unit in a file; a document.
File Maintenance	In the FileSearch system, the process of creating and updating the film file; includes processing, verifying, distributing and structuring the coded film.
Fixed Field	A term used to denote the first eight descriptors on each Minicard (an Eastman Kodak system) roughly equivalent to FileSearch's factored field.
Flag	See Timing Mark.
Flexowriter	An electric typewriter used in the FileSearch system which can punch index or request cards and/or read the cards; uses the FileSearch binary configuration to represent characters; can also punch and/or read paper tape.
Forced Hit	A hit forced into a register by means of a switch on the console of the Retrieval Unit; used to broaden a request by cancelling the corresponding descriptor from the request.
Frame	In the FileSearch system, a length of 35mm film (0.375 inches) which has two sets of sprocket holes and can be blank or can contain an image, code, or both image and code.
Frame Marker	A small vertical strip adjacent to the code block in each frame of film containing image and/or code. The strip is slightly longer than the code block, extending beyond it at both top and bottom. It tells the Retrieval Unit's film scanner that it is about to read a block of code, and finally indicates that it has finished reading the block of code.

Function	Each request register can function (make comparisons between the code on the request card and the code on the film) in three ways. It can make NORMAL (equal) comparisons, LIMIT (greater than or less than) comparisons, and CONTINUED comparisons (in which two registers are connected and a comparison made for equality on adjacent descriptors read from film).
Hit	A match: (1) A hit occurs when a coded request descriptor and a coded index descriptor on the film compare in a specified way; the corresponding register's indicator light is turned on; (2) A hit occurs when a document's indexing satisfies the request; the document is retrieved or counted.
Hit Counter	An optional Retrieval Unit feature which can be used to count and display a running total of the number of documents which match a request.
Index Card	A specially printed card which contains the indexing information about a document; punched by the Flexowriter.
Index Term	A descriptor used to index a document.
Indexing	The process of analyzing a document to determine its information content: identifying and evaluating the concepts discussed in the document; describing important information and the document with index terms, according to the procedures set forth for the system.
Jimal Switches	A series of switches for registers C, D, E and F that link each of these registers directly with register A, without regard for intervening registers.
Library	Any collection of documents; often referred to as a file.
Open Field	A term used to denote the field following the fixed field on each Minicard (an Eastman Kodak system) roughly equivalent to FileSearch's phrased field.
Parity Error	A machine-readable character which contains an even (rather than odd) number of 1-bits is said to have a parity error.

Phrase	A group of related index terms demarcated by a phrase boundary; used to prevent false drops caused by the association of unrelated index terms.
Phrase Boundary	The control codes (tab-Q-tab) which appear in the indexing at the beginning and end of a phrase and can be recognized by the Retrieval Unit (an optional feature); carriage return characters can be substituted for the tabs.
Phrased Descriptor	A descriptor assigned to the phrased field because it might be indexed into a phrase; in any one indexing system, the assignment is fixed according to the kind of descriptor (tag).
Phrased Field	The second of two variable length fields in a document's coded index which contains the phrased, rather than factored, descriptors.
Phrase Gate	A phrase boundary.
Q-tag	See Phrase Boundary.
Recording Unit	In the FileSearch system, a specially designed camera which records both a document and its associated index code on 35mm microfilm; Record Unit.
Register	One of six storages in the Retrieval Unit which holds a request term for comparison during search; the request term is entered by means of the punched request card or optional request console.
Request Card	A specially printed card which contains the request information: the request terms, register functions and associations; punched on the Flexowriter.
Retrieval	In the FileSearch system, the recovery of a requested document in the form of an image on a screen, a hard copy, a duplicate film copy or a count.

Retrieval Unit	In the FileSearch system, the equipment which can search the coded film and retrieve requested documents.
Search	In the FileSearch system, the process of locating requested documents on the Retrieval Unit by moving the coded film in the forward direction; scanning, comparing, and associating index terms to determine if a document should be retrieved; positioning a matching document for retrieval.
Stop Code	A control character which: 1) On the film, tells the Retrieval Unit that it has come to the end of a code block; any code block containing less than 56 characters must have a stop code as the last character. 2) On a request card, indicates unused registers and must be entered as the association character of the first register not used. 3) On a punched card, halts the reading operation of the Flexowriter. The stop code is the last character punched into any card, request or index.
Superdigit	A control character, punched as an asterisk (*) in a request term, which causes comparison results of the corresponding character in the index terms scanned from film to be ignored. Any character (letter, number, symbol) occupying a position within a descriptor on film corresponding to that of a superdigit in a request descriptor is considered as matching. The superdigit is used to fill out unused positions of registers and to make categorical searches.
Tab	A control character punched by the Flexowriter when the tab key is depressed, thereby moving the carriage of the Flexowriter to the next tab stop position; an end-of-word character.
Tab Error	If more than seven index characters are scanned from film without sensing an end-of-word control character (tab or carriage return) a tab error is said to exist.
Tag	The first character of a descriptor which identifies the kind of descriptor; during search, a tag is always compared for equality.

APPENDIX B

ADMINISTRATIVE DIVISIONS

OF

SOUTHEAST ASIA

INTRODUCTORY STATEMENT

These listings of the administrative divisions of the countries of Southeast Asia have been compiled to assist in studies on the area. Names given in Roman type are the divisions used by the governments at the time of compilation. Italic type is used for former names and for the names of disestablished divisions.

ADMINISTRATIVE DIVISIONS OF THE REPUBLIC OF VIỆT NAM

The Republic of Việt Nam is divided into provinces (*tỉnh*), which are further subdivided into districts (*quận*). Five cities (Cam Ranh, Đà Lạt, Đà Nẵng, Huế and Vũng Tàu) and the Sài Gòn prefecture are autonomous and do not fall within the provinces. The Côn Sơn islands are administered under Sài Gòn. In this geographic listing the provinces have been arranged from north to south.

A five-digit area code is assigned to each administrative division, except the autonomous cities. The first digit (1-5) represents the corps tactical zone in which the province is located; "5" represents the Sài Gòn-Gia Định capital region. The second and third positions contain the province number, assigned from north to south. The last two positions represent specific districts.

Names for the administrative divisions used by the Việt Cộng are included, in italic type, along with the approximate correspondence of GVN areas.

(This Page Intentionally Left Blank)

APPENDIX B

ADMINISTRATIVE DIVISIONS OF THE REPUBLIC OF VIET NAM

(SOUTH VIET NAM)

GEOGRAPHICAL LISTING

10100 Quảng Trị province

10108 Ba Lòng district

10101 Cam Lộ district

10102 Gio Linh district

10103 Hải Lăng district

10104 Hương Hóa district (*VC Hương Hóa district includes Trung Lương district.*

10104 *Hy Lập district, VC district in the southern part of Hương Hóa district*

10105 Mai Lĩnh district (*Created by decree 880 on 11 June 1965.*)

10104 *Mông Cổ district, VC district in the southern part of Hương Hóa district*

10106 Triệu Phong district

10107 Trung Lương district (*Included in VC Hương Hóa district.*

6000 *Vĩnh Linh district, variant name of Vĩnh Linh region, North Việt Nam*

10200 Thừa Thiên province

10201 Hương Điền district

10202 Hương Thủy district

10203 Hương Trà district

10204 Nam Hòa district (*Nam Hòa is divided into three areas* [vùng] *by the VC, designated A, B and C.*)

10205 Phong Điền district

10206 Phú Lộc district

10207 Phú Thứ district (*Included in VC Phú Vang district.*)

10208 Phú Vang district (*VC Phú Vang district includes Phú Thứ and Vĩnh Lộc districts.*)

10209 Quảng Điền district

10210 Vĩnh Lộc district (*Included in VC Phú Vang district.*)

10300 Quảng Nam province (*VC Quảng Đà province. The VC use the name Quảng Nam for GVN Quảng Tín province.*)

10301 Đại Lộc district

10302 Điện Bàn district (*VC Điện Bàn district includes Hiếu Nhơn district.*)

10303 Đức Dục district (*Divided by the VC among Đại Lộc, Duy Xuyên and Quế Sơn districts.*)

10304 Duy Xuyên district

10305 Hiếu Đức district

10306 Hiếu Nhơn district (*Included in VC Điện Bàn district.*)

10307 Hòa Vang district

10308 Quế Sơn district

10309 Thương Đức district

10400 Quảng Tín province (*VC Quảng Nam province*)

10401 Hậu Đức district (*VC Trà My 2 district*)

10402 Hiệp Đức district

10403 Lý Tín district (*Called Tam Kỳ Nam by the VC.*)

10402 *Phước Sơn district, VC district covering northern Hiệp Đức district*

10404 Tam Kỳ district (*Called Tam Kỳ Bắc by the VC.*)

10405 Thanh Bình district

10406 Tiên Phước district

10402 *Trà My 1 district, VC district covering the southern part of Hiệp Đức district*

10401 *Trà My 2 district, VC district corresponding to Hậu Đức district.*

10500 Quảng Ngãi province

10501 Ba Tơ district

10502 Bình Sơn district

10503 Đức Phổ district

10501 *Gia Vut district, VC district covering half of Ba Tơ district, probably created in early 1967*

10504 Minh Long district

10505 Mộ Đức district

10506 Nghĩa Hạnh district

10507 Sơn Hà district (*The VC consider the western part of Sơn Hà as part of Kontum province.*)

10508 Sơn Tịnh district

10509 Trà Bồng district

10510 Tư Nghĩa district

20600 Bình Định province

20605 *An Lão district, VC district corresponding roughly to western Hoài Nhơn district*

20601 An Nhơn district

20602 An Tức district (*Included in VC Gia Lai province.*)

20603 Bình Khê district

20604 Hoài Ân district

20605 Hoài Nhơn district (*VC Hoài Nhơn district covers only the eastern part of the GVN district.*)

20606 Phù Cát district

20607 Phù Mỹ district

20608 Tuy Phước district (*VC Tuy Phước district covers only the eastern part of the GVN district.*)