

衣服設計用データベースの評価（第2報）： 検索評価の予備実験

著者	田中 早苗, 赤見 仁
雑誌名	東京家政大学研究紀要 2 自然科学
巻	34
ページ	111-116
発行年	1994
出版者	東京家政大学
URL	http://id.nii.ac.jp/1653/00010546/

衣服設計用データベースの評価(第2報) 検索評価の予備実験

田中早苗*, 赤見 仁**

(平成5年9月30日受理)

Evaluation of a Database System for Clothing Design (Part 2) Preliminary Test for Evaluation of Retrieval

Sanae TANAKA* and Hitoshi AKAMI**

(Received September 30, 1993)

1. 緒 言

データベースの評価は一般に次のふたつの側面から行われている。ひとつはシステムの利用者によって判断される検索結果に対する満足度で、もう一方はデータベースの設計者による機能評価である。機能評価では、検索された適合件数と対象とするファイルの全適合件数による再現率 (recall ratio) や、検索結果の全件数の中で検索目的に合致した情報の適合率 (precision ratio) を尺度として検索精度を調べる方法がある。¹⁾ 本研究では後者の立場から、データベースに入力されているキーワードと実際の検索でユーザーが指定するキーワードとの的中率 (hit ratio) を尺度とし、SKIRT3.DBF²⁾ の検索評価実験を試作的に行った。実験では検索条件の論理演算子を全て乗算の形式にし、メニュー方式で検索を行うプログラムを作成して用いた。実験の結果フィールドごとのキーワードの適合性が表され、本検索評価法が有効であると確認できたので報告する。

2. SKIRT3.DBF の改良

2.1 メニュー式検索プログラムの導入

検索実験ではオペレータによる作為的な検索条件の偏りや操作ミスをできるだけ排除し、被験者を時間的に拘束する負担を軽減するためのプログラムを作成した。

これは画面の中にキーワードを当てはめていくメニュー方法で、キーワードをタイプせずに表示された中から選択して検索条件を設定する。したがってユーザー自身のキー操作による検索の実行が可能である。また、プログ

* 服飾美術科 被服構成学実験研究室

** 服飾美術学科 被服構成システム研究室

ラムのユーザーヘルプ機能(操作方法やメニューの説明をする画面)を利用して実験の進行を補助する初期画面や各フィールドごとにキーワードを参照できる HELP 画面を作成した。キーワードはアルファベット順に並べ、30字程度の説明を付けた。

2.2 キーワードとデータ構造の変更

SKIRT3.DBF の収録データは、既定の雑誌で'93年7月までに掲載された495資料を加えて2389レコードに更新した。前報²⁾の結果から変更したキーワードは NAME1 と LENGTH のフィールドにあり、NAME1 ではスカート名を5つの名称に統合し、LENGTH は長さの段階分けを3段階から5段階に分化した。またデータ構造は、メニュー式検索に伴う次のような制約から変更した。

(1) 検索条件を設定するとき、各フィールドは1回しか使うことができない。また、1つのフィールドでは1つのキーワードしか選択できない。

(2) キーワードは各フィールドに固有でなければならない。同じキーワードが他のフィールドの別のレコードにあった場合、そのキーワードをもつ全てのレコードを検索することはできない。(これは3.2で後述する。)

(3) 各フィールドのキーワード数はユーザーが選択可能な数量でなければならない。

これらのことから、フィールド NAME1 では出現頻度数が多く他のスカート名との重複が多いジャンパースカートとラップスカートを NAME2 に移行した。DETAIL はフィールドをもうひとつ設けて (DETAIL 3) 出現頻度の多い順にキーワードを分配した。その他のフィールドにおいても出現頻度数が10に満たないキーワードは 'another' をキーワードの頭に付けて、another で検索

できるようにするなどの統合を行った。SKIRT3.DBFのキーワードを表1に示した。

3. 方 法

3.1 実験方法

実験は検索する情報をあらかじめ1件設定し、それに適するキーワードを被験者に選択させ検索したとき、キーワードと入力されているキーワード（以後データと記す）との合致性を調べる方法で行った。

設定情報はSKIRT3.DBFに収録されているスカート資料2389枚から被験者に2枚ずつ選択させ、合計20枚について1回ずつ検索を行った。対照実験として同資料から無作為に3枚（NO.2348,2152,1953）抽出し、各々を10名または5名の被験者で繰り返し検索を行った。

被験者は本学で被服を学ぶ学生で、被験者1名の検索件数は3件、所要時間は資料選択から検索過程のプリン

トアウトを含め約20分を予定した。

実験では、データの入出力と検索処理を行う通常のDBMSの機能を備えたパソコン装置一式（以後パソコンAとする）と、HELP画面でキーワードをディスプレイするパソコン（これをパソコンBとする）の2台を用いた。これは使用したアプリケーションのHELP機能が本来メニュー説明に使われるのでメニューをオープン（検索条件の入力画面に）した後にHELPをオープンするにはメニューに戻らなくてはならない。つまり入力画面を度々クローズしなければならないという難点を補うためである。パソコンAには検索条件と結果をプリントアウトするプリンターを接続し、パソコンBは出来るだけパソコンAのディスプレイの近くで参照できるようにラップトップ式を用いた。パソコンAの操作は主に被験者自身が行なった。被験者はカーソルキーでRECORD検索、LIST出力、RECORDCOUNT

表1 SKIRT3.DBFのキーワード

[NAME1]	[SILHOUETTE]	[BUTTON]	[BOTTOM]	[DETAIL3]
1. CULOTTES.	1. A-LINE	1. -	1. FRILL	1. DREAP
2. FLARES.	2. BALLEL	2. B.	2. GUSSET	2. EMBROIDERY
3. GATHERS.	3. BALOON	3. F.	3. ILLEGULAR	3. FASTENER
4. PLEATS.	4. BELL	4. S.	4. PLEAT	4. FRILL
5. TIGHTS.	5. BIG	5. another	5. SCALLOP	5. FRINGE
	6. CIRCULAR	6. B. (W)	6. SLIT	6. LACE
[NAME2]	7. DIRNDLE	7. F. (W)	7. STITCH	7. LACEUP
1. ACCORDIONPLEATS.	8. DOME	8. S. (W)	8. TRIMMING	8. PANEL
2. BALOONS.	9. PEGGED	9. B. BUTTNDOWN	9. VENT	9. TRIMMING
3. BOXPLEATS.	10. MERMAID	10. F. BUTTNDOWN	10. another	10. V-CUT
4. CIRCULARS.	11. SEATH	11. S. BUTTNDOWN		11. another
5. FLARES.	12. SEMITIGHT	12. B. POCKET	[DETAIL1]	12. HALTERNECK
6. FLAREGATHERS.	13. STRAIGHT	13. F. POCKET	1. DART	13. LOWNECK
7. GATHERS.		14. S. POCKET	2. FLARE	14. SURPLICNECK
8. GORED.	[WAIST]	15. SUSPENDER	3. GATHER	15. ROUNDNECK
9. INVERTEDPLEATS.	1. -		4. GORE	16. SQUARENECK
10. JUMPERS.	2. BELTLESS	[POCKET]	5. PLEAT	17. U-NECK
11. ONEWAYPLEATS.	3. DRAWSTRING	1. -	6. TIER	18. V-NECK
12. OVERS.	4. ELASTIC	2. CROSS	7. TUCK	19. FLATCOLLAR
13. PANELS.	6. HIGHBELT	3. FLAP		20. TALLEOREDCOLLAR
14. SEMIFLARES.	7. HIGHYOKE	4. KANGALOO	[DETAIL2]	
15. SEMITIGHTS.	8. HIPBONE	5. PATCH	1. RIBBON	[LENGTH]
16. SUSPENDERS.	9. INBELT	6. PIPING	2. SALOPETTE	1. MINI
17. TIEREDS.	10. SHIRRING	7. ROUND	3. SEAMDESIGN	2. SHORT
18. TUCKS.		8. SEAM	4. SUSPENDER	3. KENN
19. WRAPS.		9. SIDE	5. STITCH	4. MIDI
20. YOKES.		10. WELT	6. WRAP	5. LONG
		11. anoth	7. YOKE	
			8. BACKOPEN	
			9. FRONTOPEN	
			10. SIDEOPEN	

の3つのメニューを選択し、任意のフィールド位置でスペースキーを押し続けると次々とキーワードが現われ、リターンキーで実行となる、3つのキーで操作した。操作方法はDB設計者が被験者の側でアシストした。DB設計者は、被験者が回答したデザインの特徴と関連のあるフィールドのHELP画面をパソコンBにディスプレイし、検索過程のプリントアウトを行った。プリントアウトは、①検索条件が入力された状態と、検索条件に該当するレコードが存在した(‘検索条件 [有]’と画面右上に表示された)場合の件数を示す②RECORDCOUNT、そして核当したレコードナンバーを表示する③LISTの画面をハードコピーして記録とした。図1に検索実験の手順を示した。

通なキーワードが存在したとき、そのキーワードのあるレコード全てを検索することができる。しかし、本実験では検索目的を1件と定め、ユーザーが選んだキーワードとデータとの一致性を調べるものであるから検索結果を不必要に増やすべきではない。よって、 $A \times B \times C$ の条件式を用いた。

3.3 分析方法

被験者が設定した検索条件のHARDCOPYと検索目的のデータリストから、検索に使ったキーワードとインプットデータが一致していた、すなわち的中していたフィールド数を調べ、選択した20枚と繰り返し検索の3枚それぞれのフィールドごとに的中率(Hit ratio)を計算し、全フィールドの平均値をその資料集合の的中率とした。

$$\text{的中率(\%)} = \frac{\text{一致したフィールド数}}{\text{使用したフィールド数}} \times 100$$

4. 結果と考察

1回の検索に使われたフィールド数は、最高7フィールド、最も少なくても3フィールドで、平均4.5フィールドであった。選択した20枚の結果から、グラフで突出したNAME1とLENGTHを除いて、およそどのフィールドも平均的に使われていた。図3に検索に使ったフィールド数(input)と的中したフィールド数(hit)を、また表2に的中率を示した。

繰り返し検索のNO.2348は当初10回の検索を行った

表2 キーワードの的中率

Field	choice 20	Hit ratio(%)		
		NO. 2348	NO. 2152	NO. 1953
NAME1	89.5	80.0	100.0	100.0
NAME2	66.6	80.0	0.0	-
SILHOUETTE	75.0	100.0	0.0	25.0
WAIST	62.5	-	0.0	66.6
BELT	37.5	86.0	100.0	80.0
BUTTON	50.0	77.0	-	-
POCKET	83.3	50.0	0.0	100.0
BOTTOM	80.0	0.0	-	100.0
DETAIL1	100.0	0.0	-	100.0
DETAIL2	50.0	33.0	100.0	-
DETAIL3	100.0	0.0	-	-
LENGTH	61.5	16.0	50.0	100.0
Mean	72.4	63.0	47.8	77.2

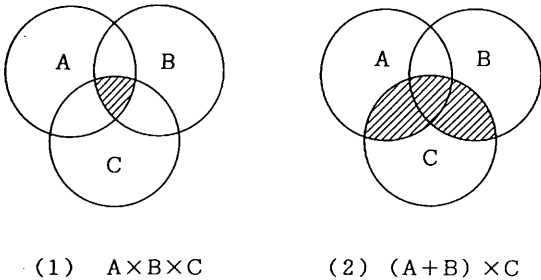


図2 検索条件の集合関係

3.2 プログラムの条件式

プログラムはDBMSの仕様に従って作成した。検索コマンドは論理演算子(2つ以上の条件を組み合わせる接続部分)を全て乗算にしたので検索結果の件数はかなり制限されるものと予測した。これは検索の条件をA, B, Cの3つであるとすれば $A \times B \times C$ であり、ベン図で表すと図2(1)のようになる。³⁾スカート为例に、条件をA:フレアスカート, B:吊りのついた, C:前にボタンがあるとすれば、‘前にボタンのある吊り付きのフレアスカート’が検索結果となる。しかし通常の検索では図2(2)の $(A + B) \times C$ の検索式なども使われ、こちらは‘フレアスカートかあるいは吊り付きのスカートで前にボタンが付いているもの’であるから、前にボタンさえ付いていれば吊りのないフレアスカートも出力される。この場合1つのコマンドで同じフィールドを何回も使うことができ、またいくつかのフィールドに共

衣服設計用データベースの評価(第2報) 検索評価の予備実験

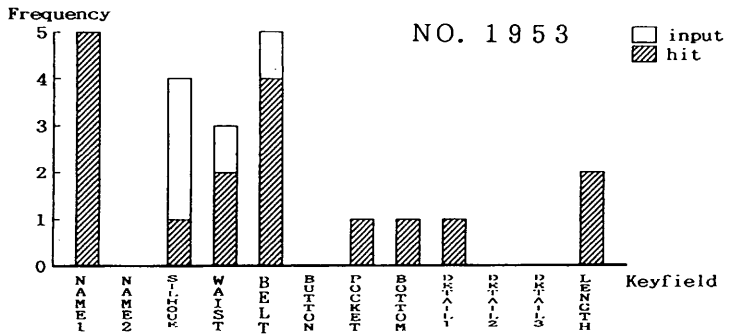
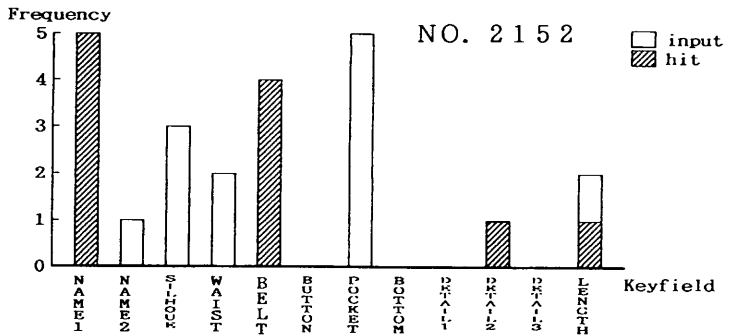
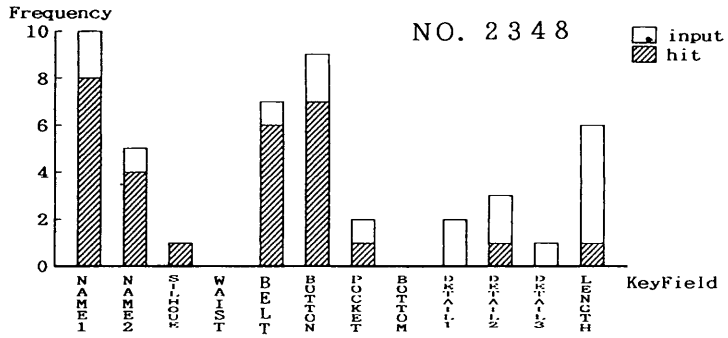
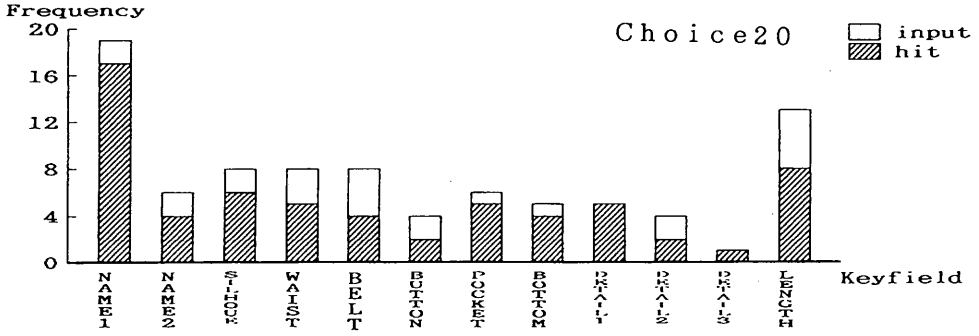


図3 検索に用いたフィールド数

が、5回までの平均的中率が61.0%、後の5回までが64.2%であった。このことから、繰り返し5回の検索で資料的中率が得られると考えた。繰り返し検索では、資料のデザインによって当然使われるフィールドも異なる。的中率50%を目処に、これを下るフィールドのキーワードは被験者の認識違いも考えられるが、認識違いを招きやすいキーワードであるかまたは不適切であることが明らかである。例えばNO.2348のLENGTH='MIDI',NO.2152のSILHOUETTE='A-LINE',WAIST='HIGH',POCKET='CROSS',NO.1953のSILHOUETTE='SEATH'などである。検索の際被験者は、デザインの特徴を見だしにくい資料の場合、LENGTHをキーワードに選ぶ傾向がある。これは前回の調査でLENGTHを表現したことばの出現頻度数が多かったことに一致するが、今回スカート長さを5段階に分けてキーワードを入力する際、筆者らは型紙に印された実寸法から段階分けしたのに対し、被験者はスタイル画のスカート丈から判断していた。この実寸法とスタイル画の長さ表現のくい違いが的中率を低くしたものと考えられる。選択した20枚の結果から、的中率が80%以上のNAME1,POCKET,BOTTOMのフィールドは適切なデータ設定

であったと言える。

5. 総括

試行した検索評価実験はSKIRT3.DBFのフィールドごとに入力されたデータの適合性を調べるために有効であった。今後、平均的中率が80%以上であるようなデータ設定にするためには、多くの資料でこの評価実験を繰り返し行い、データ更新することによって実現可能である。

謝辞

本研究にご協力下さいました徐廷権先生と今村純子さんそして本学の学生に心から感謝いたします。

文献

- 1) 藤田節子：データベース設計入門，日外アソシエーツ，（東京）1992付録P.12,13,18
- 2) 田中早苗，赤見仁：衣服設計用データベースの評価，東京家政大学研紀要，33,83-90（1993）
- 3) 三輪真木子：サーチャーの時代，丸善（東京）1992 P.73