

# データビジュアライゼーション

加藤 博己, 伏見 俊彦

## 1. まえがき

企業活動には、いわゆる PDCA (Plan, Do, Check, Action) と呼ばれるビジネスサイクルがある。まず、企業や部門のビジネスゴールに向けた戦略を策定 (Plan) し、日々の業務の中で絶えず施行 (Do)、検証 (Check) を繰り返すのである。一般に、データマイニングはトップダウンによる戦略策定過程を支援するものである。商品戦略、販売戦略、あるいは投資計画を策定するために、データマイニングを活用し、時として大きな効果を上げてきている。しかし、市場や社会の変化は、さらに早く、複雑になってきている。苦労して策定したきわめて効果的な販売戦略もすぐに他社に追従されてしまうのである。このような、状況の中で、近年、現場でのタイムリーな分析的活用が重要視され始めている。

OLAP ツールは、その容易な操作性と運用性で、現場での情報の管理/分析ツールとして定着し始めている。OLAP は、一般には多次元表の形で、データを種々の切り口で加工/整理しながら情報を検証するための道具である。OLAP のドリリングやスライス&ダイスといった操作は、人間の仮説検証型の思考パターンに馴染みやすい。OLAP は現場の業務 (ワークフロー) に密着した帳票イメージで情報を整理し、分析的活用を実現するものである。そのため、現場での非専門家にとって受け入れやすい分析ツールであろう。

しかし、一方、「人」「モノ」「金」が流れるビジネスの世界では、通常、100次元の管理すべき切り口があるとされている。年度別/期別/月別といった時間の切り口、店舗別/支店別/地域別といった場所の

切り口など、非常に多くの切り口のバリエーションでビジネスが管理される。したがって、日々の業務の中では、より多くの切り口を効率的にレイアウトした複雑な帳票が一般的に用いられる。この複雑な帳票から問題点を適切に把握し、ビジネスの意思決定に結び付けるには、経験と熟練が不可欠である。

そのため、データを自由な切り口で加工し、集約された情報としてとりだすことのできる OLAP ツールとともに、情報を適確かつ容易に把握し、新しい発想を促すことのできるデータビジュアライゼーションの技術が重要になっている。

本稿では、データビジュアライゼーションの技術の 1 つである、平行座標 (Parallel Coordinates) に着目し、多次元分析における有用性を説明する。

## 2. 平行座標とは

平行座標 (Parallel Coordinates) とは、複数の切り口 (一般に軸あるいは次元と呼ばれる) を平行に並べ、各データを連続したひとつながりの折れ線で表現したグラフィカル表現である。平行座標では、一度に複数の次元の情報を表現することができ、データ全体の概観を直感的に捉えることができる。

平行座標上では、OLAP による多次元分析と同様に、いくつかの分析手法を適用することができる。主な分析手法として、以下の 3 つがある。

### 1) フィルタリング

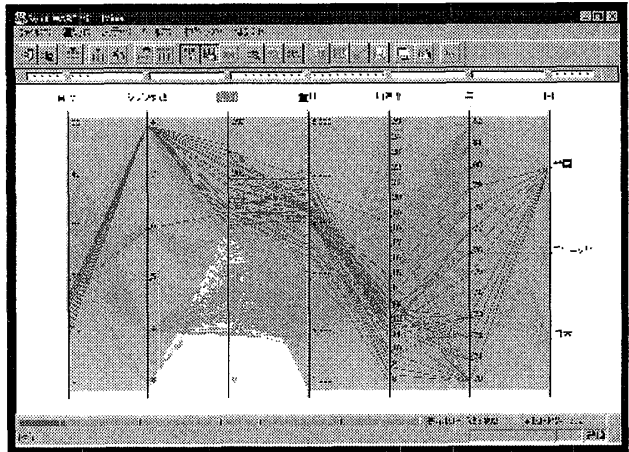
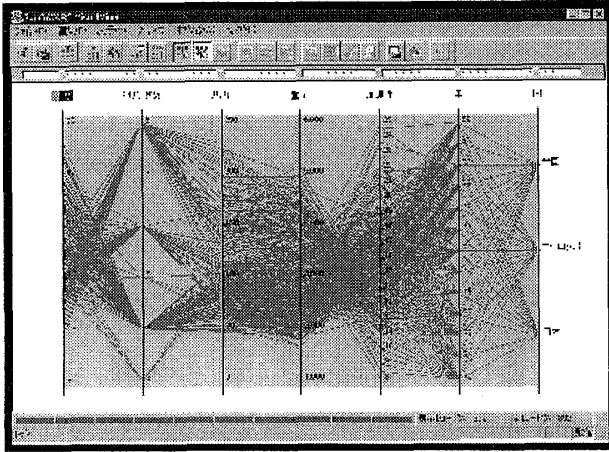
軸ごとに着目すべき範囲を限定する操作である。

これによって、特定の軸に対する範囲限定を所与の条件として、不特定任意の軸間の関連性の強さを視覚的に捉えることができる。図 1 にフィルタリングの操作例を示す。

### 2) ズーミング

特定のデータの情報だけを拡大する操作である。

かとう ひろみ, ふしみ としひこ  
株式会社富士通ソフトウェア生産技術研究所  
〒250-0011 小田原市栄町 1-6-4



馬力の高い部分にフィルタリングすると、その車の特長が浮き彫りになります。

図1 フィルタリング

これによって、データが集中している部分だけを拡大して分析することや特異データを除去して分析することが容易になる。ズームは、分析者にとっての“ノイズ”を除去し、情報の浄化（クレンジング）を可能にする。図2にズームの操作例を示す。

### 3) マグニファイ

一般的に、ビジネス分野におけるデータは多くのカテゴリによって、もともと要約されている場合が多い。そのため、視覚的な情報量が小さく、特徴的な情報を得ることができない。

マグニファイ操作は、粒度の高い任意の軸によって、カテゴリカルな軸の視覚的な情報量を大きくすることによって、隠れた情報をひきだすことができる。図3にマグニファイの操作例を示す。

## 3. ビジュアライジング OLAP

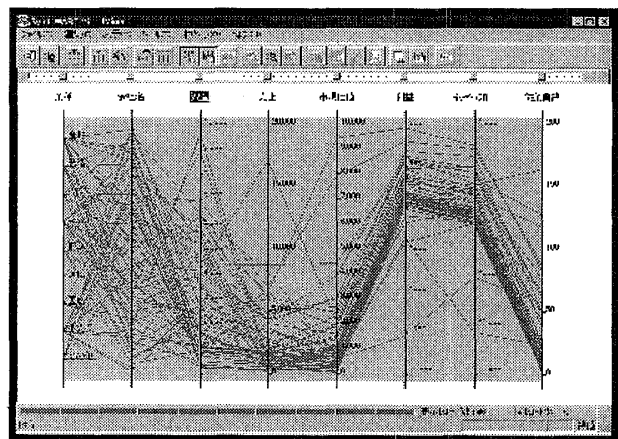
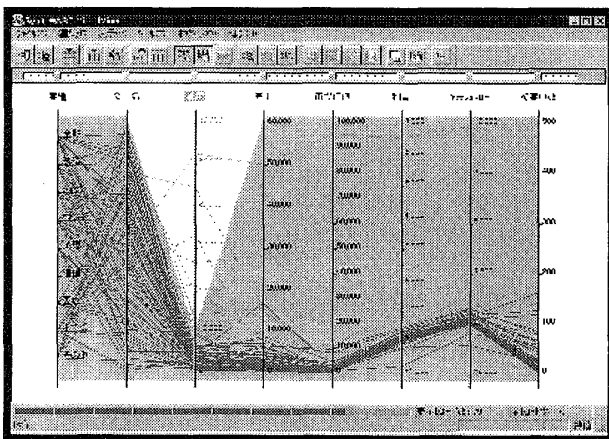
ここでは、ビジュアルに表現した多次元データに対

し、種々の操作を行うことによって、傾向や特性を直感的に把握できるビジュアライジング OLAP について説明する。

一般に、データに埋もれた問題点を明確にしたり、その原因を探るためには、広い視野から問題点を概観し、注目すべき点を探したりすることが必要となる。多くの場合、ここで人の発想や洞察が重要になる。発想や洞察は、直感的な理解と連想から生まれる。したがって、多次元にわたる情報全体を直感的に把握し、着目すべき点を即座に掘り下げたり、別の着目点に即座に移動したりする、いわゆる、フライスルーが可能なビジュアルな操作イメージが必要となる。

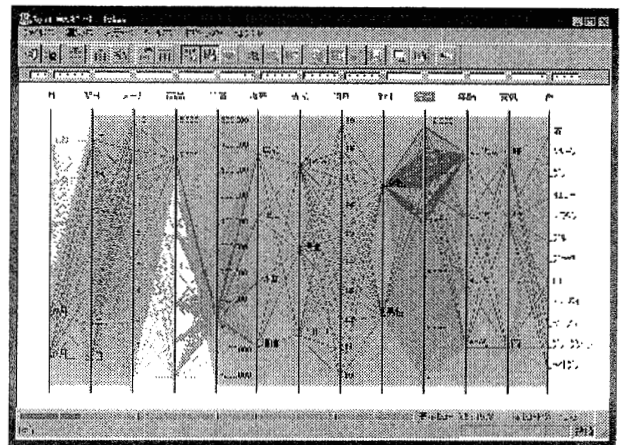
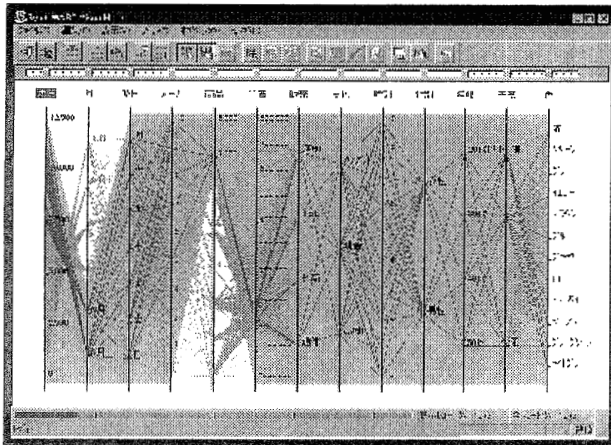
平行座標によるビジュアル表現を用い、かつ、フィルタリング、ズーム、マグニファイ操作を組み合わせることによって、上記の要件を十分に満足するビジュアライジング OLAP を実現することができる。

ビジュアライジング OLAP には、2つのデータビジュアライジング方法がある。1つは、いくつかの切



資本金の低い部分でズームすると、密集していた低資本金のデータが拡大されます。

図2 ズーム



スペクトルバーで購入顧客のプロフィールをマグニファイすると、傾向が浮き彫りになります。

図3 マグニファイ

り口で集約されたセグメントごとにビジュアライズするもので、セグメントビジュアライゼーションと呼ぶ。2つ目は、生のデータをそのままビジュアライズするディテールビジュアライゼーションである。

セグメントビジュアライゼーションでは、セグメントごとのパターンの類似性や特徴を認識するのに役立つ。また、ディテールビジュアライゼーションは、生のデータの集合（線の集中度）によって、軸間の相関や特異データを探るのに役立つ。図4にセグメントビジュアライゼーションの例を、図5にディテールビジュ

アライゼーションの例を示す。図4は、地域と顧客層を限定すると、購入商品のパターンが見えてくることを示す。また、図5は、5月～6月の売れ筋商品と、その購買層が強調されることを示す。

ビジュアライジング OLAP は、他の統計解析やデータマイニングによって得られた解を軸の形で連動することによって、より幅広い分析が可能になる。たとえば、クラスタ分析と連動して、新たに生成された分類軸を平行座標に加えることによって、各グループの特性を直感的に捉えることができる（図6参照）。ま

店名	曜日	時刻(時)	性別	年齢	商品	金額	支払	時刻	性別	年齢	色
1. 藤心堂東店	SUN	02-01	小学生	12	15	1	10	6			
2. 藤心堂東店	SUN	02-01	小学生	1	6	1	1	1			
3. 藤心堂東店	SUN	02-01	小学生	1	3	1	1	1			
4. 藤心堂東店	SUN	02-01	小学生	14	21	6	25	20			
5. 藤心堂東店	SUN	02-01	小学生	3	2	5	5	1			
6. 藤心堂東店	SUN	01-03	小学生	7	19	2	15	12			
7. 藤心堂東店	SUN	01-03	小学生	1	1	1	1	1			
8. 藤心堂東店	SUN	01-03	小学生	1	1	1	1	1			
9. 藤心堂東店	SUN	03-05	小学生	1	1	1	1	1			
10. 藤心堂東店	SUN	03-05	小学生	1	1	1	1	1			
11. 藤心堂東店	SUN	03-05	小学生	1	1	1	1	1			
12. 藤心堂東店	SUN	03-05	小学生	1	1	1	1	1			
13. 藤心堂東店	SUN	03-05	小学生	1	1	1	1	1			
14. 藤心堂東店	SUN	05-07	小学生	1	1	1	1	1			
15. 藤心堂東店	SUN	05-07	小学生	1	1	1	1	1			
16. 藤心堂東店	SUN	05-07	小学生	1	1	1	1	1			
17. 藤心堂東店	SUN	05-07	小学生	1	1	1	1	1			
18. 藤心堂東店	SUN	05-07	小学生	1	1	1	1	1			
19. 藤心堂東店	SUN	05-07	小学生	1	1	1	1	1			
20. 藤心堂東店	SUN	05-07	小学生	1	1	1	1	1			
21. 藤心堂東店	SUN	07-09	小学生	1	1	1	1	1			
22. 藤心堂東店	SUN	07-09	小学生	1	1	1	1	1			

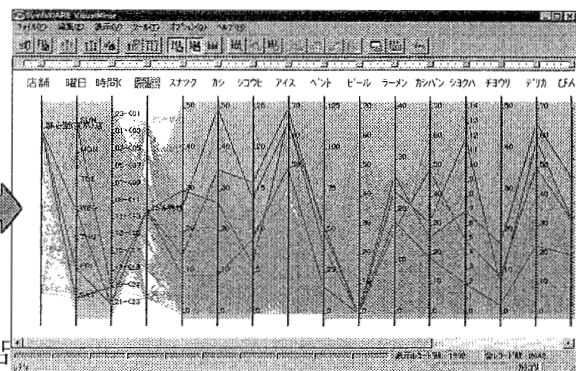


図4 セグメントビジュアライゼーション

キー	月	曜日	メーカー	商品	単価	販売	支払	時刻	性別	年齢	色
1. 5月	5月	月	1,400	31,600	店舗	現金	19	男性	20代以下	アタラシ	
2. 6月	6月	月	3,100	7,900	店舗	現金	14	女性	20代以下	アタラシ	
3. 7月	7月	月	9,100	19,300	店舗	現金	15	女性	30代	アタラシ	
4. 8月	8月	月	11,005	16,100	店舗	現金	10	女性	30代	アタラシ	
5. 9月	9月	月	9,400	31,600	店舗	現金	10	女性	30代	アタラシ	
6. 10月	10月	月	3,300	45,100	店舗	現金	18	女性	30代	アタラシ	
7. 11月	11月	月	10,205	43,300	店舗	現金	17	女性	30代	アタラシ	
8. 12月	12月	月	1,100	80,800	店舗	現金	13	男性	40代	アタラシ	
9. 1月	1月	月	1,100	13,600	店舗	現金	16	男性	30代	アタラシ	
10. 2月	2月	月	1,400	360	店舗	現金	19	男性	30代	アタラシ	
11. 3月	3月	月	10,100	12,800	店舗	現金	11	女性	30代	アタラシ	
12. 4月	4月	月	1,100	6,500	店舗	現金	19	女性	30代	アタラシ	
13. 5月	5月	月	3,300	31,600	店舗	現金	19	男性	30代	アタラシ	
14. 6月	6月	月	1,100	13,600	店舗	現金	16	男性	30代	アタラシ	
15. 7月	7月	月	1,400	360	店舗	現金	19	男性	30代	アタラシ	
16. 8月	8月	月	10,100	12,800	店舗	現金	11	女性	30代	アタラシ	
17. 9月	9月	月	1,100	6,500	店舗	現金	19	女性	30代	アタラシ	
18. 10月	10月	月	6,100	202,600	店舗	現金	13	男性	40代	アタラシ	
19. 11月	11月	月	1,100	11,900	店舗	現金	15	男性	40代	アタラシ	
20. 12月	12月	月	5,400	2,800	店舗	現金	19	女性	40代	アタラシ	
21. 1月	1月	月	10,100	7,900	店舗	現金	18	男性	30代	アタラシ	

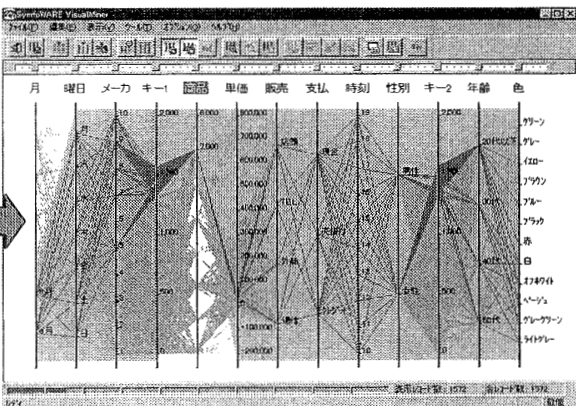
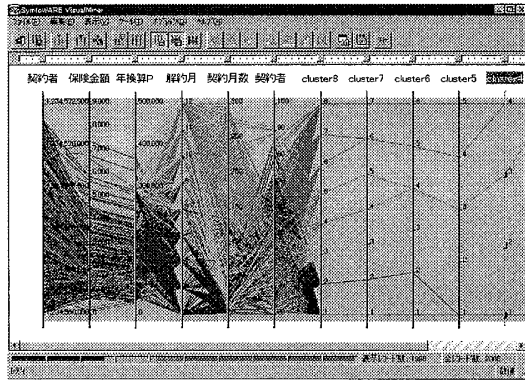


図5 ディテールビジュアライゼーション



4 クラスタに分類したとき、各クラスタの特性が分かる。

図6 クラスタ分析結果による特性把握

た、多変量解析の主成分（因子）分析と連動して、新たに生成された主成分（因子）軸でマグニファイすることで、データの直感的なセグメンテーションを行うことが可能になる（図7参照）。

このように、ビジュアライジングOLAPでは、他の種々の手法によって生成された新しい情報を付加的に表示することにより、情報の表現力を高め、人の発想を支援する新しいGUIとして柔軟に拡張していくことが可能である。

#### 4. ビジュアライジングOLAPの適用例

ビジュアライジングOLAPの適用分野として以下がある。

#### ❖顧客購買特性の把握

複数の切り口から、売れ筋商品の購買層や関連購買商品を直感的に把握する。

#### ❖営業力増大に向けた営業員特性の把握

複数の切り口から、営業員の販売傾向、得意分野を直感的に把握する。

#### ❖品質管理におけるの関連要因の把握

工程ごとの品質状況と特性要因の間の関連性を直感的に把握する。

#### 購買特性分析例

販売明細データを顧客層別に集計したデータのセグメントビジュアライゼーションを用いて、購買特性分析を行ってみる。

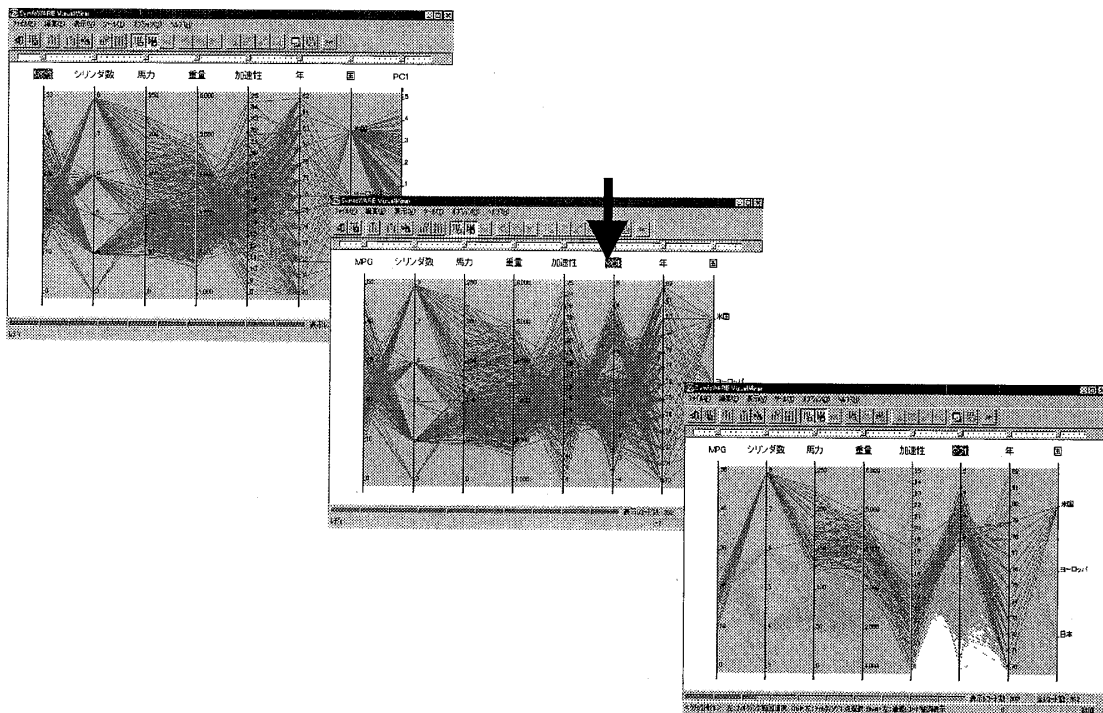
全体の傾向をみて、着目したい部分を絞り込む。たとえば、「菓子」の売り上げの高い部分に着目してみる。

その結果、選択範囲内のデータだけが強調表示される。ミドルの男性／女性が多いという傾向が分かる。

さらに、「東京都心部」に着目してみる。

その結果、都心部ではミドル男性が、菓子やアイスクリームをよく買っている傾向が直感的に分かる。また同時に、買う時間帯から「残業食」であることも想像がつく（当然、ビールは売れていない）。

このように、ビジュアライジングOLAPを使用す



第1主成分軸でマグニファイ。データの弁別が可能。

図7 主成分分析結果によるデータ弁別

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	店舗名	曜日	時間帯	客層	スナック	菓子	嗜好品	アイスクリーム	弁当	ビール	びん・缶飲料	書籍	衣料
2	都心型	日	23<-01	小学生	13	15	1	10	8		63	14	
3	都心型	日	23<-01	ヤング男性	1	6			1		22	2	
4	都心型	日	23<-01	ヤング女性	1	3	1				8		
5	都心型	日	23<-01	ミドル男性	14	21	6	25	28		72	11	
6	都心型	日	23<-01	ミドル女性	3	2	5	6	1		11	2	
7	都心型	日	01<-03	小学生	7	15	2	15	12		57	6	2
8	都心型	日	01<-03	ヤング女性	1			1	1		2		
9	都心型	日	01<-03	ミドル男性								4	
10	都心型	日	03<-05	小学生									1
11	都心型	日	03<-05	女子中高生							1		
12	都心型	日	03<-05	ヤング男性	5	5	2	8	22		16	12	
13	都心型	日	03<-05	ヤング女性	2	3		1	9		13	2	1
14	都心型	日	05<-07	小学生		3	1	9	9		4	1	
15	都心型	日	05<-07	男子中高生			1	1	4		3	2	
16	都心型	日	05<-07	女子中高生		3			7		2		
17	都心型	日	05<-07	ヤング男性		6		6	27		27	9	1
18	都心型	日	05<-07	ヤング女性	1	6	3	6	24		26	10	3
19	都心型	日	05<-07	ミドル男性		1			11		3	1	
20	都心型	日	05<-07	ミドル女性	4	6	2	10	25		18	7	
21	都心型	日	07<-09	小学生	6	7	1	5	9		15	6	1
22	都心型	日	07<-09	男子中高生					4				

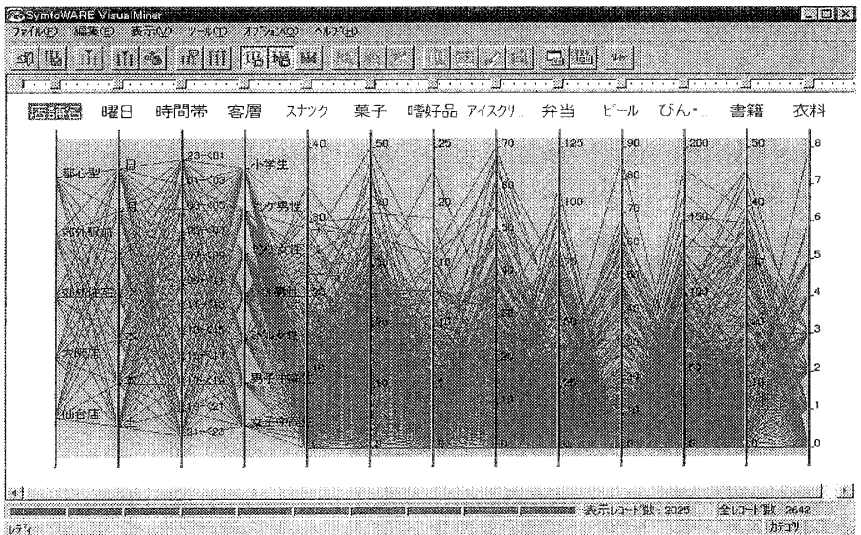


図8 客層別商品別販売実績とビジュアライズ

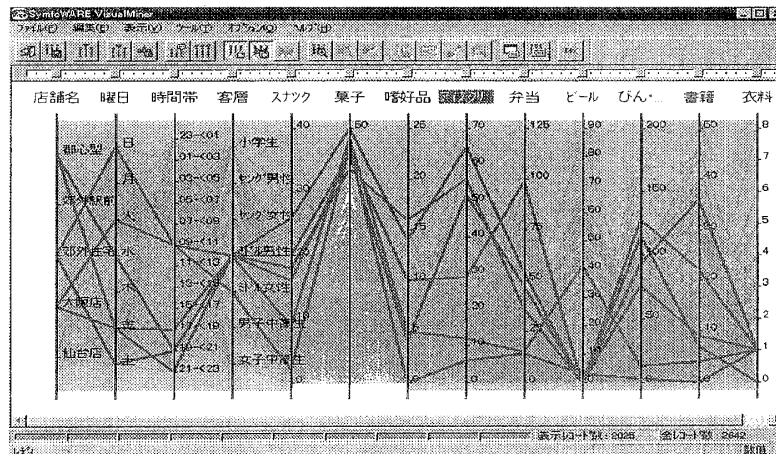


図9 特定商品でフィルタリング

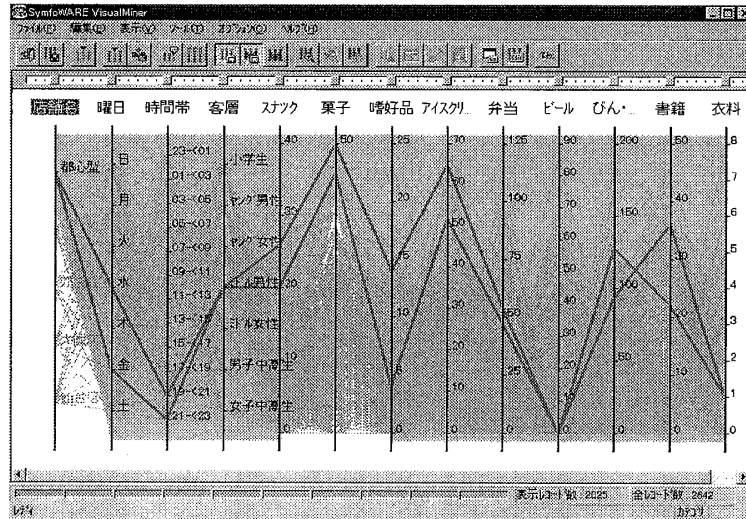


図10 特定地区でフィルタリング

ると、たくさんの切り口からなる複雑なデータから、着目すべき顧客セグメントを認識したり、売上の傾向や特性などを容易に把握することができ、販促施策のための商品レイアウトの検討に利用することができる。

参考文献

[1] W.H.Inmon: Building the Data Warehouse,

Wiley-QED, 1995, Chapt7 EIS and Data Warehouse  
 [2] Ruth Dilly (The Queen's Univ. of Belfast): Data Mining-An Introduction, 1995, Chapt2 Data Mining Functions, Chapt3 Data Mining Techniques  
 [3] 加藤ほか: 戦術的データマイニングシステム: FUJITSU