

1W-6

生体情報のデータウェアハウス構築および

OLAP によるデータ分析支援環境構築

秋野 俊治 新本 昌弘 不破 拓 和田 雄次

東京電機大学情報環境学部

1 はじめに

生体情報である視線データを計測するツールとしてはFree viewなどがあるが、測定された視線データをデータベース化する研究は進んでいるもののオンライン分析処理ツール（以後OLAP）で自由に扱えるものは調べた限り皆無であった。ところが、意思決定支援ツールとして有名なOLAPによる分析は生体情報分析にあまり適用されていないのが現状である。オープンソフトウェアを利用して可能な限り低コストでデータウェアハウス（以後DWH）を構築しOLAPによる生体データの分析を支援環境の構築を行いたい。

2 研究の目的

使用するデータは Free View にて測定された生体情報の視線データを利用する。測定されたデータは、5 分間ディスプレイ上に写る点を視線が追っていくことにより測定する。この実験の被験者は、3名（男性2名、女性1名）である。このデータをデータクレンジング作業によってDWHに格納できる形へと変形させ、構築するためのスキーマ設計を行う。そのスキーマを使ってキューブを作成する。これによりOLAPによる分析の可能性・有効性を評価し、生体情報の専門家の分析環境の元となるものを作ることが目的である。

3 OLAP

3. 1 DWH とは

DWHは、集められた大量の元データを多次元データベースに格納し、これを様々な角度から検索・集計して問題点や解決策を発見する分析ツールである。

3. 2 OLAP とは

データベースを多次元的に解析し、視覚化するシステムである。また、解析結果を必要とするエンドユ

ーザが直接分析をし、結果を得られるツールである。

3. 3 OLAP による DWH の構築

今回 OLAP ツールとして、IAF consulting 社が開発したオープンソースである OpenOLAP を使用する。

測定されたデータをクレンジング作業によって被験者3人のデータを統合し、IDを振る。また、時間表示がないため、時間をIDにより作成した。（図1）

	A	B	C	D	E	F	G	H
1	カテゴリ	角度X [deg]	角度Y [deg]	瞳孔径X [dot]	瞳孔径Y [dot]	移動速度 [deg/sec]	注視時間 [msec]	隣接 [Y/N]
2	X	0.28	-1.44	77	84	0	0.N	
3	X	0.11	-0.92	77	83	16.41	0.N	
4	X	0.24	-1.06	77	82	6.72	0.N	

↓

	A	B	C	D	E	F	G	H
1	時間[1/3秒]	被験者番号	角度X [deg]	角度Y [deg]	瞳孔径X [dot]	瞳孔径Y [dot]	移動速度 [deg/sec]	注視時間 [msec]
2	time_id	people_id						
3	t0001	1	0.48	0.25	82	80	0	0
4	t0002	1	0.6	0.38	82	79	5.31	0
5	t0003	1	0.44	0.43	82	79	0	0
6	t0004	1	0.61	0.41	82	79	5.14	0

図1 視線データのクレンジング

クレンジングされたデータを使いDWHを構築した。図2の sisen_fact をファクトテーブルとし、time_id は時間を示し、people_id は被験者の番号を示す。

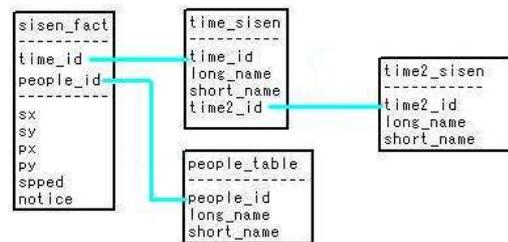


図2 視線データのDWH図

4 OLAP による評価

4. 1 分析ツールによる評価手法

データの時間を変化させ粒度毎の違いを表示できるかどうか。また、遠距離にいる研究者同士で互いのPCに接続し同様な分析が可能か。Excelよりも分析しやすいツールであるかという3つのポ

イントを評価基準と考えました。

4.2 OLAPの主な機能

OLAPでは作成したキューブを用い、ドリルダウン・アップ、スライシング、ダイシングといった機能を用いた多次元的な分析が可能である。

図3では左のデータは一分毎の合計が右のデータでは1/3秒毎のデータが表示されている。このように粒度を変化させ、詳細にみることができる操作をドリルダウンという。この逆をドリルアップという。

時間[1/3]	角度X	角度Y
時間[1/3] 合計	10.16	
1 minute	10.16	
2 minutes	9.64	
3 minutes	7.42	

図3 ドリルダウン・アップ

図4では赤の矢印のように「hikensya1」を「hikensya2」に変更することにより、被験者の変更を行うことにより、データ変化を見ることができる操作がスライシングである。

青の矢印のように「時間」軸と「メジャー」軸を入れ替える操作をする。このように複雑なデータを多次元に見るために、縦軸と横軸の入れ替えを行う操作をダイシングという。

時間[1/3]	角度X
時間[1/3] 合計	10.16
1 minute	10.16
1	0.48
2	0.60
3	0.44

図4 スライシング・ダイシング

4.3 OpenOLAPによる評価

OpenOLAPの3つの機能およびグラフ表示機能を用い、瞳孔径Y軸の大きさ、移動速度の速さと注視時間との関係をOLAPによる分析を行うとした場合、「時間」は1分目に焦点を当て、一分目のみドリルダウン

させてある。

「数値データ」は瞳孔径Yと視線速度、注視時間のみを選択し表示させている。これらを面とし「被験者番号」をスライシングで変化させることで被験者1と被験者3のデータの違いの分析を行った。図5、6のグラフ部分では、赤が瞳孔径Y、青が移動速度、緑が注視時間を示している。

また、これら分析はネットを介した遠距離からの分析も可能である。



図5 被験者番号1のOpenOLAPによる分析

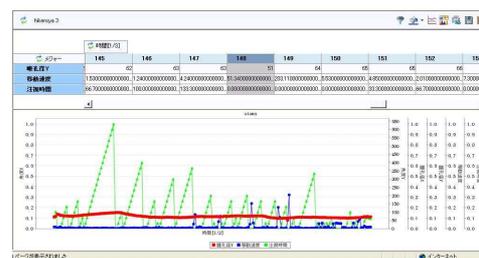


図6 被験者番号3のOpenOLAPによる分析

5 まとめ

本研究ではOpenOLAPを用いることで、Excelに比べさらに簡単な操作で分析が可能であり、専門の操作を学ばなくてもクリック・ドラッグ&ドロップといった操作で分析できる。次の研究課題として、Web上のHP内で利用者が注視した特定フィールドの順番を測定し、注視したパターンを抽出・分析することが考えられる。

今後、生体情報の専門家とのさらなる情報交換を通じたOLAPの適用評価を行っていく必要があると考える。

参考文献

[1] IAF Consulting

HP:<http://www.iafc.co.jp/products/openolap.htm>