



Методы визуализации многомерных медицинских данных на основе обобщенных графических образов

Берестнева О.Г., Воловоденко В.А.,
Гергет О.М., Шаропин К.А.,
Осадчая И.А.

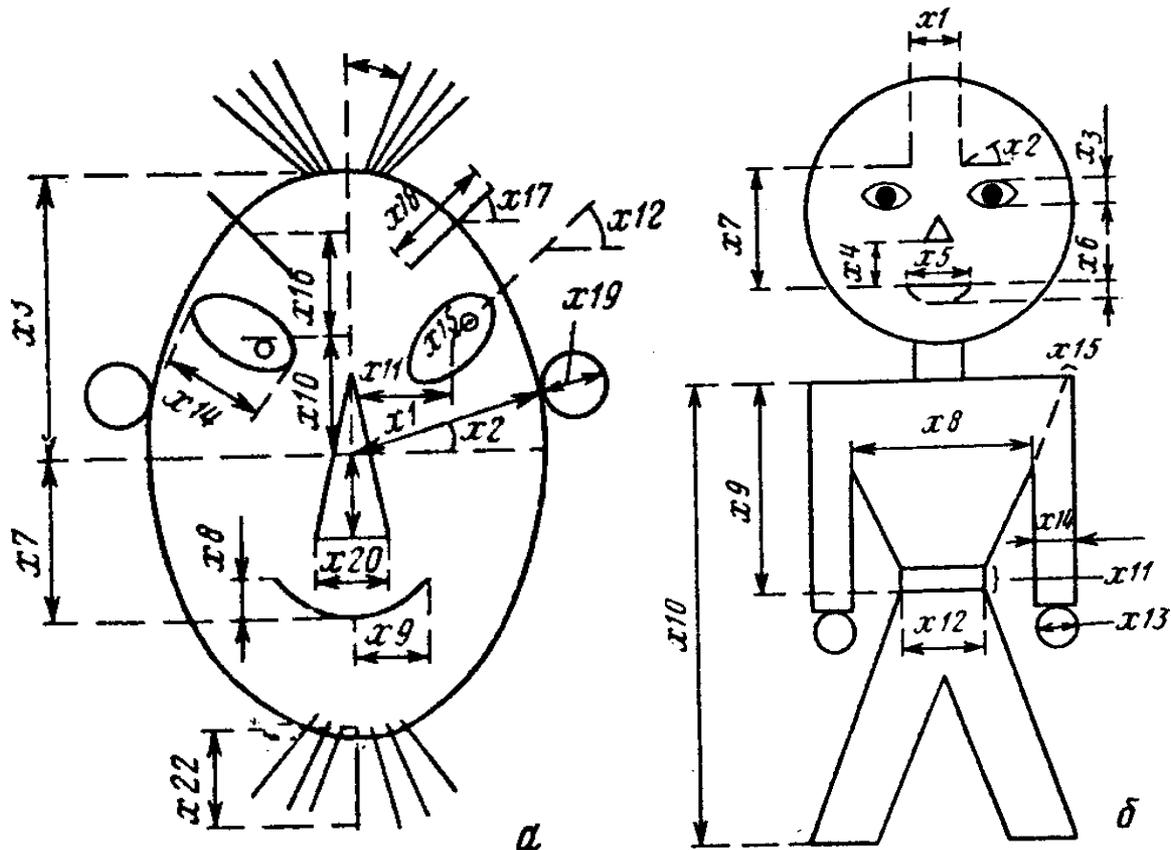
ИОИ-2014, Греция, Крит

- Визуализация данных – задача, с которой сталкивается в своей работе любой исследователь.
- К задаче визуализации данных сводится проблема представления в наглядной форме данных эксперимента или результатов теоретического исследования.
- Возможности инструментария визуализации определяется областями его применения, и конкретизируются совокупностями управленческих задач, решаемых с его помощью.



- Современные компьютерные технологии используют широкий спектр методов визуализации информации. Легкость построения графиков и диаграмм с помощью ЭВМ все заметнее меняет когнитивные навыки исследователя. Современные пакеты анализа информации позволяют строить сотни типов различных графиков и диаграмм, в том числе так называемые пиктографики «Лица Чернова».

Примеры построения пиктографики «Лица Чернова»



Лица Чернова представляют собой схематичное изображение лиц, определенным чертам которых соответствуют относительные значения заданных характеристик.



ВИЗУАЛИЗАЦИЯ МНОГОМЕРНЫХ ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ

- Основной задачей визуализации данных является задача получения визуального образа, однозначно соответствующего набору данных.
- В докладе рассмотрен подход, который позволяет провести визуализацию основных линейных конструкций.



- Общая форма представления данных является вектор конечномерного пространства R

$$A = (a_0, a_1, a_2, \dots, a_{n-1}) \in R_n \quad (1)$$

- Для перехода от данного вектора к визуальному образу будет использоваться базис ортонормированных функций $\{\varphi_i(\tau)\}_{i=0}^{\infty}$
- В качестве такого базиса можно использовать известные функции, в частности $\{l_i(\tau)\}_{i=0}^{\infty}$ ортонормированные полиномы Лежандра на отрезке $[0, 1]$, множество которых мы обозначим через
- Точке A можно поставить в соответствие функцию

(2)

$$F_A(\tau) = \sum_{i=0}^{n-1} a_i l_i(\tau)$$



- Между (1) и (2) устанавливается взаимно-однозначная связь.
- Если ввести в рассмотрение второй вектор

$$B = (b_0, b_1, b_2, \dots, b_{n-1}) \in R_n$$

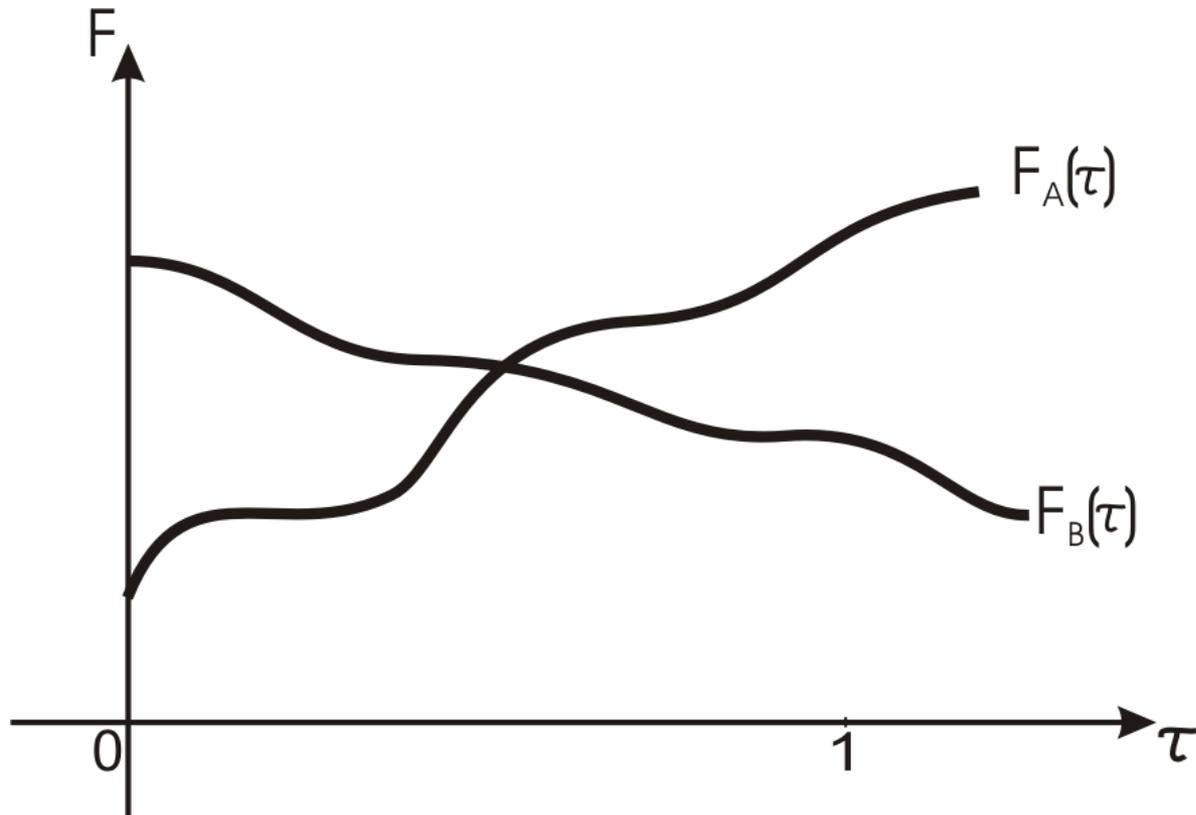
то ему ставится в соответствие функция

$$F_B(\tau) = \sum_{i=0}^{n-1} b_i l_i(\tau)$$

- Функции $F_A(\tau)$ и $F_B(\tau)$ являются визуальными образами точек A и B , принадлежащих пространству R_n

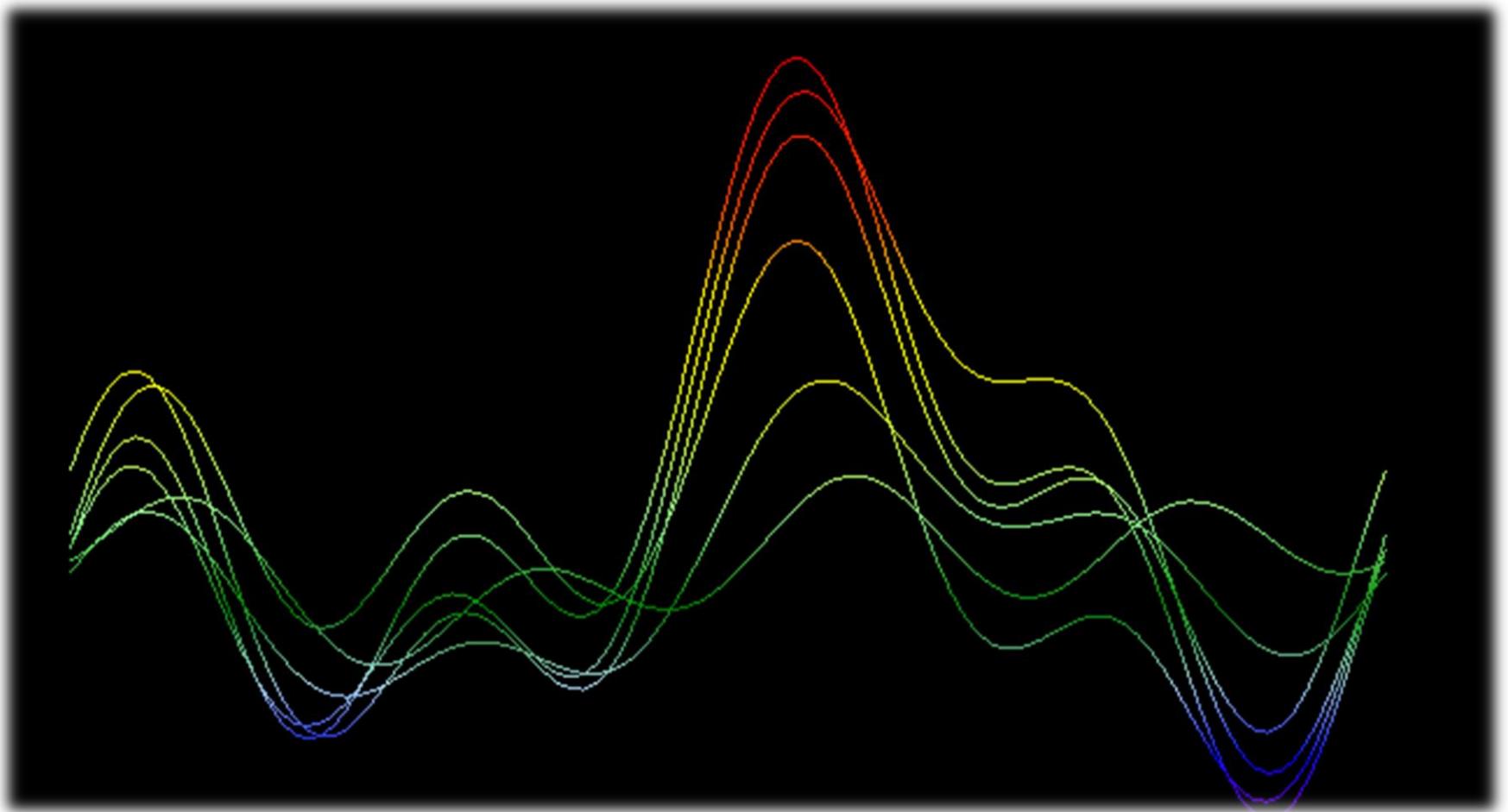


ВИЗУАЛЬНЫЕ ОБРАЗЫ ТОЧЕК A и B ,
ПРИНАДЛЕЖАЩИХ ПРОСТРАНСТВУ R_N



- Данный подход реализован в пакете *NovoSpark Visualizer* (www.novospark.com).
- Таким образом, основой визуализационного подхода *NovoSpark* является представление *многомерного* наблюдения в виде *двумерного* образа – кривой.
- При этом гарантируется, что близким по значениям наблюдениям *A* и *B* будут соответствовать визуально близкие образы-кривые; для сильно различающихся по значениям наблюдений их образы-кривые будут заметно отличаться.
- *NovoSpark Visualizer* – это современный инструмент визуализации, который соединяет в себе интуитивность визуального анализа с описательной мощностью статистических методов и предоставляет наиболее полную информацию о многомерных данных.





ВИЗУАЛИЗАЦИЯ ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ В ЗАДАЧЕ ОЦЕНКИ ПАРАЛЛЕЛЬНОСТИ ВАРИАНТОВ ТЕСТОВЫХ ЗАДАНИЙ

- В современной системе образования независимая аттестация студентов является наиболее объективной оценкой их знаний, потенциала их умственных возможностей.
- В связи с этим в Томском политехническом университете разработана система независимой оценки качества знаний студентов по общеобразовательным дисциплинам.
- Контрольно-измерительные материалы по дисциплинам представлены в нескольких вариантах. Например, по математике, имеется двадцать один вариант тестовых заданий.
- Из этого вытекает проблема параллельности («одинаковости») этих вариантов тестовых заданий, и, как следствие, качество оценки знаний студентов, и ее объективность.



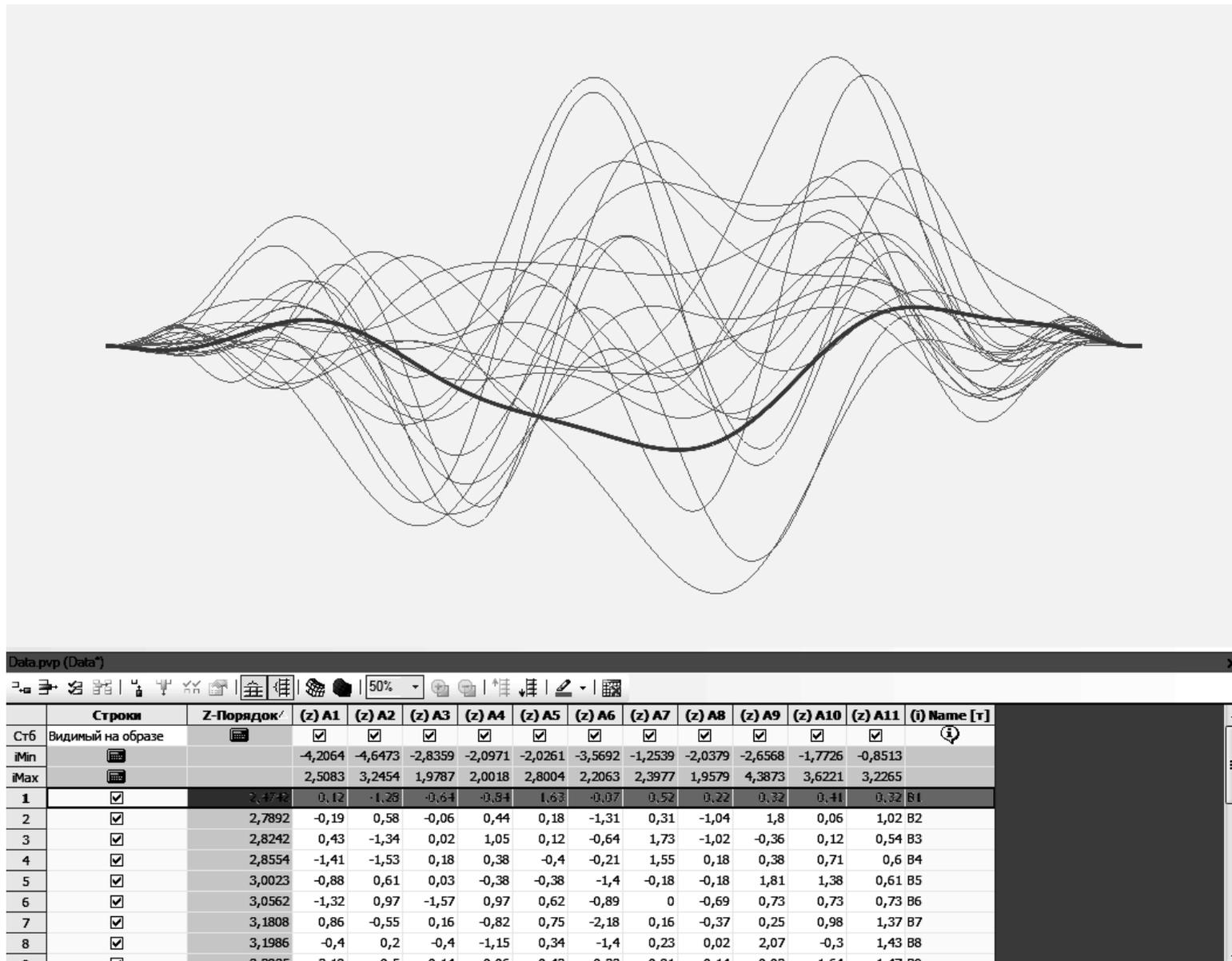
- Экспериментальные данные представляли собой результаты текущего контроля и оценки качества знаний по математике у студентов-первокурсников Томского политехнического университета.

Таблица 1. Стандартизированные значения трудности заданий теста A_i по вариантам B_j [5]

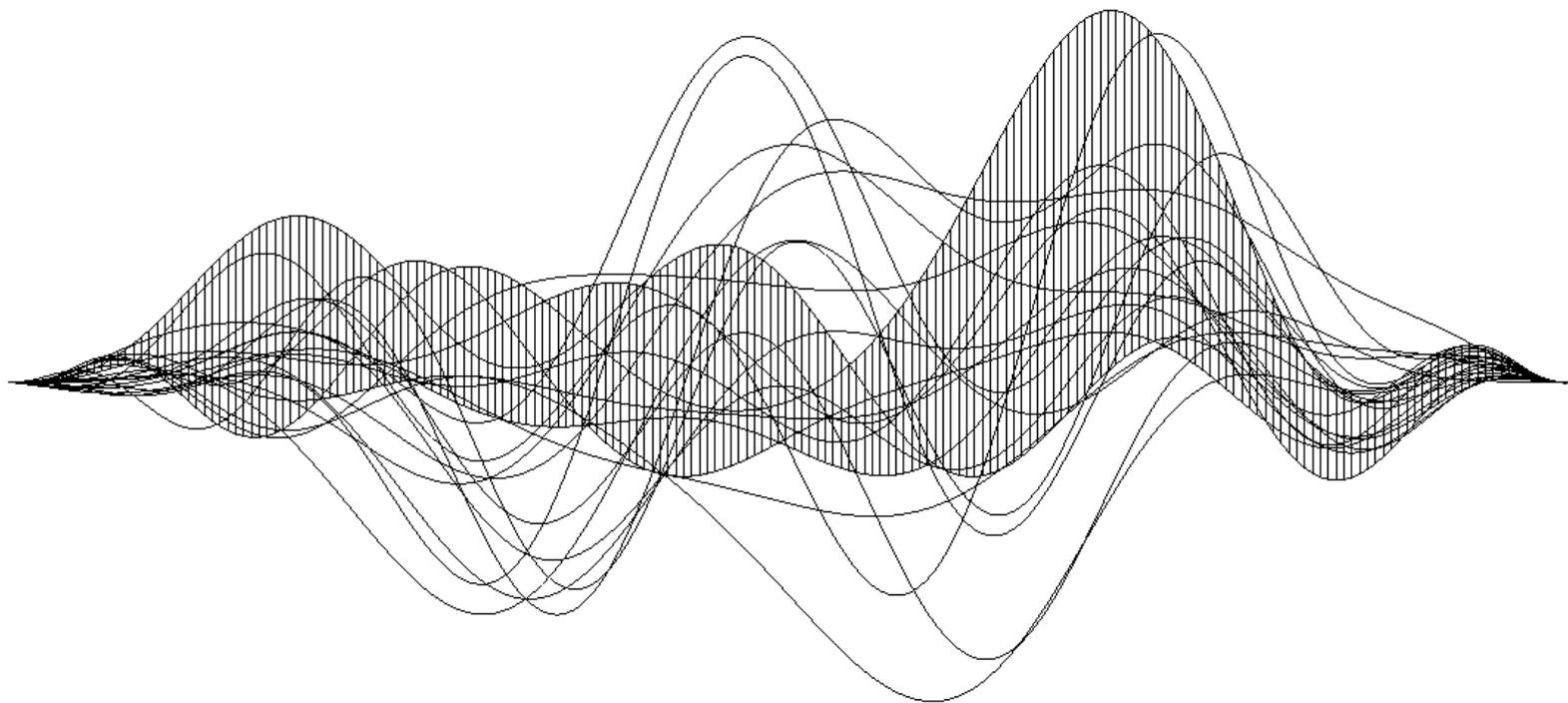
	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11
B1	-1,41	-1,53	0,18	0,38	-0,4	-0,21	1,55	0,18	0,38	0,71	0,6
B2	-1,32	0,97	-1,57	0,97	0,62	-0,89	0	-0,69	0,73	0,73	0,73
B3	-0,94	-1,18	-0,31	-0,31	-0,23	-0,39	-0,55	1,54	1,38	2,14	0,45



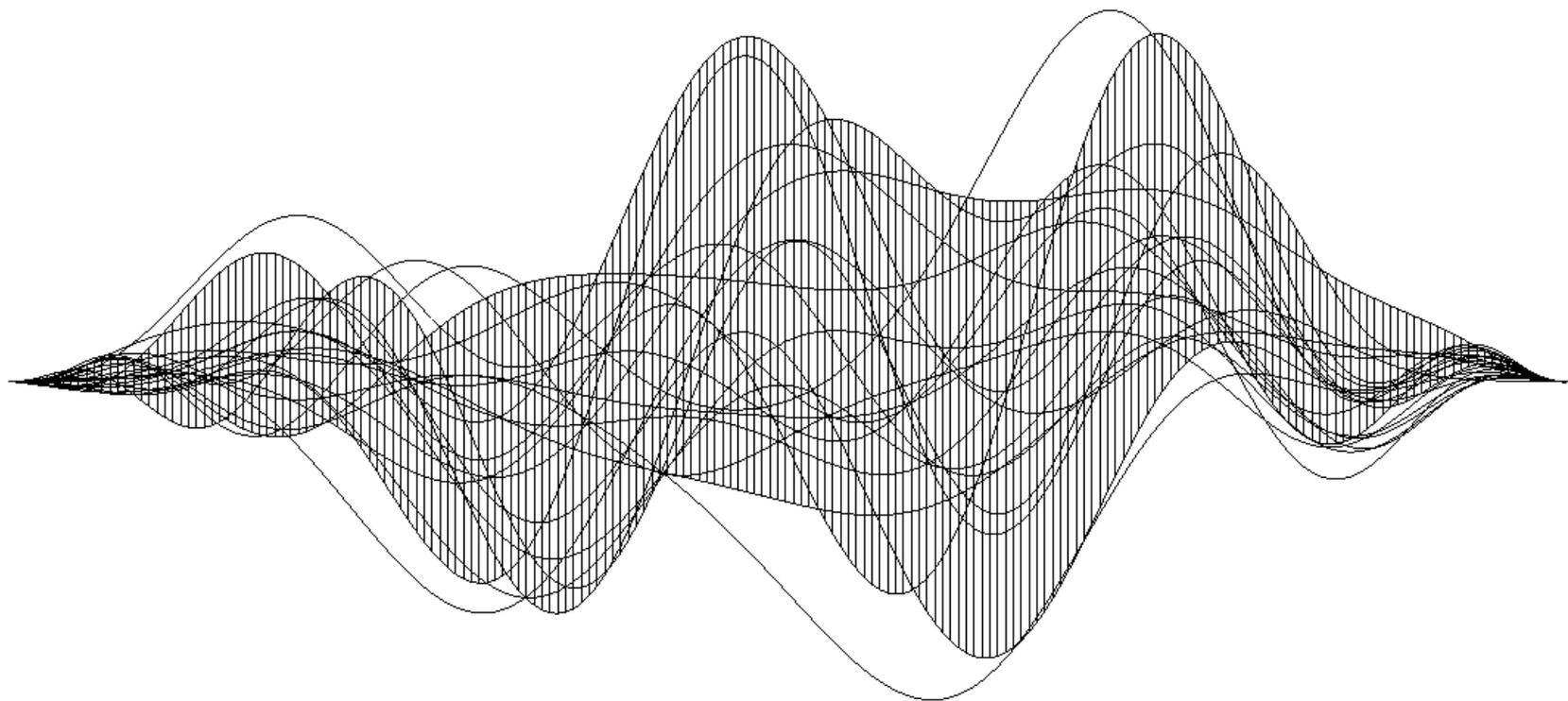
ИСХОДНЫЙ НАБОР ДАННЫХ



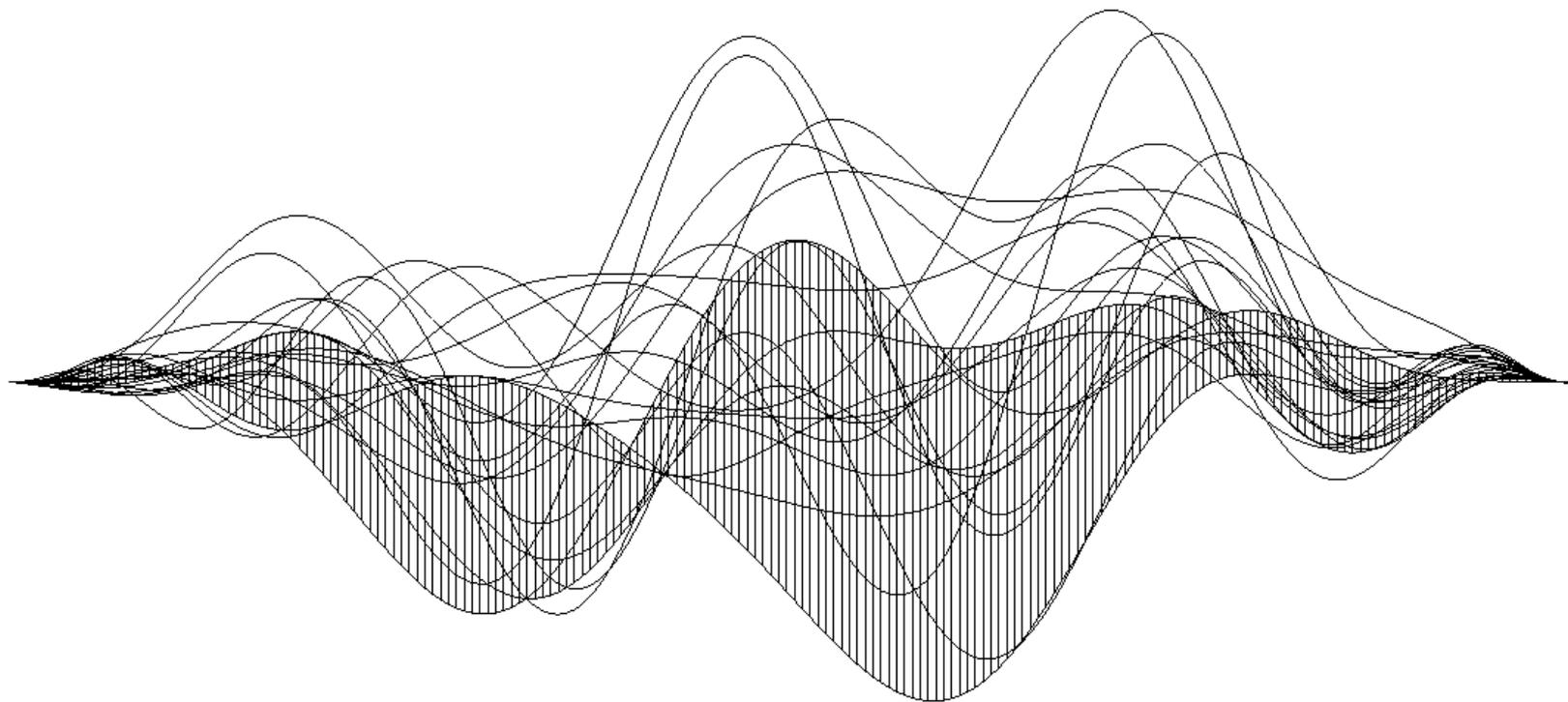
КЛАСТЕР № 1 – ВАРИАНТЫ 2, 5, 6, 8, 15, 21



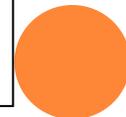
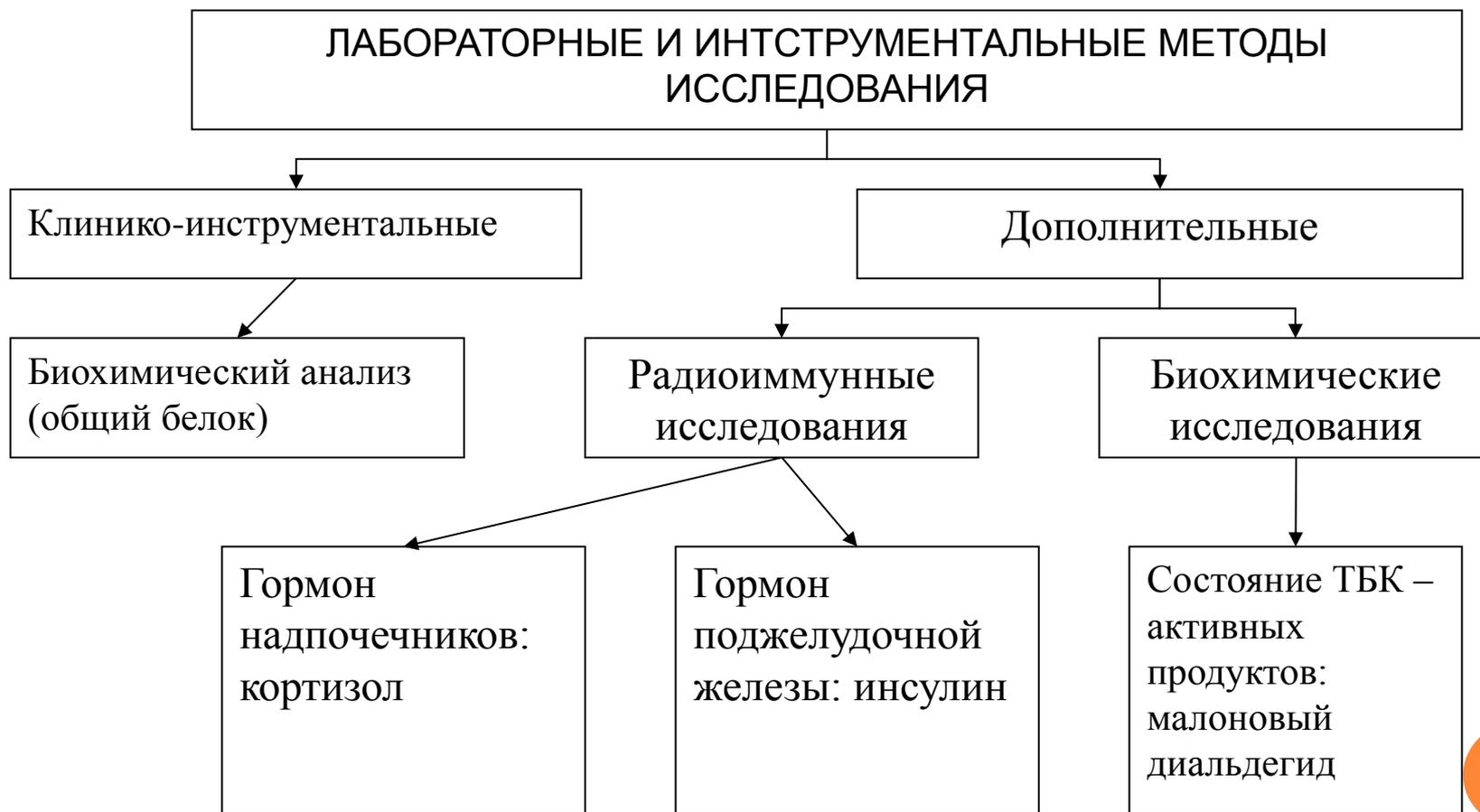
КЛАСТЕР № 2 – ВАРИАНТЫ 1, 4, 9, 10, 12, 14, 16, 17, 18, 19, 20



КЛАСТЕР № 3 – ВАРИАНТЫ 3, 7, 11, 13



ВИЗУАЛИЗАЦИЯ ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ В ЗАДАЧЕ МОНИТОРИНГА ФИЗИОЛОГИЧЕСКОГО СОСТОЯНИЯ БЕРЕМЕННЫХ ЖЕНЩИН



Экспериментальные данные

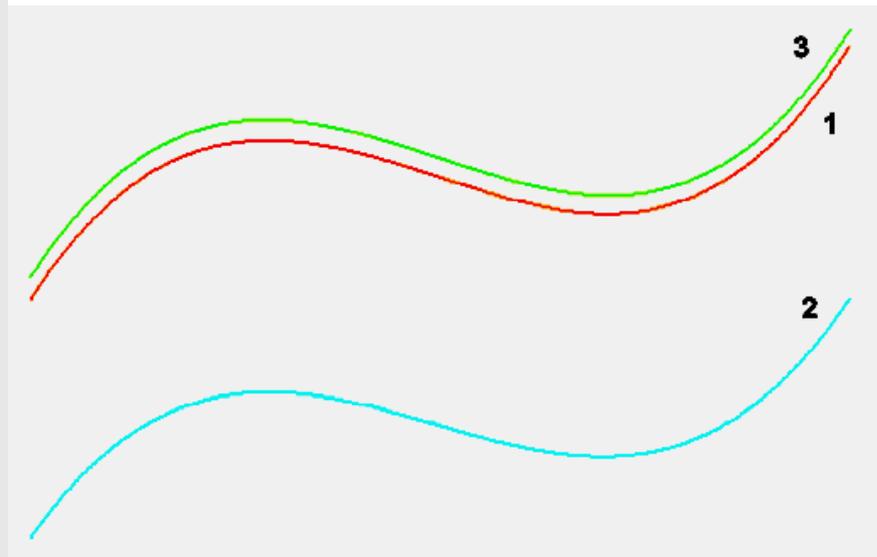
- **Контрольная группа (гр.1)** - практически здоровые беременные женщины.
- **Группа сравнения (гр.2)** - беременные женщины, с соматическими заболеваниями.
- **Основная группа (гр.3)** - беременные женщины, имеющие соматические заболевания, с которыми проводился комплекс оздоровительных мероприятий.

ФИО	гр	кор1	кор3	инс1	инс3	мда1	мда3	бел1	бел3
Ал-дер И.А.	1	465,08	685,03	16,43	25,45	2,87	2,65	66,52	68,05
Бал-ва Е.М.	1	436,00	760,00	14,78	24,05	3,00	2,95	71,05	68,55
Вер-на Т.С.	1	520,58	780,45	15,07	24,54	2,87	2,66	64,05	70,30
Вес-ва Г.А.	1	498,09	680,80	14,44	25,35	2,78	3,00	66,01	64,02

кор1, кор3 – уровень кортизола в I и III триместрах,
инс1, инс3 – уровень инсулина в I и III триместрах,
мда1, мда3 – уровень малонового диальдегида в I и III триместрах,
бел1, бел3 – уровень общего белка в I и III триместрах.



Анализ динамики состояния беременных женщин (I и III триместры)



1 - объект из контрольной группы,
2 - объект из группы сравнения,
3 – объект из основной группы



- Для сравнения индивидуальных наблюдений можно использовать их "спектральное" представление, которое подчеркивает отличительные характеристики каждой кривой и помогает более детально исследовать их визуальные свойства.

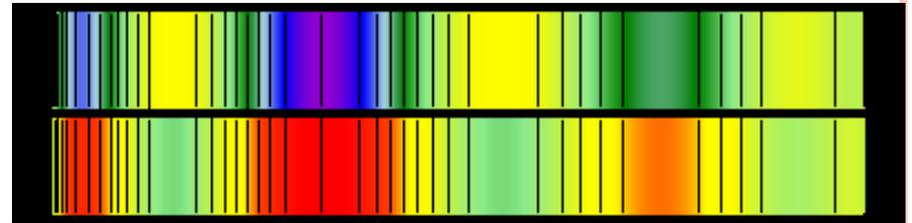
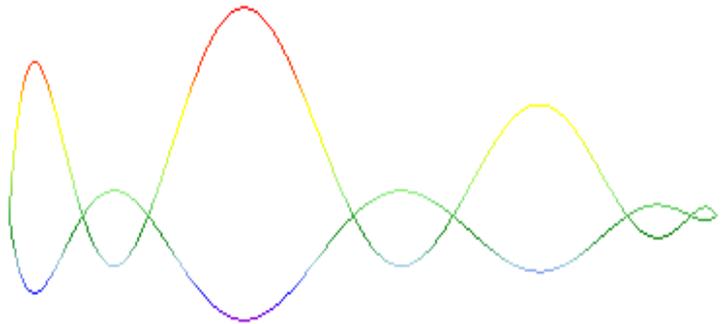


Рисунок 3.14 – «Спектры» многомерных наблюдений



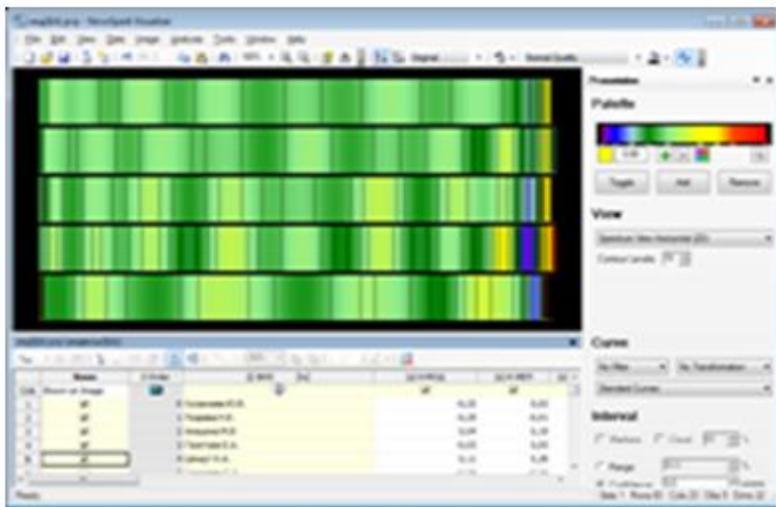


Fig. 2. The spectral representation of the data on patients diagnosed with BAPI

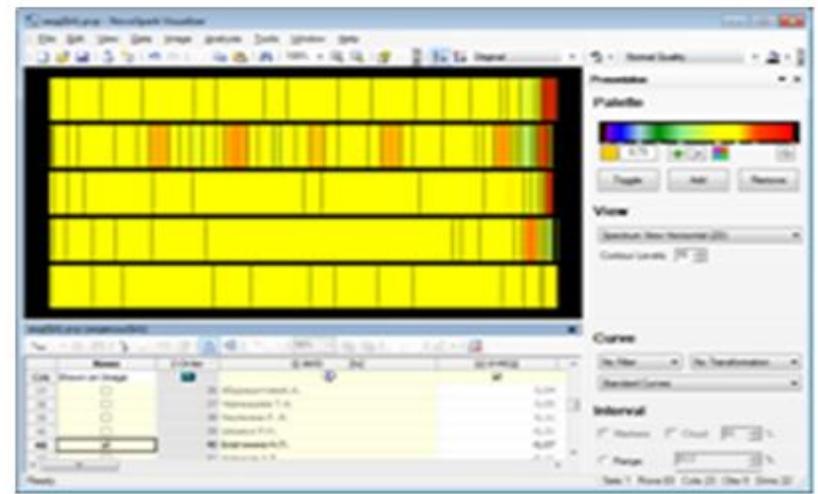


Fig. 3. The spectral representation of the data on patients with a diagnosis of D A C D

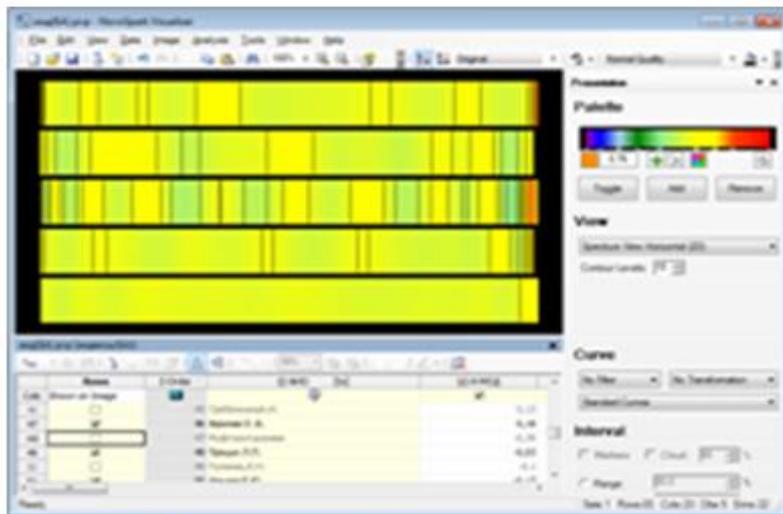


Fig. 4. The spectral representation of the data on patients diagnosed with BANP

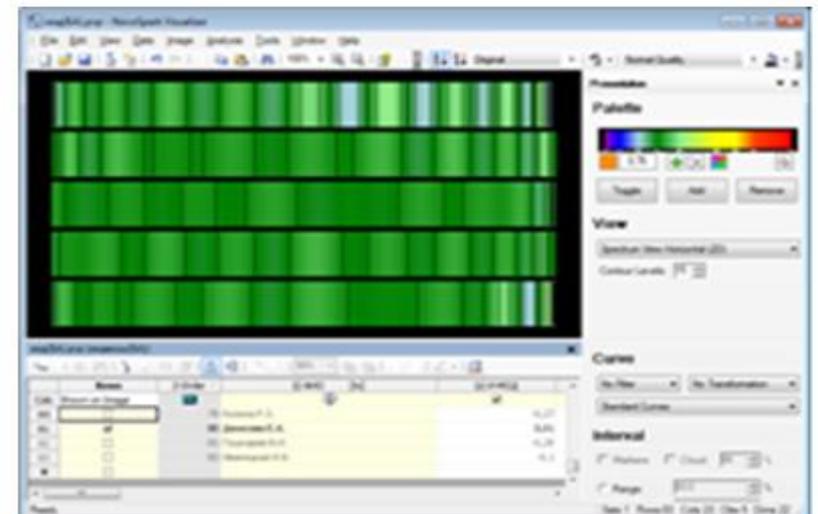


Fig. 5. The spectral representation of the data on patients diagnosed with PD

СПАСИБО ЗА ВНИМАНИЕ!!!

