

**Чернова О.Ф.**

ведущий научный сотрудник отдела НМОПЭ  
ФБУ РФЦСЭ при Минюсте России,  
доктор биологических наук

**Нилова М.В.**

ведущий государственный судебный эксперт  
лаборатории судебно-почвоведческих и биологических экспертиз  
ФБУ РФЦСЭ при Минюсте России,  
кандидат биологических наук

## **ОПЫТ ПРИМЕНЕНИЯ МЕТОДОВ МНОГОМЕРНОГО СТАТИСТИЧЕСКОГО АНАЛИЗА ПРИ ЭКСПЕРТИЗЕ ДОКУМЕНТОВ НА БУМАЖНОМ НОСИТЕЛЕ**

На конкретных примерах биологических экспертиз идентичности бумажных носителей, выполненных традиционными методами, показана возможность применения методов многомерного статистического группового анализа, в частности метода Уорда, для подтверждения и уточнения результатов экспертизы. Предлагаемый метод позволяет не только контролировать субъективные оценки, приводимые экспертом, но и давать более подробную характеристику сходства и различия тестируемых объектов.

**Ключевые слова:** биологическая экспертиза, статистический анализ, волокнистый состав бумаги

---

### **O. Chernova**

Lead research associate

Forensic Research Methodology Department Russian Federal Center of Forensic Science of the Ministry of Justice of the Russian Federation, DSc (Biology)

### **M. Nilova**

Lead forensic examiner Laboratory of Forensic Biology and Soil Analysis

Russian Federal Center of Forensic Science of the Ministry of Justice of the Russian Federation, PhD (Biology)

## **APPLYING MULTIVARIATE STATISTICAL ANALYSIS IN THE FORENSIC EXAMINATION OF HARD COPIES OF QUESTIONED DOCUMENTS**

Casework examples are used to demonstrate the application of multivariate statistical cluster analysis, and Ward's method in particular, to verify and clarify the results of questioned document examination obtained via traditional methods of biological analysis of document materials (paper substrate). The proposed method helps to control for subjective expert evaluation, and to characterize the differences and similarities between tested objects in a more detailed way.

**Keywords:** biological analysis, statistical analysis, fiber content of paper.

Одной из часто возникающих задач биологической экспертизы служит идентификация бумаги, на которой напечатан тот или иной документ, иллюстрация, рисунок, этикетка и т.д. В ходе экспертизы обычно исследуют абсолютный и относительный волокнистый состав бумаги (волокна, трахеиды, сосуды, редко встречающиеся элементы), что позволяет оценить сходство или различие изученных проб на групповом уровне. В результате тестирования накапливаются значительные по объему морфометрические данные, обработка которых трудоемка, а оценка субъективна. При обработке репрезентативных цифровых данных выявление различия и сходства тестируемых объектов становится более достоверным в случае применения адекватного алгоритма многомерного статистического анализа, с помощью которого успешно анализируются множественные промеры. Опираясь на наш опыт применения методов статистического анализа (4), мы предприняли новую попытку использования этих методов, на этот раз для анализа массива морфометрических данных, полученных в ходе экспертизы документов на бумажном носителе.

Цель нашей работы — разработка методических подходов к применению многомерных статистических методов для их внедрения в практику биологической экспертизы, на примере бумажного носителя документа.

### Материал и методы

Для достижения поставленной цели применен метод кластерного анализа, проведенного при помощи статистической программы STATISTICA 6, разработанной компанией «StatSoft» (США) (Халафян, 2008). Объектом исследования послужили документы А—Е на бумажном носителе, поступившие на исследование и прошедшие судебно-ботаническую экспертизу с помощью традиционных методов исследования, оформленную в виде заключений экспертов лаборатории судебно-почвоведческих и биологических экспертиз ФБУ РФЦСЭ при Минюсте России. Для определения волокнистого состава бумаги из листов документов взяты пробы сходного размера, которые затем подвергнуты мацерации. После мацерации материал обработан (непосредственно на предметном стекле) реактивом Херцберга (раствор хлор-цинк-йода) или реактивом Зеллегера (ГОСТ, 1987). Изготовленные временные препараты изучены с использованием микроскопа «Микромед-2» при увеличении от 100 до 400. Для определения абсолютного количественного состава бумаги подсчитаны от 600 до 1000 элементов (волокон, трахеид и сосудов). Для каждого образца вычислен относительный (в %) количественный состав элементов бумаги. С целью повышения достоверности определения относительного состава редко встречающихся элементов (сосудов) сосчитаны все сосуды (или их фрагменты) на всей площади препарата. На основании полученных результатов эксперт представил свои заключения. Для объективного контроля заключений экспертов построены фенограммы, в основу которых заложены промеры для каждой пробы (Табл. 1–5). Причем, при отсутствии промера элемента, в матрицу, подготовленную для составления фенограммы, внесены нулевые значения. Фенограмма создана по методу Уорда, при этом за метрическую основу приняты евклидовы расстояния. Сущность метода Уорда сводится к тому, что в качестве расстояний между кластерами используют прирост суммы квадратов расстояний до центров кластерных групп. На каждом шагу алгоритма объединяются такие два кластера, которые приводят к минимальному увеличению целевой функции, то есть внутригрупповой суммы квадратов /5/.

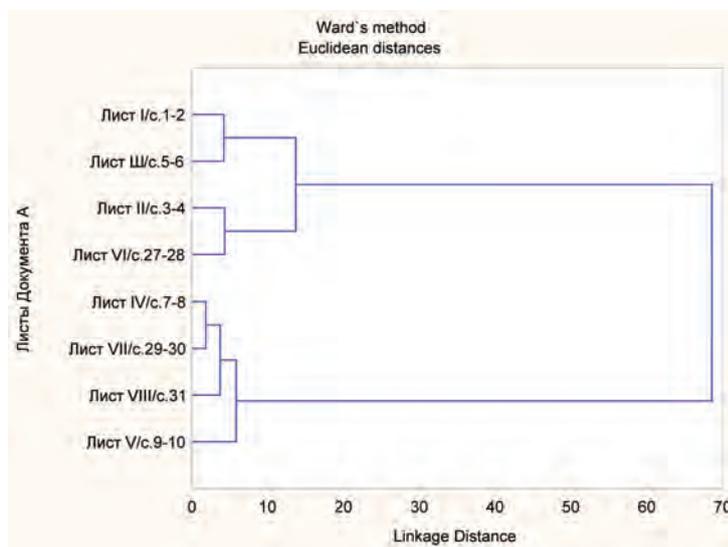
### Результаты

**Документ А.** Исследовали ряд пронумерованных листов для сравнения волокнистого состава бумаги. По данным эксперта, по составу бумаги листы разделяются на две группы (Табл. 1). (1) Лист Ш входит в общую группу с листами I, II и VI, но отличается от листов IV, V, VII и VIII. (2) Лист VII сходен с листами IV, V и VIII (имеющиеся различия в количественном соотношении элементов несущественны), но отличается по данному признаку от листов I, II, III и VI.

**Таблица 1. Качественный и относительный количественный состав элементов бумаги тестируемого Документа А (%% общего числа элементов)**

№ лист/ стр.	Состав бумаги					Соотношение сосудов		
	Элементы					Осина		
	древесной массы		целлюлозы					
	Волокна (и единичные трахеиды лиственных пород)	Сосуды	Волокна и трахеиды лиственных пород	Трахе- иды хвойных пород	Сосуды	Древес- ная масса	Целлю- лоза	Береза (цел- люлоза)
I /1-2	5,48	1,02	75,65	14,41	3,44	14,52	2,15	83,33
II /3-4	6,13	1,13	73,99	15,36	3,39	18,01	2,48	79,51
III /5-6	6,19	0,76	77,18	13,06	2,81	11,65	3,42	84,93
IV /7-8	0,19	0	79,96	15,36	4,48	0	9,04	90,96
V /9-10	0,1	0	80,23	15,17	4,5	0	12,35	87,65
VI /27-28	7,47	0,63	73,27	15,39	3,24	18,43	5,07	76,5
VII/ 29-30	0	0	81,34	14,85	3,81	0	8,5	91,5
VIII/ 31	0	0	79,42	17,21	3,37	0	10,06	89,94

В созданной нами фенограмме оперативной единицей служит определенный лист, состоящий из разных страниц. Фенограмма представляет собой древовидный граф, сформированный восемь терминальными группами (по числу листов) (рис. 1).



*Рис. 1. Фенограмма морфометрических данных промеров элементов бумажного носителя Документа А.*

Фенограмма свидетельствует о разделении тестируемых образцов на два кластера, которые полностью идентичны группам, выделенным экспертом, и подтверждают выводы экспертизы. Можно утверждать, что и сами группы неоднородны. Так, в первом (верхнем на фенограмме) кластере большее сходство между собой имеют с одной стороны Лист I и Лист III, а с другой — Лист II и Лист VI. Во втором кластере (нижнем на фенограмме). Лист IV более сходен с Листом VII и отличается от Листов VIII и V.

**Документ Б.** На экспертизу поступили образцы бумаги расписок. По качественному и количественному волокнистому составу бумаги образцы разделены на три группы (Табл. 2).

**Таблица 2. Качественный и относительный количественный состав элементов бумаги тестируемого Документа Б (% общего числа элементов)**

№ образца	Состав бумаги				Соотношение сосудов	
	Элементы				Осина	Береза
	древесной массы	целлюлозы				
Волокна и трахеиды лиственных пород		Трахеиды хвойных пород	Сосуды			
I	1,4	1,1	97,5	0	0	0
II	0,7	1,8	97,5	0	0	0
III	1,1	0,9	98	0	0	0
IV	0,4	0,4	99,2	0	0	0
V	0,2	86	7,2	6,6	58,7	41,3
VI	5,9	3	91,1	0	0	0
VII	0	83,9	5,6	10,5	69,1	30,9
VIII	0	82,7	10,6	6,7	58,1	41,9
IX	0	84,3	9,9	5,8	65,3	34,7

Бумага из Группы 1 и Группы 2 изготовлена практически только из древесины хвойных пород (сосны и ели). Группа 1 из проб I, II, III и IV характеризуется единичным присутствием элементов древесной массы, и почти 99% содержанием целлюлозы. Группа 2 представлена пробой VI. Древесная масса составляет почти 6%, целлюлоза — 94%, бумага изготовлена практически только из древесины хвойных пород (сосны и ели). Группа 3 объединяет пробы V, VII, VIII и IX. Вся бумага изготовлена из целлюлозы, а древесная масса отсутствует (несколько элементов обнаружено в пробе V). Бумага примерно на 90% состоит из элементов древесины лиственных пород.

Соответствующая фенограмма полностью подтверждает основные выводы эксперта, внося лишь некоторые уточнения (рис. 2). Так, группа 1 неоднородна, в ней проба I ближе к пробе III, а проба II ближе к пробе IV. Проба VI стоит особняком, но ближе к группе 1, чем к группе 3. Группа 3 распадается на два кластера — с одной стороны сходны пробы V и VIII, а с другой — пробы VII и VIII.

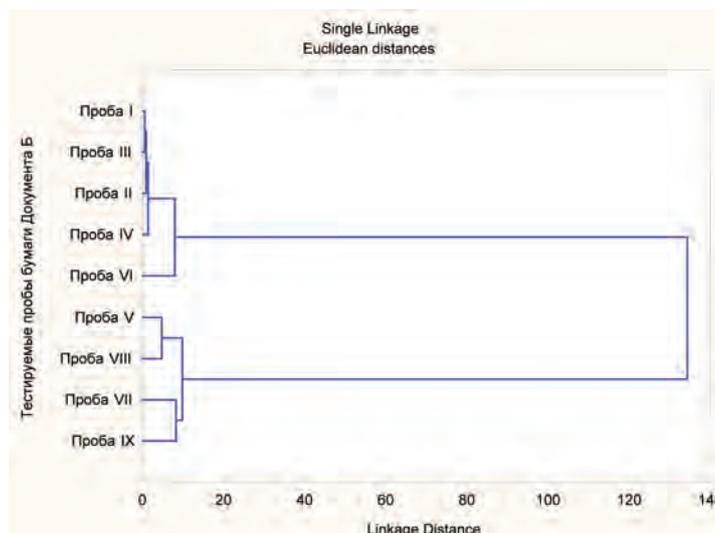


Рис. 2. Фенограмма морфометрических данных промеров элементов бумажного носителя Документа Б.

**Документ В.** Исследовали ряд пронумерованных объектов для сравнения состава бумаги. Установлено, что бумага всех объектов состоит только из целлюлозы хвойных и лиственных пород, а древесная масса отсутствует (Табл. 3).

**Таблица 3. Качественный и относительный количественный состав элементов бумаги тестируемого Документа В (%% общего числа элементов)**

№ образца	Состав бумаги			Соотношение сосудов	
	Волокна и трахеиды лиственных пород	Трахеиды ранней древесины хвойных пород	Сосуды	Береза	Осина (и единич. др. древесные породы с простыми перфорационными пластинками сосудов)
1	82,82	9,27	7,91	34,53	65,47
2	81,36	12,57	6,07	37,57	62,43
19	80,73	12,42	6,85	38,65	61,35
3	86,98	7,91	5,11	51,78	48,22
29	83,49	10,47	6,04	64,25	35,75
30	83,94	10,22	5,84	63,25	36,75
4	78,67	15,47	5,85	48,00	52,00

В результате экспертизы установлено сходство образцов 2 и 19, а также 29 и 30. Контрольная фенограмма позволила уточнить и расширить эти выводы (рис. 3). Все пробы распадаются на три кластера. Первый объединяет пробы 1, 2 и 19, что соответствует выводу эксперта. Но в этом кластере наиболее сходны пробы 2 и 19. Второй кластер обособляет пробы 3 и 4, что не было отмечено экспертом. Третий кластер состоит из проб 29 и 30, что соответствует выводу эксперта. Причем, этот кластер ближе ко второму кластеру (пробы 2 и 19), чем к первому кластеру (пробы 1, 2 и 19).

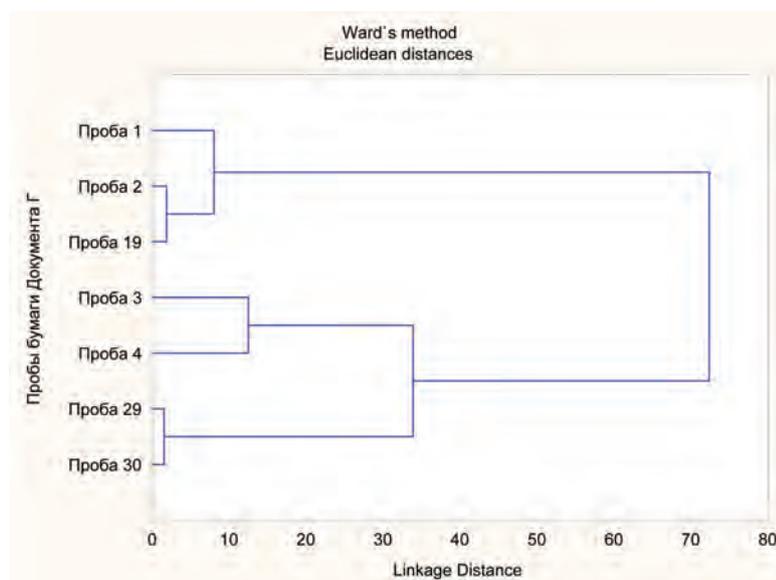


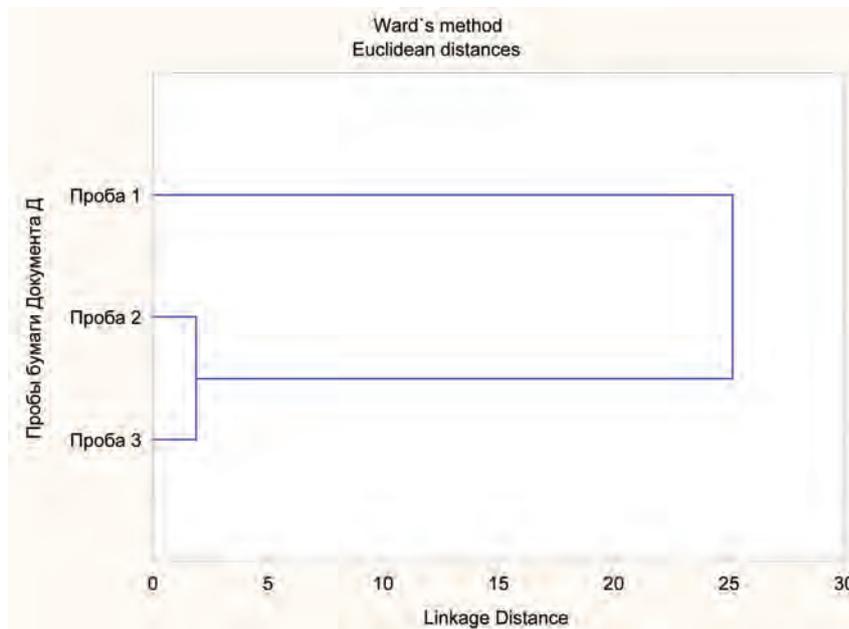
Рис. 3. Фенограмма морфометрических данных промеров элементов бумажного носителя Документа В.

**Документ Г.** Из каждого их трех листов Документа Г для определения состава бумаги отобраны соответственно образцы 1, 2 и 3. Экспертом установлено, что бумага всех образцов состоит практически только из целлюлозы хвойных и лиственных пород (с единичными элементами древесной массы) (табл. 4).

**Таблица 4. Качественный и относительный количественный состав элементов бумаги тестируемого Документа Г (%% общего числа элементов)**

Номера проб	Состав бумаги			Соотношение сосудов	
	Волокна и трахеиды лиственных пород	Трахеиды древесины хвойных пород	Сосуды	Береза	Осина
1	88,6	6,1	5,3	83,6	16,4
2	85,1	11,8	3,1	70,4	29,6
3	84,6	12,4	3,0	71,6	28,4

Установлено, что бумага всех образцов имеет в целом сходный качественный состав, но проба 1 отличается от проб 2 и 3 по соотношению элементов. Соответствующая фенограмма полностью подтверждает вывод эксперта (рис. 4). Пробы 2 и 3 образуют единый кластер, отстоящий от пробы 1 на значительном расстоянии.



*Рис. 4. Фенограмма морфометрических данных промеров элементов бумажного носителя Документа Г.*

**Документ Д.** Из четырех листов текста Договора поручительства и трех листов с текстом Дополнительного соглашения отобрали пробы. Установлено, что бумага проб 1, 2 и 4 состоит только из целлюлозы хвойных и лиственных пород, а бумага проб 3, 5, 6 и 7, наряду с целлюлозой, содержит древесную массу хвойных и лиственных пород (Табл. 5).

**Таблица 5. Качественный и относительный количественный состав элементов бумаги тестируемого Документа Д (%% общего числа элементов)**

№ проб	Состав бумаги							
	Древесная масса хвойных и лиственных пород			Целлюлоза хвойных и лиственных пород				
	Волокна и трахеиды хвойных и лиственных пород	Фрагменты сосудов		Волокна и трахеиды лиственных пород	Трахеиды древесины хвойных пород	Сосуды		
		Осина	Береза			Осина	Береза	Липа
1	0	0	0	81,29	14,37	0,37	3,97	0
2	0	0	0	76,97	17,87	0,33	4,83	0
3	7,34	1,22	0,15	67,28	20,49	0,61	2,91	0
4	0	0	0	78,93	16,22	0,5	4,35	0
5	5,69	1,1	0,14	63,8	23,86	0,28	5,13	0
6	5,74	1,44	0,32	65,39	22,01	0,32	4,78	0
7	5,87	1	0	64,6	22,65	0,34	5,37	0,17

По заключению эксперта, пробы 5, 6 и 7 (Дополнительного соглашения) имеет одинаковый качественный и количественный состав. Среди проб 1, 2, 3 и 4 (Договора поручительства), сходны пробы 1, 2 и 4, а проба 3 отличается от них наличием элементов древесной массы. Контрольная фенограмма в целом подтверждает и уточняет заключение эксперта. Она демонстрирует два отчетливых кластера (рис. 5). Первый состоит из сходных проб 1, 2 и примыкающей к ним пробы 4, причем пробы 2 и 4 более сходны между собой, чем с пробой 1. Второй кластер состоит из проб 5, 6 и 7, и в нем более сходны между собой пробы 5 и 7. Особняком стоит проба 3, которая более сходна со вторым кластером (именно с пробой 6), но отличается от проб первого кластера.

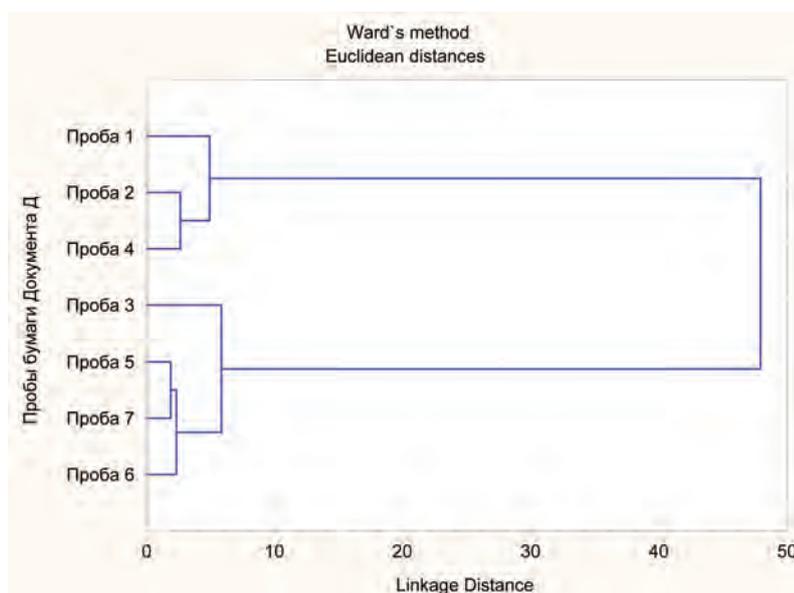


Рис. 5. Фенограмма морфометрических данных промеров элементов бумажного носителя Документа Д.

**Документ Е.** Провели сравнительное исследование бумаги этикеток ампул лекарственного средства. На исследование отобраны образцы из этикеток с 37 ампул (двух разных серий выпуска препарата) и 6 этикеток из контрольных об-

разцов ампул каждой серии (находящихся в досье на лекарственное средство) (табл. 6).

**Таблица 6. Качественный и относительный количественный состав элементов бумаги тестируемого Документа Е (%% общего числа элементов)**

Пробы	Состав бумаги			Соотношение сосудов			
	Волокна и трахеиды лиственных пород	Трахеиды древесины хвойных пород	Сосуды	Береза	Осина		
Серия 260808	Контроль 1	73,4	22,6	4,0	95,2	4,8	
	Контроль 2	74,8	22,0	3,2	95,9	4,1	
	Контроль 3	80,0	19,3	0,9	95,7	4,3	
		75,0	21,8	3,2	95,0	5,0	
		76,0	20,2	3,8	97,9	2,1	
		73,3	22,9	3,8	97,4	2,6	
		74,5	22,8	2,7	95,5	4,5	
		73,6	23,3	3,1	95,2	4,8	
		73,7	22,7	3,6	94,5	5,5	
		74,1	21,6	4,3	96,3	3,7	
		75,0	21,0	4,0	95,9	4,1	
		74,0	23,2	2,8	94,8	5,2	
		74,5	22,5	3,0	96,6	3,4	
		74,4	22,0	3,6	95,4	4,6	
		73,4	22,9	3,7	96,8	3,2	
		73,9	22,7	3,4	96,2	3,8	
		74,1	23,1	2,8	97,0	3,0	
		74,8	21,6	3,6	96,8	3,2	
	Серия 290808		74,3	22,4	3,3	94,9	5,1
			74,4	21,6	4,0	95,8	4,2
		75,7	21,3	3,0	94,7	5,3	
		73,6	23,0	3,4	97,5	2,5	
		74,2	22,4	3,4	95,6	4,4	
Контроль 4		74,6	21,9	3,5	94,8	5,2	
Контроль 5		75,5	20,9	3,6	94,2	5,8	
Контроль 6		74,5	22,4	3,1	96,7	3,3	
		75,8	21,4	2,8	96,1	3,9	
		75,1	21,8	3,1	97,0	3,0	
		76,0	20,6	3,4	95,2	4,8	
		75,6	20,5	3,9	95,3	4,7	
		76,0	20,9	3,1	95,1	4,9	
		75,4	20,9	3,7	94,2	5,8	
		76,2	20,8	3,0	95,0	5,0	
		75,7	21,0	3,3	96,7	3,3	
		74,5	22,4	3,1	96,2	3,8	
		75,3	21,0	3,7	95,6	4,4	
		76,2	20,9	2,9	97,5	2,5	
		75,7	20,4	3,9	96,2	3,8	
	75,5	20,7	3,8	95,8	4,2		
	75,1	21,4	3,5	97,4	2,6		
	75,8	21,4	2,8	94,7	5,3		
	74,5	22,4	3,1	95,4	4,6		
	74,8	21,3	3,9	94,4	5,6		

Установлено, что бумага всех исследованных образцов состоит только из целлюлозы хвойных и лиственных пород, древесная масса отсутствует. Эксперт пришел к выводу, что бумага сравниваемых объектов имеет одинаковый (качественный и количественный) волокнистый состав.

Полученные фенограммы это подтверждают (рис. 6 и 7)

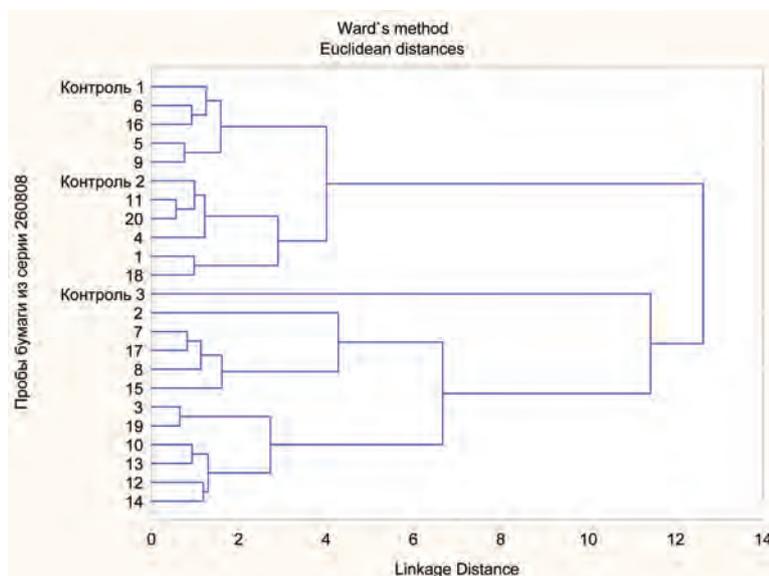


Рис. 6. Фенограмма морфометрических данных промеров элементов бумажного носителя Документа Е (серия 260808).

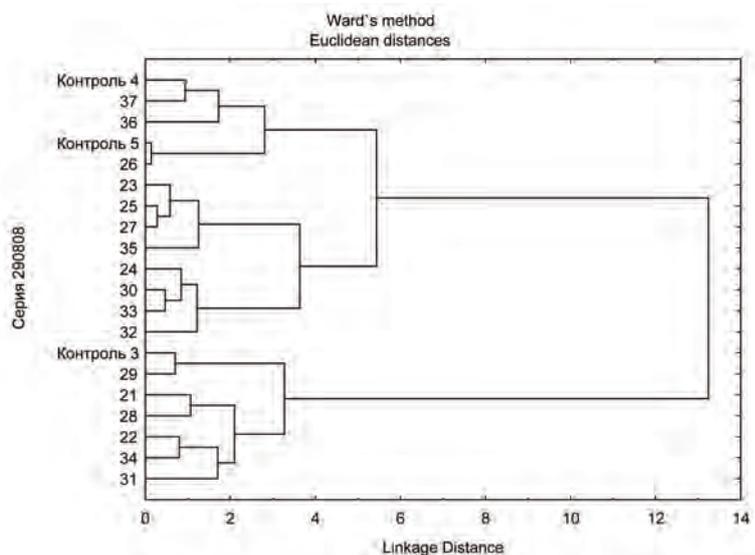


Рис. 7. Фенограмма морфометрических данных промеров элементов бумажного носителя Документа Е (серия 290808).

### Выводы

1. Применение методов многомерного статистического анализа и в частности метода группового анализа данных по методу Уорда целесообразно использовать при проведении биологической экспертизы для анализа большого массива цифровых данных.

2. Предлагаемый метод позволяет не только контролировать субъективные оценки, приводимые экспертом, но и давать более подробную характеристику сходства и различия тестируемых объектов.

### Литература

1. Бумага и картон. Методы определения состава по волокну. ГОСТ 7500-85. М. 1987. 48 с.
2. Халафян А.А. STATISTICA 6. Статистический анализ данных. 3-е изд. М.: ООО «Бином-Пресс». 2008. 512 с.
3. Чавчавадзе Е.С., Брянцева З.Е., Гончарова Е.В., Нехлюдова М.В., Горбачева Г.Н., Коржицкая З.А. Атлас древесины и волокон для бумаги. М.: Ключ. 1992. 336 с.
4. Чернова О.Ф., Перфилова Т.В., Киладзе А.Б., Сорокин П.А. Идентификация волос подвидов и гибридных форм леопарда *Panthera pardus*: алгоритм применения статистических методов // Теория и практика судебной экспертизы. 2015. № 2. с.с. 156-162.
5. Чубукова И.А. Data Mining. М.: Интернет-университет информационных технологий; Бином. Лаборатория знаний. 2008. 384 с.