ПЕДАГОГИКА
(шифр научной специальности: 5.8.7)

Научная статья
УДК 37
doi: 10.18522/2070-1403-2024-102-1-178-183

ТЕОРИЯ ОБРАБОТКИ РЕЗУЛЬТАТОВ ИССЛЕДОВАНИЙ В ОБЛАСТИ РАЗВИТИЯ ПРОСТРАНСТВЕННОГО МЫШЛЕНИЯ

© Виктор Анатольевич Нефедовский[1], Юрий Александрович Савицкий[2], Людмила Михайловна Медяникова[3]
[1, 2, 3]Краснодарское высшее военное авиационное училище летчиков имени Героя Советского Союза А.К. Серова, г. Краснодар, Россия
[1]victor_anna@mail.ru [2]yura_helen@mail.ru [3]mlm22@mail.ru

Аннотация. Рассматривается проблема обработки данных полученных при решении типовых задач при развитии пространственного мышления. Описаны примеры этапов развития пространственного мышления. Представлена структура полученных данных в ходе выполнения упражнений. Рассмотрены противоположные подходы при различных методах группировки данных. Предложены алгоритмы действий при обработке полученных результатов. Показаны подходы при анализе результатов тестирования. Произведён анализ достоинств и недостатков используемых методов анализа результатов тестирования. Сделан вывод, для получения достоверных результатов необходимо, чтобы выборка была достаточно большой, правильно подобраны признаки для анализа, чтобы они разделяли объекты на кластеры и нужно проводить анализ чувствительности к выбору метода и параметров.

Ключевые слова: пространственное мышление, пространственные образы, информация, кластерный анализ, иерархические методы.

Для цитирования: Нефедовский В.А., Савицкий Ю.А., Медяникова Л.М. Теория обработки результатов исследований в области развития пространственного мышления // Гуманитарные и социальные науки. 2024. Т. 102. № 1. С. 178-183. doi: 10.18522/2070-1403-2024-102-1-178-183

PEDAGOGY
(specialty: 5.8.7)

Original article

Theory of processing research results in the field of spatial thinking development

© Viktor A. Nefedovskiy[1], Yuri A. Savitskiy[2], Lyudmila M. Medyanikova[3]
[1, 2, 3]Krasnodar Air Force Institute for Pilots named after Hero of the Soviet Union A.K. Serov, Krasnodar, Russian Federation
[1]victor_anna@mail.ru [2]yura_helen@mail.ru [3]mlm22@mail.ru

Abstract. The article discusses the problem of processing data obtained when solving typical problems in the development of spatial thinking. Examples of the stages of spatial thinking development are described. The structure of the data obtained during the exercises is presented. The opposite approaches for different data grouping methods are considered. Algorithms of actions for processing the obtained results are proposed. Approaches to the analysis of test results are shown. The advantages and disadvantages of the methods used for analyzing test results are analyzed. It is concluded that in order to obtain reliable results, it is necessary that the sample be large enough, the signs for analysis are correctly selected so that they divide objects into clusters and sensitivity analysis to the choice of method and parameters should be carried out.

Key words: spatial thinking, spatial images, information, cluster analysis hierarchical methods.

For citation: Nefedovskiy V.A., Savitskiy Yu.A., Medyanikova L.M. Theory of processing research results in the field of spatial thinking development. *The Humanities and Social Sciences*. 2024. Vol. 102. No 1. P. 178-183. doi: 10.18522/2070-1403-2024-102-1-178-183

Введение

Пространственное мышление играет важную роль в различных сферах деятельности человека, таких как архитектура, дизайн, инженерия, география и другие. Оно помогает человеку ориентироваться в пространстве, представлять себе объекты и их отношения, анализировать их геометрические свойства и работать с различными пространственными моделями и планами.

Развитие пространственного мышления начинается с детства и происходит в тесной связи с развитием визуального восприятия и моторики. Детям предлагаются различные игры и упражнения, которые помогают им узнавать и исследовать пространственные объекты, узнавать их форму и размеры, расположение в пространстве и возможное взаимодействие с другими объектами.

Взрослые могут тренировать пространственное мышление с помощью решения логических задач, конструирования моделей или путешествий в новые места. Это помогает развивать способность видеть и анализировать пространственные отношения, делать предположения о расположении и движении объектов, а также решать задачи, требующие вычислений и прогнозирования результатов.

Пространственное мышление также может быть развито через обучение специальным методам, таким как геометрические построения, черчение и работа с трехмерными моделями. Эти методы позволяют человеку лучше понять и воспроизвести пространственные конструкции и решать задачи, связанные с их созданием и анализом.

В целом развитие пространственного мышления является важным аспектом когнитивного развития человека и помогает облегчить решение задач, связанных с ориентацией и оперированием в пространстве. Это навык, который может быть усилен и развит на протяжении всей жизни.

Обсуждение

Решение типовых задач при развитии пространственного мышления требуют определенных алгоритмов действий. Для оценки возможности развития пространственного мышления используется модифицированный тест Амтхауэра [3], состоящий из шести этапов. Первый этап – оценка начального уровня пространственного мышления. Второй этап – упражнения на развитие зрительной памяти и работы с пространственными образами. Третий этап – тренировка оперирования пространственными образами и развитие зрительной памяти. Четвертый и пятый этапы – модификации теста Амтхауэра для оценки использования алгоритмов при изменении заданий и мыслительной выносливости. Шестой этап – контрольный, повторяющий задания первого этапа для оценки эффективности тренировок и алгоритмов решения.

Эти этапы подробно описаны в наших предыдущих публикациях: «Исследование в области развития пространственного мышления будущего военного летчика» [3], «Развитие пространственного мышления как фактора, влияющего на подготовку операторов управления беспилотных летательных аппаратов» [2].

По окончании выполнения упражнений каждого этапа данные всех участников по каждому упражнению обрабатываются программой P_Collection. В результате получается таблица, в которой для каждого участника формируется отдельная строка.

Структура таблицы, получаемой программой P_Collection, будет следующей:

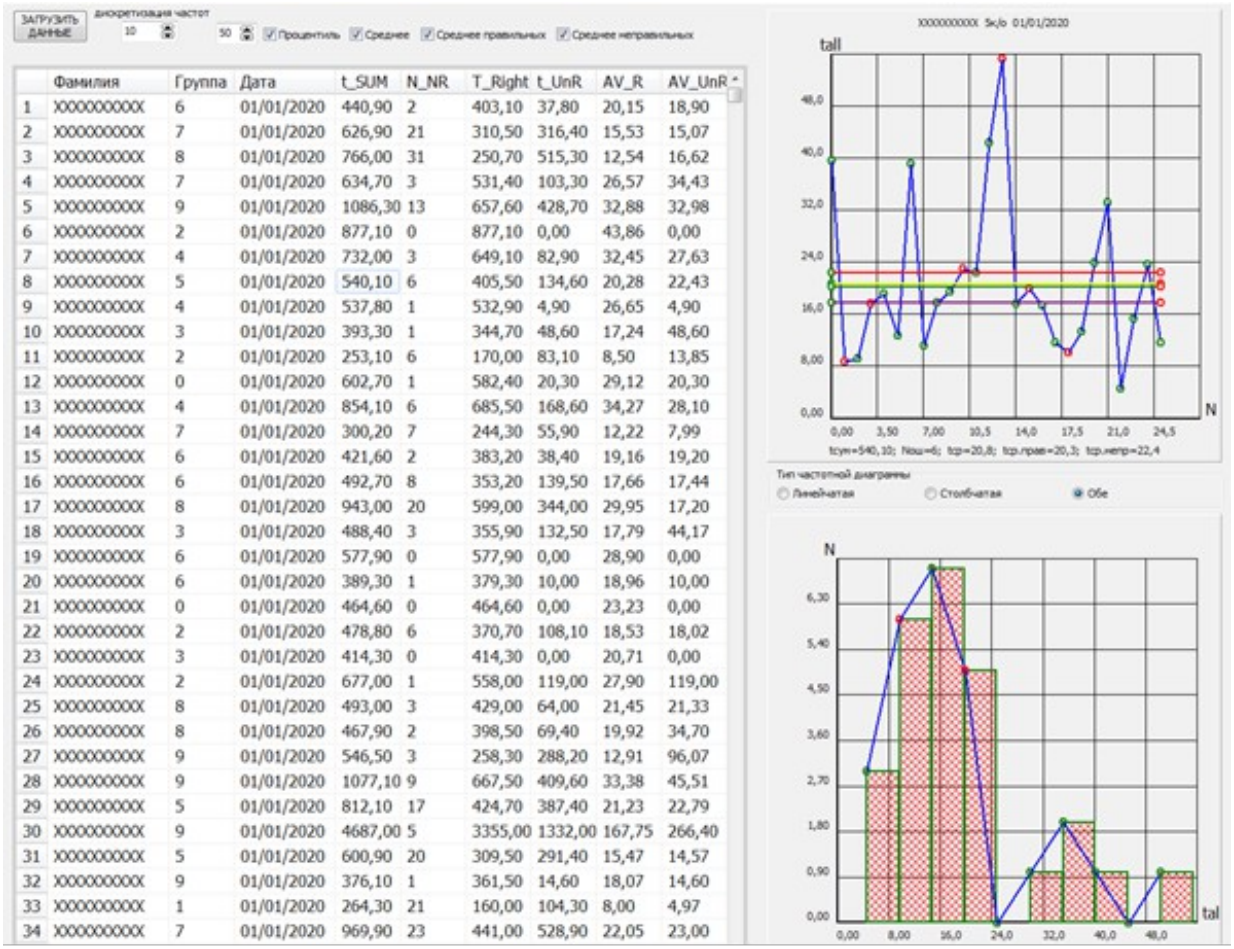
1. Имя файла – имя файла с результатами выполнения упражнения для каждого участника.
2. Фамилия – фамилия участника, указанная в имени файла.
3. Группа – группа участника, указанная в имени файла.
4. Дата выполнения – дата выполнения задания, указанная в имени файла.
5. Порядковый номер вопроса – порядковый номер вопроса из задания.
6. Порядковый номер вопроса в текущем упражнении очередного этапа – порядковый номер вопроса, так как вопросы выбираются случайным образом для каждого упражнения.
7. Действительный номер вопроса в текущем задании – действительный номер вопроса из всего задания.
8. Правильность ответа – указание правильности ответа (1 – правильный, 0 – неправильный).

9. Время, затраченное на ответ – время, затраченное участником на ответ.

Таким образом, каждая строка таблицы будет содержать информацию о выполненном вопросе одного участника. Количество строк в таблице будет равно общему количеству выполненных вопросов для всех участников. Таблица позволит удобно анализировать данные и сравнивать результаты участников.

Таблица

**Анализ выполнения упражнений каждого этапа,
данные всех участников, обработанные программой P_Collection**



	Фамилия	Группа	Дата	t_SUM	N_NR	T_Right	t_UnR	AV_R	AV_UnR
1	XXXXXXXXXX	6	01/01/2020	440,90	2	403,10	37,80	20,15	18,90
2	XXXXXXXXXX	7	01/01/2020	626,90	21	310,50	316,40	15,53	15,07
3	XXXXXXXXXX	8	01/01/2020	766,00	31	250,70	515,30	12,54	16,62
4	XXXXXXXXXX	7	01/01/2020	634,70	3	531,40	103,30	26,57	34,43
5	XXXXXXXXXX	9	01/01/2020	1086,30	13	657,60	428,70	32,88	32,98
6	XXXXXXXXXX	2	01/01/2020	877,10	0	877,10	0,00	43,86	0,00
7	XXXXXXXXXX	4	01/01/2020	732,00	3	649,10	82,90	32,45	27,63
8	XXXXXXXXXX	5	01/01/2020	540,10	6	405,50	134,60	20,28	22,43
9	XXXXXXXXXX	4	01/01/2020	537,80	1	532,90	4,90	26,65	4,90
10	XXXXXXXXXX	3	01/01/2020	393,30	1	344,70	48,60	17,24	48,60
11	XXXXXXXXXX	2	01/01/2020	253,10	6	170,00	83,10	8,50	13,85
12	XXXXXXXXXX	0	01/01/2020	602,70	1	582,40	20,30	29,12	20,30
13	XXXXXXXXXX	4	01/01/2020	854,10	6	685,50	168,60	34,27	28,10
14	XXXXXXXXXX	7	01/01/2020	300,20	7	244,30	55,90	12,22	7,99
15	XXXXXXXXXX	6	01/01/2020	421,60	2	383,20	38,40	19,16	19,20
16	XXXXXXXXXX	6	01/01/2020	492,70	8	353,20	139,50	17,66	17,44
17	XXXXXXXXXX	8	01/01/2020	943,00	20	599,00	344,00	29,95	17,20
18	XXXXXXXXXX	3	01/01/2020	488,40	3	355,90	132,50	17,79	44,17
19	XXXXXXXXXX	6	01/01/2020	577,90	0	577,90	0,00	28,90	0,00
20	XXXXXXXXXX	6	01/01/2020	389,30	1	379,30	10,00	18,96	10,00
21	XXXXXXXXXX	0	01/01/2020	464,60	0	464,60	0,00	23,23	0,00
22	XXXXXXXXXX	2	01/01/2020	478,80	6	370,70	108,10	18,53	18,02
23	XXXXXXXXXX	3	01/01/2020	414,30	0	414,30	0,00	20,71	0,00
24	XXXXXXXXXX	2	01/01/2020	677,00	1	558,00	119,00	27,90	119,00
25	XXXXXXXXXX	8	01/01/2020	493,00	3	429,00	64,00	21,45	21,33
26	XXXXXXXXXX	8	01/01/2020	467,90	2	398,50	69,40	19,92	34,70
27	XXXXXXXXXX	9	01/01/2020	546,50	3	258,30	288,20	12,91	96,07
28	XXXXXXXXXX	9	01/01/2020	1077,10	9	667,50	409,60	33,38	45,51
29	XXXXXXXXXX	5	01/01/2020	812,10	17	424,70	387,40	21,23	22,79
30	XXXXXXXXXX	9	01/01/2020	4687,00	5	3355,00	1332,00	167,75	266,40
31	XXXXXXXXXX	5	01/01/2020	600,90	20	309,50	291,40	15,47	14,57
32	XXXXXXXXXX	9	01/01/2020	376,10	1	361,50	14,60	18,07	14,60
33	XXXXXXXXXX	1	01/01/2020	264,30	21	160,00	104,30	8,00	4,97
34	XXXXXXXXXX	7	01/01/2020	969,90	23	441,00	528,90	22,05	23,00

В качестве примера на рисунке приведены графические зависимости времени, затраченного на каждую попытку, медианное значение этого времени каждой попытки, среднее значение времени попытки, среднее время правильных и неправильных ответов, частотная характеристика времени каждой попытки.

Кластерный анализ позволяет группировать участников на основе их характеристик, в данном случае – времени выполнения заданий и количества ошибок. Это помогает выделить схожие характеристики участников и классифицировать их в группы.

Количество кластеров выбирается исходя из необходимости максимального различия между группами участников и минимальных отличий внутри каждой группы. В данном случае было принято решение о четырех кластерах, так как меньшее количество кластеров не даст достаточно информации о различиях между участниками, а большее количество усложнит анализ.

Цель анализа заключается в нахождении влияющих факторов, которые помогут объяснить различия между группами участников. Анализ также может выявить участников, не входящих в какую-либо группу, что также важно для дальнейшего анализа.

Итак, общая задача анализа заключается в том, чтобы разбить множество из k участников на m кластеров (k, m – целые, $k>m$) таким образом, чтобы каждый участник принадлежал только одному кластеру с однородными признаками, и чтобы объекты из разных кластеров имели разнородные признаки. Задача выделения кластеров по показателям расстояния между признаками в группах участников решается выполнением следующих условий:

$$\sum_{i=1}^{k} = \sum_{j=1}^{k} \delta_{ij} d_{ij} \rightarrow min ; \quad (1)$$

$$\sum_{i=1}^{k} = \sum_{j=1}^{k} (1-\delta_{ij}) d_{ij} \rightarrow max , \quad (2)$$

где k – количество объектов; d_{ij} – расстояние между i и j объектами; δ_{ij} – символ Кронекера; $\delta_{ij}=1$, если i и j объекты входят в один и тот же кластер, и $\delta ij=1$, если они входят в разные кластеры.

Признаки могут представляться в натуральных единицах измерения, в нормализованной или стандартизированной форме. Представление признаков в естественных единицах измерения наименее желательно, так как натуральные величины всех признаков могут существенно отличаться и признак с наибольшей натуральной величиной будет оказывать наибольшее влияние на распределение объектов по кластерам, а влияние на результат признака с наименьшим значением будет минимальным. Способов разбиения (кластеризации) очень много, и далеко не все из них приспособлены для автоматизации. Наиболее используемыми являются иерархические (агломеративные или дивизимные) и итеративные методы группировки. Иерархические методы могут использовать два противоположных подхода.

При первом подходе (агломеративном) все объекты считаются самостоятельными кластерами и в процессе работы алгоритма наиболее похожие кластеры объединяются. Наиболее известными представителями таких алгоритмов являются односвязные, полно связные алгоритмы, алгоритмы взвешенных и невзвешенных центроид, алгоритм Уорда [4]. Все эти алгоритмы отличаются способом объединения кластеров и полученным результатом, преимуществами и недостатками. Например, односвязный алгоритм реализуется наиболее быстро, однако в результате того, что группировка кластеров производится на основании только одной связи, кластеры получаются слишком большие и «продолговатые». При использовании алгоритма Уорда наоборот, используется большое число признаков и кластеры группируются по признаку наименьшего вклада в функцию качества. При этом все кластеры имеют форму гиперсфер и приблизительно равный размер.

При втором подходе (дивизимном) все исходные объекты считаются принадлежащими одному кластеру, который на последующих шагах делится на меньшие, и в результате образуется последовательность расщепляющихся групп.

Иерархические методы используются при сравнительно небольших наборах данных и признаков, основным их достоинством является наглядность. При увеличении количества исследуемых объектов или связей между ними иерархические методы теряют свою эффективность, а в некоторых случаях могут стать непригодными. В этом случае целесообразно применять итеративные методы разбиения, которые предполагают разделение набора данных на заранее заданное количество кластеров и последовательное их формирование до тех пор, пока не будет выполнено правило остановки.

Наиболее известным итеративным методом группировки является алгоритм k-средних. При использовании этого алгоритма количество кластеров считается предварительно заданным (произвольно либо с использованием таблицы Фортьера – Соломона) [1]. Каждый объект относится к тому кластеру, расстояние до которого минимально. На нулевом шаге такого алгоритма в качестве центров кластеров выбирают или случайно выбранные точки, или точки, расстояние между которыми максимально, и каждому кластеру присваивают единичный вес. На следующем шаге каждую точку относят к тому кластеру, расстояние до которого минимально, и затем пересчитывается вес каждого кластера и его центр тяжести. На всех последующих ша-

гах этот алгоритм повторяется либо до получения стабильной конфигурации кластеров, либо до достижения заранее заданного количества итераций. Недостатком такого алгоритма является существенное изменение конфигурации кластеров при появлении точек-«выбросов», однако простота и скорость реализации этого алгоритма нивелирует его недостатки.

В качестве расстояний могут быть использованы следующие величины:

Евклидово расстояние

$$d_{ij} = \sqrt{\sum_{n=1}^{k} (x_{in} - x_{jn})^2} \quad (3)$$

или квадрат Евклидова расстояния

$$d_{ij} = \sum_{n=1}^{k} (x_{in} - x_{jn})^2 \quad (4)$$

Такое расстояние чаще всего используется для стандартизированных данных или данных, измеряемых в одних и тех же единицах измерения. Считается, что оптимальным является использование не стандартизированных данных, так как при появлении точек-«выбросов» расстояние между объектами не изменяется.

Манхеттенское расстояние (расстояние городских кварталов, хеммингово расстояние), применяется для номинальных или качественных переменных.

$$d_{ij} = \sum_{n=1}^{k} |x_{in} - x_{jn}| \quad (5)$$

Расстояние Чебышева

$$d_{ij} = max |x_{in} - x_{jn}| \quad (6)$$

Такой способ измерения расстояния применяется тогда, когда необходимо определить два объекта как различные, если они сильно отличаются только по одному параметру.

Расстояние Минковского (обобщенное степенное расстояние)

$$d_{ij} = \sqrt[r]{\sum_{(n=1)}^{k} |x_{in} - x_{jn}|^p} \quad (7)$$

Такое измерение расстояния используется в случае, когда возникает необходимость прогрессивно уменьшить или увеличить вес, относящийся к параметру, по которому соответствующие объекты сильно отличаются. В этом выражении r и p – определенные значения параметров (r,p = 1…4), определяющих «поведение» метрики (расстояния) Минковского [1].

Исходное разбиение будет производиться с использованием метода k-средних, в качестве метрики (расстояния) будет использоваться Евклидово расстояние, начальное количество кластеров равно 4. Перед началом разбиения будет производиться стандартизация параметров. Разбиение будет проводиться отдельно для каждого упражнения. Выбор начальных центров кластеров будет производиться по максимальному расстоянию между ними. Результат будет считаться полученным по выполнению 50 итераций. В качестве независимых переменных выбирается общее время выполнения упражнения, количество ошибок, среднее время правильных и среднее время неправильных ответов. Все остальные переменные (время, затраченное на правильные и неправильные ответы) хоть и не являются линейно зависимыми от указанных, но включение их в расчет разбиения приведет к только к уменьшению точности разбиения.

Выводы

Эти способы деления объектов на кластеры широко применяются в различных областях, таких как маркетинг, медицина, финансы, сетевой анализ и другие. Кластерный анализ позволяет выделить группы объектов с похожими характеристиками или поведением, что может быть полезно для понимания особенностей данных и принятия решений. Существуют также методы, позволяющие оценить качество полученных кластеров, например, с использованием внутренних и внешних критериев сходства.

Однако при применении кластерного анализа необходимо учитывать некоторые ограничения. Во-первых, выборка должна быть достаточно большой и репрезентативной, чтобы гарантировать получение достоверных результатов. Во-вторых, необходимо правильно выбрать признаки для анализа, чтобы они явно разделяли объекты на кластеры. В-третьих, результаты кластерного анализа могут зависеть от выбранного метода и параметров алгоритма, поэтому важно проводить анализ чувствительности к выбору метода и параметров.

Список источников

1. *Воронцов К.В*. Алгоритмы кластеризации и многомерного шкалирования // Курс лекций. МГУ, 2007.
2. *Нефедовский В.А., Мальчинский Ф.В., Савицкий Ю.А., Терехов В.В.* Развитие пространственного мышления как фактора, влияющего на подготовку операторов управления беспилотных летательных аппаратов // Сборник научных статей Всероссийского форума с международным участием «Академические Жуковские чтения» III Всероссийской научно-практической конференции «Беспилотная авиация: состояние и перспективы развития» проводимой в ВУНЦ ВВС ВВА г. Воронеж. 23–25 ноября 2021 г.
3. *Нефедовский В.А., Савицкий Ю.А., Терехов В.В.* Исследование в области развития пространственного мышления будущего военного летчика // Гуманитарные и социальные науки 2021. № 2.
4. *Райзин Д.В.* Классификация и кластеризация. М.: Мир, 1980. 390 с.
5. *Ясукова Л.А.* Тест структуры интеллекта Р. Амтхауэра // Методическое руководство СПБ.: ГП «Иматон». 2002. 78 с.

References

1. *Vorontsov K.V.* Algorithms of clustering and multidimensional scaling // Course of lectures. Moscow State University, 2007.
2. *Nefedovsky V.A., Malchinsky F.V., Savitsky Yu.A., Terekhov V.V.* Development of spatial thinking as a factor influencing the training of unmanned aerial vehicle control operators // Collection of scientific articles of the All-Russian Forum with international participation “Academic Zhukovsky Readings” of the III All-Russian Scientific and Practical Conference "Uroneless Aviation: State and Development Prospects" held at the All-Russian Scientific and Technical Center of the Air Force VVA. Voronezh. November 23–25, 2021.
3. *Nefedovsky V.A., Savitsky Yu.A., Terekhov V.V.* Research in the field of development of spatial thinking of the future military pilot // Humanities and Social Sciences 2021. No. 2.
4. *Raizin D.V.* Classification and clustering. M.: Mir, 1980. 390 p.
5. *Yasukova L.A.* Intelligence structure test R. Amthauer // Methodological guide. SPB: SE Imaton. 2002. 78 p.

Статья поступила в редакцию 24.11.2023; одобрена после рецензирования 15.12.2023; принята к публикации 15.12.2023.
The article was submitted 24.11.2023; approved after reviewing 15.12.23; accepted for publication 15.12.23.