

УДК 004.6, 004.9, 004.827, 004.824

## ОЦЕНКА ИНФОРМАТИВНОСТИ И КАЧЕСТВА ОБРАБОТКИ ГРАФИЧЕСКОЙ ИНФОРМАЦИИ

Цветков В.Я.

*Московский государственный технический университет радиотехники, электроники  
и автоматики МГТУ МИРЭА, Москва, e-mail: cvj2@list.ru*

Статья описывает методику оценки и результаты эксперимента по оценке качества обработки графической информации. вводится понятие графических информационных единиц. Раскрывается содержание параметров применяемых для оценки информационной коллекции. Описан пример, поясняющий принципы, методику обработки и смысл коэффициентов оценки. Показано различие между репрезентативностью обработанной информации и ее информативностью. Репрезентативность связана с объемом информационной единицы. Информативность связана со смысловым содержанием информационной единицы. введено понятие информационной устойчивости обрабатываемой информации. Показано, что сложные информационные единицы являются информационно более устойчивыми чем простые информационные единицы. В тоже время сложные информационные единицы имеют более высокий коэффициент искажений в сравнении с простыми информационными единицами.

**Ключевые слова:** обработка информации, графическая информация, информативность, информационная устойчивость, информационные единицы, графические информационные единицы, информационные коллекции

## SOLUTION TASKS OF THE SECOND KIND WITH THE USE OF THE INFORMATION APPROACH

Tsvetkov V.Y.

*Moscow State Technical University of Radio Engineering, Electronics and Automation  
MSTU MIREA, Moscow, e-mail: cvj2@list.ru*

This article describes the solution of complex problems, which are called tasks of the second kind. Description of the method covers the area of decision-making in the management area and mathematical solutions of problems. The article reveals the content of the tasks of the first kind. Given their topological representation. The content of the tasks of the second kind. Given their topological model. Shows the difference between the objectives and tasks of the first kind of the second kind. Solution of the problem of the second kind on the basis of the known ant algorithm. We introduce a modification to the algorithm. Modification involves the transition from many points of the search for solutions to the same. Modification involves obtaining resources for incremental stages of the decision and the accumulation of experience in dealing with. The experience is stored in the database. This model can be used to search for mathematical and management decisions

**Keywords:** problem solving, the ant algorithm, tasks of the second kind, the complexity

Обработка графической информации применяется в разных направлениях. Рассмотрим графическую информацию, которая получается при обработки графики с помощью сканирующих устройств и последующем переводе растровой информации в векторную. Эта задача часто решается в геоинформатике при цифровании карт и при обработке данных дистанционного зондирования и при формировании геоданных [1]. На практике применяют специализированные программы векторизаторы [2], которые решают большую часть задачи преобразования информации. Эти программы имеют специальные фильтры векторизации, которые настраивают процесс распознавания на определенный тип информации [3, 4], например чертежи, текст, деловая графика географические карты и т.д. Процесс векторизации является процессом трансформации информации [5] из одной формы в другую. На практике важна семантическая составляющая такого процесса [6]. Это касается во-

просов информативности. Существующая практика оценки информативности часто связана с использованием энтропии и оценкой объема информационного объекта до и после обработки. В данной постановке задача информативности определяется по числу интерпретируемых объектов по отношению к исходным объектам.

**Основная часть.** Для проверки качества векторизации использовались четыре информационные коллекции. Информационная коллекция представляет собой совокупность однотипных объектов, которые легко сравнивать между до и после обработки. Эта совокупность позволяет набирать статистику.

Первая информационная коллекция включает регулярную сетку из 200 крестов с расстояниями между крестами 10мм и размером креста по вертикали и горизонтали 1,5 мм. Вторая информационная коллекция представляет собой совокупность 200 горизонтальных отрезков размером 1,5 мм. Третья информационная коллекция

представляет собой совокупность 160 вертикальных отрезков размером 1,5 мм. Четвертый информационная коллекция представляет собой совокупность 160 окружностей диаметром 3 мм.

Использовались следующие параметры:  $N$  – количество объектов на входе до обработки;  $N_i$  – количество искаженных интерпретируемых объектов;  $N_l$  – количество потерянных объектов;  $N_o$  – количество неискаженных объектов на выходе;  $N_{ni}$  – количество искаженных не интерпретируемых объектов;

Введены следующие коэффициенты  $K_l$  – коэффициент потерь;  $K_{inf}$  – коэффициент информативности;  $K_d$  – коэффициент искажений;  $K_{int}$  – коэффициент интерпретируемости искаженных объектов;  $K$  – коэффициент репрезентативности.

$$K_l = N_l / N; K_{inf} = (N_i + N_o) / N; K_d = (N_i + N_{ni}) / N;$$

$$K_{int} = N_i / (N_i + N_{ni}); K = N_o / N.$$

Объект характеризуется количеством элементов и «характерной точкой». В качестве такой точки для симметричных объектов может быть центр симметрии. Для несимметричных объектов такой точкой может быть центроид. Объект считается искаженным, при изменении количества элементов в нем и при изменении положения характерной точки относительно других точек объекта. По существу речь идет о смещении точки. На рис. 1 приведены примеры характерных искажений. Характерная точка выделена штриховкой.

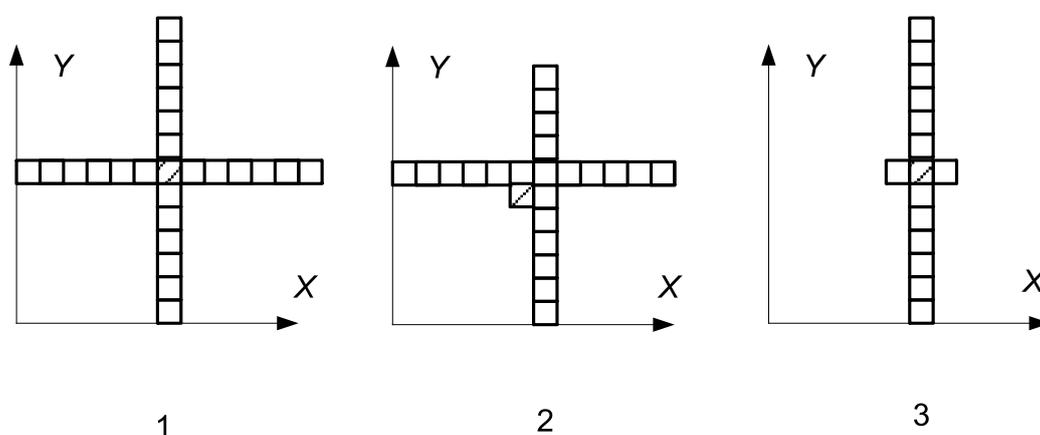


Рис. 1. Исходный объект и его искажения

Объект показан в целочисленных координатах, которые применяют при обработке растровых изображений. На рис. 1 под цифрой 1 показан исходный объект. Координаты его характерной точки в условной системе координат, показанной на рисунке равны  $X = 7, Y = 7$ .

Цифрой 2 показан объект с потерей двух элементов. Эта потеря приводит к смещению характерной точки. Координаты характерной точки в этом случае равны  $X = 6, Y = 6$ . Напомним, что в целочисленных системах координат нет дробной части. Это тип искажения называют асимметричным искажением.

Под цифрой 3 показан характерный объект, полученный при сканировании, когда направление сканирования или направление векторизации направлено вдоль оси  $Y$ . В этом случае иногда происходит потеря элементов на линиях перпендикулярных направлению сканирования. Симметрия сохраняется, но потеря элементов налицо. Это

тип искажения называют симметричным искажением.

На рис. 2 приведена информационная ситуация [7, 8], характеризующая фрагмент первой информационной коллекции. Стрелкой показано направление сканирования или векторизации.

Информационные коллекции строятся по матричному типу. Это упрощает процедуру их кодирования и анализа. Цифры располагаются в порядке типичном для целочисленной системы координат. Таким образом, по фрагменту, характеризующему результаты обработки на рис. 2, можно сделать следующие выводы. Используем матричную кодировку объектов. Объект 1,7 потерян. Объекты 2,2; 1,3; 3,4; 2,5 – искажены. Все остальные объекты неискажены. Количество искаженных интерпретируемых объектов  $N_i = 2$ . Это объекты 2,2 и 1,3. количество искаженных не интерпретируемых объектов  $N_{ni} = 2$ . Это объекты 3, 4 и 2, 5.

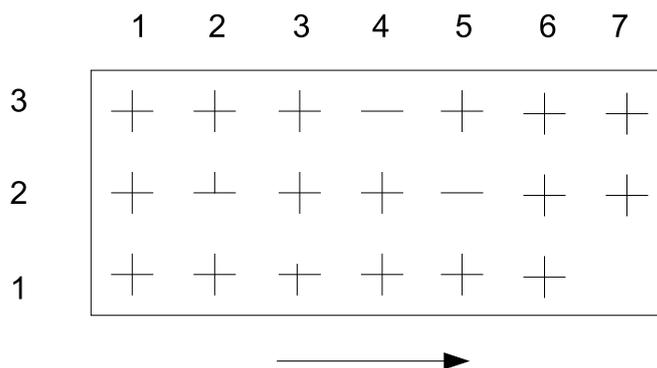


Рис. 2. Фрагмент информационной коллекции после обработки

Для информационной ситуации, показанной на рис. 2, имеем результаты, приведенные в табл. 1.

**Таблица 1**  
Результаты оценки качества и информативности по макету

Качества	Количественные характеристики
N	21
N <sub>1</sub>	1
N <sub>i</sub>	2
N <sub>ni</sub>	2
N <sub>o</sub>	16
K	0.76
K <sub>1</sub>	0.05
K <sub>d</sub>	0.19
K <sub>inf</sub>	0.86

Результаты полного исследования на четырех коллекциях приведены в табл. 2. Следует отметить, что кроме сканирования существует ряд алгоритмов обработки информации, которые также обрабатывают информацию построчно, это программы векторизаторы и программы построения цифровых моделей по методу развешивающей линии (sweep-line algorithm) [9]. Поэтому результаты исследования и методика охватывают сканерную обработку и алгоритмическую обработку по тем же принципам.

**Обсуждение.** Коэффициент репрезентативности является основной характеристикой качества обработки информации. Он показывает, какая часть объектов сохраняет свои семантические и метрические свойства. Он же показывает, какая часть объектов сохраняет свой информационный объем. Коэффициент информативности является основной семантической характеристикой обработки информации. Он показывает, какая часть информации сохраняет содержательность, безотносительно

к искажениям. Коэффициент интерпретируемости искаженных объектов является условной характеристикой. Он показывает, какая часть искаженных объектов сохраняет семантические свойства.

**Таблица 2**  
Результаты оценки качества обработки на 4-х информационных коллекциях

	Кол 1	Кол 2	Кол 3	Кол 4
N	200	200	160	160
N <sub>1</sub>	22	28	0	0
N <sub>i</sub>	12	0	14	80
N <sub>ni</sub>	12	16	16	8
N <sub>o</sub>	77	72	65	72
K	0.77	0.72	0.81	0.45
K <sub>1</sub>	0.11	0.14	0.0	0
K <sub>d</sub>	0.12	0.08	0.23	0.55
K <sub>inf</sub>	0.83	0.72	0.9	0.95
K <sub>int</sub>	0.5	0.0	0.47	0.91

Рассмотренные элементарные объекты в теории САПР называют «графические примитивы». В семантической теории информации такие объекты являются информационными единицами [10, 11], а для графической информации «графическими информационными единицами».

Обобщенно данная методика сводится к следующему. Выбираются однородные наборы графических информационных единиц. Задается структура их расположения (локализация), что позволяет проводить дополнительный апостериорный контроль. Эти наборы образуют структурированную систему информационных единиц, или информационную коллекцию, которая пропускается через систему обработки (программа, устройство). Выходные данные поэлементно сравнивают с входными данными и вычисляют коэффициенты.

Данная методика применима для информационных коллекций с разным числом

информационных единиц. Она позволяет проводить сравнение результатов обработки независимо от количества единиц.

Результаты, приведенные в табл. 2, дают основание предполагать, что направление сканирования или обработки объектов влияет на качество обработки и информативность. Результаты, приведенные в таблице, дают основание предполагать, что существуют пространственные отношения [12] между направлением сканирования и качеством обработки информации. Результаты, приведенные в таблице, дают основание предполагать, что морфологически сложные информационные единицы информативно более устойчивы к преобразованиям, хотя менее устойчивы к искажениям, чем более простые информационные единицы. Это обусловлено тем, что в сложных информационных единицах существуют дополнительные связи между элементами этих единиц, которые повышают информационную устойчивость при обработке.

Результаты исследований дают основание ввести параметр «информационная устойчивость». Этот параметр определяет степень сохранения семантики информационной коллекции, информационной модели или информационной единицы при обработке. Этот параметр характеризует информационную коллекцию с одной стороны, с другой стороны характеризует систему обработки. Информационная устойчивость сложных информационных единиц выше за счет дополнительных связей, которые играют роль резервирования. Средство обработки информации имеет более высокое качество по сравнению с подобными, если оно обеспечивает высокую информационную устойчивость по сравнению с другими средствами на тех же информационных коллекциях.

### Выводы

Данная методика применима для исследования различных графических образов в системах обработки. Она позволяет получать характеристики информационных коллекций и характеристики средств обработки.

Результаты представляют интерес не только для технических приложений по физической обработке информации, но и для теоретических исследований в области пространственного знания. Информативность в данной методике, оценивается не по информационному объему, а по числу интерпретируемых объектов после обработки. Это основано на семантической теории информации. Исследование целесообразно расширить в область пространственного знания [13] и искусственного интеллекта [14].

### Список литературы

1. Савиных В.П., Цветков В.Я. Геоданные как системный информационный ресурс // Вестник российской академии наук. – 2014. – Т. 84, № 9. – С. 826–829.
2. Темнов К.А. Повышение эффективности векторизации конструкторской документации электронных средств // Вестник Московского авиационного института. – 2010. – Т. 17, № 3. – С. 27.
3. Цветков В.Я. Классификация и типизация в геоинформатике // Геодезия и аэрофотосъемка. – 2004. – № 2. – С. 80–87.
4. Замулин А.В. Типы данных в языках программирования и базах данных. – Изд-во «Наука», Сибирское отделение, 1987.
5. Косс В.А. Модель трансформации информации в цикле управления сложной системы // Математические машины и системы. – 2005. – № 4. – С. 39–48.
6. Цветков В.Я. Семиотический подход к построению моделей данных в автоматизированных информационных системах // Геодезия и аэрофотосъемка. – 2000. – № 5. – С. 142–145.
7. Соловьев И.В. Применение модели информационной ситуации в геоинформатике // Науки о Земле. – 2012. – № 01. – С. 54–58.
8. Tsvetkov V.Ya. Information Situation and Information Position as a Management Tool // European Researcher. – 2012. – Vol.(36), № 12-1. – P. 2166–2170.
9. Maiorov A., Kong N. Effective algorithm for constructing a Delaunay triangulation // Izvestia vuzov. Geodesy and aerophotography. – 2011. – № 1. – P. 101–108.
10. Цветков В.Я. Информационные единицы как средство построения картины мира // Международный журнал прикладных и фундаментальных исследований. – 2014. (Часть 4) – № 8 – С. 36–40.
11. Tsvetkov V.Ya. Information Units as the Elements of Complex Models // Nanotechnology Research and Practice. – 2014. – Vol.(1), № 1. – P. 57–64.
12. Майоров А.А., Цветков В.Я. Геоференция как применение пространственных отношений в геоинформатике // Известия высших учебных заведений. Геодезия и аэрофотосъемка. – 2012. – № 3. – С. 87–89.
13. Benjamin Kuipers. Modeling Spatial Knowledge (1978) // Cognitive Science – № 2. – P. 129–153.
14. Antony Galton. Spatial and temporal knowledge representation // Earth Science Informatics, September, 2009, Volume 2, Issue 3, P. 169–187.