BIOSCIENCES BIOTECHNOLOGY RESEARCH ASIA, November 2014. Vol. **11**(Spl. Edn.), p. 211-213

Assessment of Quality of Vectorization Based Information

V.Y. Tsvetkov

Professor, Doctor of technical sciences, Moscow State University of Radio Engineering, Electronics and Automation (MSTU MIREA), 119454, Russia, Moscow, Prospect Vernadskogo st., 78, MIREA.

doi: http://dx.doi.org/10.13005/bbra/1464

(Received: 27 September 2014; accepted: 10 October 2014)

The presented paper discusses results of the experiment on an assessment of quality of information, derived from raster to vector conversion. Contents of the test object are disclosed. Three test object of different qualitative content are discussed. An assessment of information quality was carried out on a basis of a comparison with initial information with information obtained after vectorization. Coefficients are introduced which describe quality of information after processing.

Key words: Image processing, image scanning, Information vectorization, Assessment of quality of information

During image processing[1, 2] vectorization, which is a procedure of a conversion of raster information into vector information, is of importance[3]. It is used in design[4], geo-informatics, where it serves as a basis for a formation of Geodata[5]. In practice, specialized software for vectorization[6] is implemented, which solves the major part of the problem of information conversion. That software use special filters[7], which set up the process of recognition for a certain type of information[8], such as drawings, text, business graphics, geographical maps, etc. The vectorization is a process of a transformation of information[9] from one form to another. In practice a semantic component of the process plays an important role[10]. The result of information processing using such software is not always unambiguous. Furthermore, because of the specific characteristics of an algorithm, it is possible to identify some vectorization's systematic errors. Special testing allows to identify and to take into account such errors. Special test objects are used for that.

Methodology

In order to test vectorization's quality, three test objects were used. The test object is an object, which semantic and metric contents are known prior to processing. Such an object allows to compare processing result with initial information and to make a conclusion about processing quality basing on that. An equivalent of test objects are bench marks[11, 12], which are used for testing of software's performance. The test object includes a set of primitive object which define statistics. Using that statistics, it is possible to conduct statistical estimations characterizing processing quality.

The first test object was an aggregate of primitive objects for recognition, which were forming a regular grid of 100 crosses with 10mm vertical distance and 1.5 mm horizontal distance between them. The second test object was an aggregate of 100 horizontal segments with 1.5 mm

* To whom all correspondence should be addressed.
E-mail: cvj2@mail.ru

size. The third test object was an aggregate of 80 vertical segments with 1.5 mm size.

The following parameters were used: N – number of input objects before processing; N_i – number of distorted interpreted objects; N_l – number of lost objects; N_{on} – number of output non distorted objects; N_{ni} – number of distorted not interpreted objects;

Following coefficients were introduced K_l – coefficient of losses; K_{inf} – coefficient of informativeness; K_d – coefficient of distortion; K_{int} – coefficient of interpretability of distorted objects; – coefficient of representativeness.

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

The object was considered as a distorted, if there was a change of number of elements in it and if its center of symmetry shifted relatively to an initial position.

RESULTS

Software scan was carried out in a vertical direction. Sweepline algorithm was implemented[13, 14]. The results are presented in the table.

Table 1. Vectorization results in Vectory 3.1 software with three test objects

	Test 1 100	Test 2 100	Test 3 80
N_l	11	14	0
N_i	6	0	7
N_{ni}	6	8	8
N_o	77	72	65
κ	0.77	0.72	0.81
K_l	0.11	0.14	0.0
K_d	0.12	0.08	0.23
K_{inf}	0.83	0.72	0.9
K_{int}	0.5	0.0	0.47

DISCUSSION

Coefficient of representativeness is the main characteristic of quality of processed information. It shows, which part of objects remains its semantic and metric properties.

That procedure for quality assessment of information is applicable for test objects with different number of recognizable objects. It is applicable for an assessment of information quality not only software, but also scanning devices.

The results, presented in the table, allow to presume, that objects' scan direction affects quality of information. The results presented the table suggest, that there are spatial relationships between scan direction and scan quality of converted information. The results are of interest not only for engineering applications, but also for theoretical research in a field of spatial knowledge. Informativeness is determined not on a basis of data volume, but by a number of interpreted objects after processing. This is semantic approach.

CONCLUSION

An expansion of types of recognition objects in a form of other shapes, such as circles, is of certain interest. Moreover, recognition objects, which were formed as crossing of primitive objects or their tangential junction, are also interesting. Is is reasonable to expand research area to spatial knowledge[15, 16], which is investigated using artificial intelligence methods[17].

REFERENCES

1. Jensen, J.R. *et al.*, Introductory digital image processing: a remote sensing perspective. Second edition. Prentice-Hall Inc, 1996.
2. Takeda, H., S. Farsiu and P. Milanfar, Kernel regression for image processing and reconstruction. Image Processing, IEEE Transactions on, 2007; **2**(16): 349-366.
3. Congalton, R.G., Exploring and evaluating the consequences of vector-to-raster and raster-to-vector conversion. *Photogrammetric Engineering and Remote Sensing*, 1997; **4**(63): 425-434.
4. Gaidar', O.G., Vektorizatziya rastrovogo ozobrazheniya v sisteme KOMPAS [Vectoring of raster image in COMPAS CAD system]. Donetsk: GRAFIK, DonNTU, 2006: 28.
5. Savinykh, V.P. and V.Ya. Tsvetkov, Geodata as a Systemic Information Resource. *Herald of the Russian Academy of Sciences*, 2014; **5**(84): 365-368.
6. Aniskov, A.V., O.N. Ponomarenko and V.A. Fisun, Konvertori-vektorizatori [Vectorizer-convertors]. Yaziki i parllel'nie EVM. Algoritmi i algorimicheskie yaziki. [Languages and parallel

computers. Algorithms and algorithmic languages], 1990; 38-61.
7. Bier, E.A., Method and apparatus for composing layered synthetic graphics filters. Patent 6072501 USA, 2000.
8. Csisz, I. *et al.*, Information-type measures of difference of probability distributions and indirect observations. *Studia Sci. Math. Hungar,* 1967; **2**: 299-318.
9. Koss, V.A., Model' transformatzii informatzii v tzikle upravlenia slozhnoi sistemi [Model of information's transformation in a control cycle of a complex system]. Matematicheskie mashini i sistemi [Mathematical machines and systems], 2005; 4.
10. Tsvetkov, V.Ya., Semioticheski podhod k postroeniyu model'ei dannih v avtomatizirovannih informatzionnih sistemah [Semiotic approach to a creation of data models in automatized information systems]. Geodeziya i aerofotosemka [Geodesy and aerophotography], 2000; **5**: 142-145.
11. Bailey, D.H. *et al.*, NAS parallel benchmark results. Proceedings of the 1992 ACM/IEEE conference on Supercomputing. IEEE Computer Society Press, 1992; 386-393.
12. Gray, J., Benchmark handbook: for database and transaction processing systems. Morgan Kaufmann Publishers Inc, 1992.
13. Maiorov, A. and N. Kong, Effective algorithm for constructing a Delaunay triangulation. Izvestia vuzov. *Geodesy and aerophotography,* 2011; **1**: 101-108.
14. Jin, L. *et al.,* A sweepline algorithm for Euclidean Voronoi diagram of circles. *Computer-Aided Design*, 2006; **3**(38): 260-272.
15. Kuipers, B., Modeling Spatial Knowledge. *Cognitive Science,* 1978; **2**: 129-153.
16. Tsvetkov, V.Ya., Prostrtanstvennie znania: formirovanie i predstavlenie [Spatial knowledge: formation and representaion]. Germany, Saarbrücken: LAP LAMBERT Academic Publishing GmbH & Co. KG, 2013; 107.
17. Galton, A., Spatial and temporal knowledge representation. *Earth Science Informatics*, 2009; **3**(2): 169-187.