IOP Conf. Series: Materials Science and Engineering **177** (2017) 012047 doi:10.1088/1757-899X/177/1/012047

New Methods of Three-Dimensional Images Recognition Based on Stochastic Geometry and Functional Analysis

N G Fedotov[1], A V Moiseev[2], A A Syemov[1], V G Lizunkov[3] and A Y Kindaev[2]

[1] Penza State University, 40, Krasnaya Street, Penza, 440017, Russia
[2] Penza State Technological University, 1a/11, Baidukova Passage/Gagarina Street, Penza, 440039, Russia
[3] Yurga Institute of Technology, TPU Affiliate, 26, Leningradskaya Street, Yurga, 652055, Russia

E-mail: vladeslave@rambler.ru

Abstract. A new approach to 3D objects recognition based on modern methods of stochastic geometry and functional analysis is proposed in the paper. A detailed mathematical description of the method developed on the approach is also presented. The 3D trace transform allows creating an invariant description of spatial objects, which better resist distortion and coordinate noise than the one, obtained as a result of the object normalization procedure, does. The ability to control properties of developed features increases intellectual capacities of the 3D trace transform significantly, which can be mentioned as its undeniable advantage. The justification of the proposed theory and mathematical model is a variety of worked out theoretical examples of hypertriplet features that have particular described properties. The paper considers in detail scan techniques of the hypertrace transform and its mathematical model as well as approaches to developing and distinguishing informative features.

1. Introduction

Current progress in science and technology has resulted in an annually growing number of people involved in information processing. One of the most important tasks connected with information systems creation is automating the process of pattern recognition. Ample researches are being conducted to perform the task and among them there are those that help us to try and to understand one of the main abilities of a human brain – the ability to recognize [1-3].

They develop and generate decisive prerequisites for constructing intelligence systems.

It is worth noting that imitating human brains activity is not the only method of creating such systems [4-8]. Technology has its own ways of solving the problem, the ways that do not copy natural abilities of a human being technically, but take into account merits and particular properties of natural visual systems. The paper discusses some of the above-mentioned ways using stochastic geometry and functional analysis.

Here is the description of the method mathematical structure with extensive comments where they are necessary.

2. Techniques of 3D image scanning and choice of scanning element

Let F denote a source model of a 3D image. We define plane $B(\eta, r) = \{x \mid x^T \cdot \eta = r\}$ as the tangent to the sphere that has the center at the origin of coordinates and radius r, the plane passes through given point X at distance r from the origin of coordinates with given angles ω and φ, where $\eta = [\cos\omega \cdot \sin\varphi,\ \sin\omega \cdot \sin\varphi,\ \cos\varphi]$ is the unit vector in R^3, ω is the angle between axis 0x and the projection of interval OX on plane 0xy, φ is the angle between axis 0z and interval OX (fig. 1).

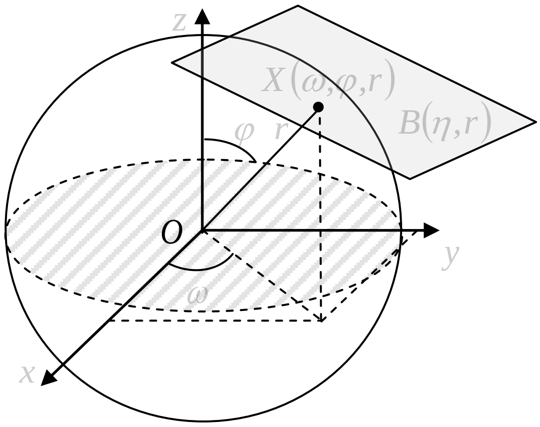
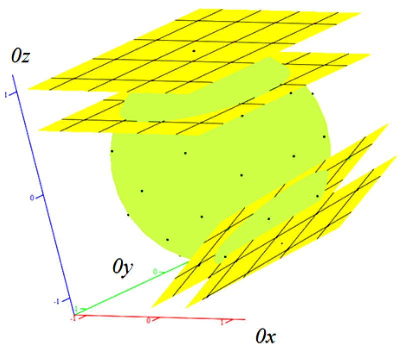


Figure 1. Defining spherical coordinates of the plane

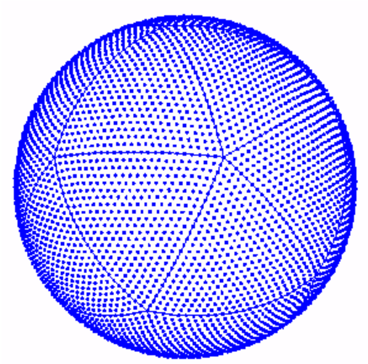
3D images will be scanned by planes. The choice is justified by the following: intersection of a plane with any other extensive geometric primitive is explicit in Higher Mathematics and has strict analytical representation that will be useful for developing the analytic structure of a feature.

Moreover, a grid of parallel planes helps to give invariant description of an object as it provides us with the condition necessary for developing features that are invariant to a group of motions of a 3D image.

One should aim for the situation where all scanning grids of parallel planes at different angles ω and φ of viewing a recognized 3D image will coincide with each other after any its spatial shift or rotation (discrete step disregarded).



a



b

Figure 2. a – a sphere base grid and grids of parallel scanning planes conformed with it; b – an example of a sphere uniform grid

In terms of topology for the continuous case, stereotyped search of all angles ω and φ that identify every scanning grid of parallel planes gives a model of concentric spheres with the center at the origin of coordinates. Every scanning grid of parallel planes on a sphere will be matched to a point that will be a point of tangency with a sphere of the plane that is parallel to planes of the given grid (separately for each pair (ω, φ) of view angles). A set of such points on a sphere will make a grid that will be called a base one (fig. 2a).

It should be noted that a pair of angles (ω, φ) uniquely defines a base grid node that matches up the only tangent plane to the sphere in the point and, therefore, the only scanning grid of parallel planes.

For a discrete case on a standard global map, one can find a thicker concentration of points near the Pole than near the Equator. So if, while rotating, we superpose the Pole and a point at the Equator, then deviation of points of the source grid and the rotated one will be evident. As every point of a sphere base grid uniquely defines a tilting angle of a planes grid, changing the tilting angle of the planes grid will result in changing the form of sections produced. Thus feature calculation errors will be bigger and the accuracy of a 3D object recognition will be lower.

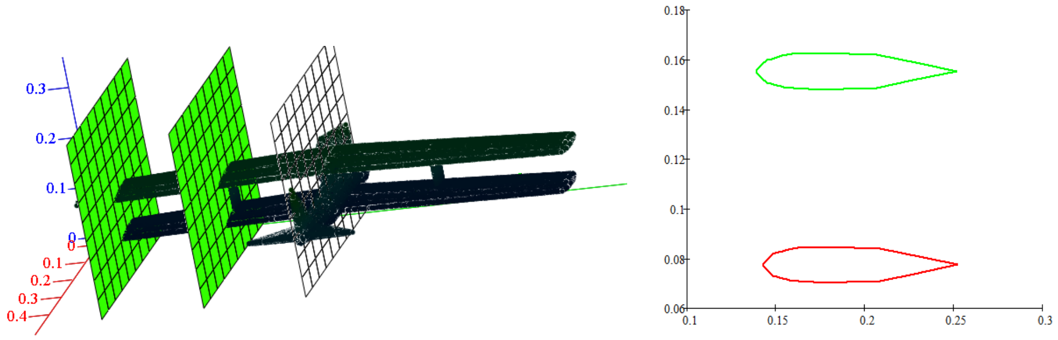
On the other hand, if, while rotating around the center of the sphere, the base grid coincides with itself, then respective grids of tangent parallel planes will completely coincide with each other and the resulting sections will be the same (will not change their form). Hence the calculated feature value will not change.

Thus one need to create a base grid that will ensure a uniform grid of a sphere in order to diminish errors in convergence of the base grid nodes while rotating it because of discrete step of scanning. Uniform distribution of base grid points on a sphere (fig. 2b) will guarantee the absence of spots with high concentration of base grid nodes on the sphere surface, those that define sections at different viewing angles of the object.

The property of uniform distribution of base grid points on a sphere is the necessary condition of invariance of developed features during rotation.

3. Mathematical model of hypertrace transform

Scanning 3D images is done with a grid of parallel planes having distance Δr between the planes and given angles ω and φ (fig. 3a). Mutual location of 3D image F and every scanning plane $B(\eta(\omega,\varphi),r)$ forms a section (fig. 3b) that is defined by number G being calculated according to a certain rule: *HyperT*: $G = HyperT(F \cap B(\eta(\omega,\varphi),r))$. One can use computations of the perimeter or area of the section, the number of intersection of the plane and the source object, the properties of the neighborhood of the section and so on as such rule. In other words, functional *HyperT* defines the property of a section feature.



a b

Figure 3. a – Process of 3D object scanning with planes; b – One of the sections

From here on, scanning with a grid of parallel planes is done for a new value of angle $\omega+\Delta\omega$ and $\varphi+\Delta\varphi$ that has received discrete increments $\Delta\omega$ and $\Delta\varphi$, respectively, with the same value Δr between scanning elements of the planes grid. The same earlier chosen rule *HyperT* is used for sections of a new grid of planes $B(\eta(\omega+\Delta\omega,\varphi+\Delta\varphi),r_i)$. It is important to stress that angles are changed not randomly, but in accordance with the construction of the sphere base grid.

A computational result of the *HyperT* functional depends on three parameters of a plane (r, ω, φ). So if any 2D image received after intersecting the source 3D model with a scanning plane, certain informative feature $\Pi(F_{\sec t})$ will be given according to the rule *HyperT*. Then for numerical analysis it is convenient to present the result of the 3D trace-transform as three-dimension hypertrace matrix *3TM*, which has a vertical 0φ axis, a horizontal 0ω axis and a deep-directed $0r$ axis.

For example, every deep-directed row of the matrix contains feature-elements that are calculated using 2D images [9], which resulted from intersecting a source 3D object with scanning planes after direct search of all values of distance variable r with fixed values of angles ω and φ. If plane B does not intersect the 3D image, that is $F \cap B(\eta(\omega,\varphi),r) = \phi$, then the value of a hypertrace functional is assumed to equal zero: $HyperT(F \cap B(\eta(\omega,\varphi),r)) = 0$.

After filling the three-dimension hypertrace matrix, we process its deep-directed rows with the help of the *HyperP* functional, that can be defined, for example, as follows: $Hyper\mathrm{P} = \int G(\omega, \varphi, r)dr$. As a result, the source three-dimension hypertrace matrix evolves into two-dimension matrix *2TM*. Then we process the *2TM* matrix column after column using functional *HyperΩ*, which can be defined, for example, in the following way: $Hyper\Omega = \max_{\varphi} G(\omega,\varphi)$. As a result, we get horizontal row *1TM* (values vector) that has a continuous analog in the form of the 2π-cyclic curve. One can apply the *HyperΘ* functional to the received set of numbers and obtain a certain number – image feature *Res(F).* The functional can be defined, for example, by the amplitude of the second harmonics of Fourier series.

Hence, the hypertriplet feature of 3D image F has the structure as a combination of four functionals each of which, apart from functional *HyperT,* reduces the matrix *3TM* size by one item if we use them consequently:

$$Res(\mathrm{F}) = Hyper\Theta \circ Hyper\Omega \circ Hyper\mathrm{P} \circ Hyper\mathrm{T}(F_{\mathrm{sec}t}).$$

Every 2D image received after intersecting the source 3D model with a grid of parallel planes at different viewing angles needs scanning to elicit some relevant features (for example, the perimeter of a section figure). To define features of a two-dimension image of a section, we use 2D trace-transform [10].

The triplet feature of 2D image *Fsect* has the structure as a composition of three functionals each of which, apart from functional T, reduces the matrix *TM* size by one item if we use them consequently [10]:

$$\Pi(\mathrm{F_{sect}}) = Hyper\mathrm{T}(\mathrm{F_{sect}}) = \Theta \circ \mathrm{P} \circ \mathrm{T}(\mathrm{F_{sect}} \cap l(\theta,\rho)).$$

Combining the formulas for *Res(F)* and $\Pi(F_{\mathrm{sec}t})$, we finally receive the following analytic structure of a feature of a 3D image in the form of many functionals composition:

$$Res(\mathrm{F}) = Hyper\Theta \circ Hyper\Omega \circ Hyper\mathrm{P} \circ Hyper\mathrm{T}(\Theta \circ \mathrm{P} \circ \mathrm{T}(\mathrm{F_{sect}} \cap l(\theta,\rho))).$$

Due to the composite structure of functionals, presenting the analytic structure of features $\Pi(F_{\mathrm{sec}t})$ and *Res(F)* respectively, it is possible to receive an abundance of features. Moreover, some features have distinct geometric interpretation, which facilitates developing features and increases their discriminative power. Particular qualities of hypertriplet and triplet features allow creating features both sensitive and invariant to group of motions and scaling, which results in better intellectuality and flexibility of the 3D trace-method used for objects recognition.

A more detailed description of a stochastic scanning technique, its merits and influences on a feature development can be found in [10].

4. Conclusion

The method under discussion is rather versatile as the scanning scheme is not stipulated by geometric parameters of a source model. So the proposed technics can deal with objects of any layout or degree of complexity, which increases intellectual capacity of 3D transform.

The described method affords to give an invariant description of a source 3D object under any translation or rotation motions and homothetic transformation. Furthermore, the developed hypertriplet features are strongly resistant to coordinate noise and linear distortions. They also possess distinct geometric interpretation facilitating intellectual capabilities for analysis.

Effective intellectual analysis of hyper-trace transform is stipulated by the ability to control properties of the developed features. Moreover, sensitive features allow receiving a lot of additional information about the object using the same scanning technique (for example, information about the object transformations), which is essential for solving a certain class of practical problems (in robotic technology and the like).

Another intellectual peculiarity of the proposed method is a high level of pre-processing, processing and after-processing of a 3D image using the same scanning technique.

The authors plan to improve the method in such way that it could be applied for analyzing not only binary and monochrome 3D images, but also those having color and texture. Similar results have already been obtained while analyzing 2D images having color and texture in [10]. We expect the intellectual level of hypertrace transforms to be higher after our adding affine-invariant transforms to the theory.

When introducing a small amount of the aluminium oxide powder, the distance between the axes of the second-order dendrites and the average grain size reduces. This implies that a considerable part of the powder particles is effective crystallization centres. When increasing the nanopowder content, the structure starts coarsening relatively that which was obtained using low powder concentrations. The introduction of the large quantity of the modifier leads to its coagulation and reduction of its influence on the structure.

References

[1] Kiy K I 2015 Pattern Recognition and Image Analysis. *Advances in Mathematical Theory and Applications* **25(2)** 338-346
[2] Myasnikov V V 2016 Pattern Recognition and Image Analysis. *Advances in Mathematical Theory and Applications* **26(1)** 22-33
[3] Vasil’ev K K, Dement’ev V E and Andriyanov N A 2016 Pattern Recognition and Image Analysis, *Advances in Mathematical Theory and Applications* **26(1)** 240-247
[4] Song S and Xiao J 2014 *Computer vision* 8694 634-651
[5] Wang C and Huang K-Q 2015 *Journal of Computer Science and Technology* **30(2)** 325-339
[6] Zhang Y, Song S, Tan P and Xiao J 2014 *Computer vision* **8694** 668-686
[7] Andreux M, Rodolà E, Aubry M and Cremers D 2014 *Computer vision* **8928** 299-312
[8] Lmaati E A, Oirrak A E, Kaddioui M N, Ouahman A A and Sadgal M 2010 *International Arab Journal of Information Technology* **7(3)** 264-270
[9] Fedotov N G, Ryndina S V and Semov A A 2014 Pattern Recognition and Image Analysis. *Advances in Mathematical Theory and Applications* **24(4)** 566-574
[10] Fedotov N G 2009 The theory of pattern recognition features based on stochastic geometry and functional analysis, FIZMATLIT **304**