THE HOUGH TRANSFORM AND UNCERTAINTY

V.S.Donchenko

Abstract: The paper deals with the generalisations of the Hough Transform making it the mean for analysing uncertainty. Some results related Hough Transform for Euclidean spaces are represented. These latter use the powerful means of the Generalised Inverse for description the Transform by itself as well as its Accumulator Function.

Keywords: Uncertainty, Hough Transform, Accumulator Function, Generalised Inverse.

Introduction

This report is the attempt to represent Hough Transform (HT) [Hough 1962] as a tool for analysis of the uncertainty

Some results, besides, are represented for the vector observations in the scheme of the Hough Transform as well as for complex observations in this case.

The Hough Transform (HT) for well over forty years has been and continues to be an important tool in analysis of shape and pattern recognition. But potentially the Transform seems to be much more than engineering tool only. The idea of the Transform may be used for analyzing uncertainty in much more general cases than in its classical variant. In the paper [Donchenko 1994] the general concept of the Hough Transform within the Hough-pair of the spaces was proposed. Later, in [Donchenko, Kirichenko, 2001] [Donchenko, Kirichenko 2002] the powerful Generalized Inverse apparatus [Nashed, Votruba 1976]], [Albert 1977], [Kirichenko 1997], [Kirichenko, Lepeha 2001] applied to describe HT and Accumulator Function (AF). In the work the results from [[Donchenko, Kirichenko, 2001] and [Donchenko 1999] are extended on the vector case.

General concept of the Hough Transform – Hough-pare of spaces

General concept of the HT[Donchenko 1994] as a tool for analysis of the uncertainty may be built on the base of so called Hough-pare of spaces, which are virtually a pare of sets S_o, S_p enhanced by its subsets G_θ L_s: $G_\theta \subseteq S_o$, $L_s \subseteq S_p$, $\theta \in S_p$, $s \in S_o$, mutually indexed. One set (space) S_o is interpreted as a space of observations, another S_p – as a space of parameters. This space of parameters is interpreted as a variety of the variants for uncertainty which corresponds to observation s.

These subsets are agreed in the next sense: for any pare (s,θ) $\theta \in L_s \Rightarrow s \in G_\theta$ and conversely: $s \in G_\theta \Rightarrow \theta \in L_s$. Some additional conditions may be added to these: about type of interception for example.

The HT within the Hough-pare is determined as a transition from the observation s – or its sequence s_1, $s_2,...,s_N \in S_o$ – to subset L_S , correspondingly –sequence of subsets $L_{s_1} = L_1$, $L_{s_2} = L_2$,..., $L_{s_N} = L_N$ – of another space, indexed by the observations.

Such determination permits the description of straight and inversed HT, Hough-estimator, Fast HT as a sequential HT and Fast HT as the HT of the complex observations.

Each of the observations is supposed to be taken by the choice of the parameter $\theta_1,..., \theta_N$ and the consecutive choice of the elements $s_1, s_2,...,s_N \in S_o$ from correspondent subsets of G_θ -type. Besides, the observations may be disturbed in that or this way. In this case one says that the elements are observed with an error.

The parameters $\theta_1,..., \theta_N$ –some of them may be equal - is said to be represented in $s_1, s_2,...,s_N \in S_o$.

The set of parameters represented in the sample is supposed to be “comparatively small” and the main target of the HT-based analysis is ascertainment – estimation - of that set. Properly, this is the task of the estimation in the theory of the HT.

HT may be applied to the sequence $s_1, s_2,...,s_N \in S_o$ taken without previous choice of , $\theta_1,..., \theta_N$. In that case HT-analysis is targeted to describe “comparatively small” set of the parameters, “concentrated” the observations. This task may be called the task of the clustering in HT.

As to complex observations S, then we say it to be the subset S = $S_{i_1,\ldots,i_k}$ = $\{s_{i_1}, s_{i_2}, \ldots, s_{i_k}\}$, $\{i_1, i_2, \ldots, i_k\} \subseteq \{1,\ldots,N\}$ of the initial observations. And by HT of the complex observation S= $S_{i_1,\ldots,i_k}$ we will call the set $L_S = L_{i_1,\ldots,i_k}$, determined by the relation:

$$L_S = L_{i_1,\ldots,i_k} = \bigcap_{j=1}^{k} L_{i_j} .$$

Sometimes HT of the complex observations is called Fast HT. Fast HT can cut essentially the set of the parameters, pretended to be represented in the sample.

This variant of the Fast HT one ought to differ from another using, when the Accumulator Function (AF) of the HT consecutively calculated for some set C - rough approximation - and its consecutive – detailed - partitions.

AF is determined by the HT of the sample – original or complex observations – as a function of set C in the space of parameter in one of the next two senses: absolute A(C) or relative NA(C) –by the relations:

$$A(C) = \Sigma_{i=1}^{N} \delta(C \cap L_i), \ C \subseteq S_p , \quad (1)$$

$$NA(C) = A(C)/N = N^{-1} \Sigma_{i=1}^{N} \delta(C \cap L_i), C \subseteq S_p , \quad (2)$$

where $\delta(C)$, $C \subseteq S_p$, equal 1, if C is not empty and 0, if C – empty set.

The summing in (1), (2) for the complex observations is by the set of the complex observations under consideration.

Argument C in the AF depends on the concrete types of spaces. For the Euclidean spaces for parameters and observations set C may be: ball as in (10), (11) below; hyper-cube; compact and so on.

AF is the mean to estimate the set of parameter, which are represented in the sample or the "smallest" set of the parameters in the clustering task. Properly, such set (or sets) is the set of maximum for AF.

Hough Transform in the Euclidean spaces

In its original variant HT was determined for the case, when

- S_o, S_p are appropriate rectangles in R^2;
- parametric sets G_θ, $\theta=(\rho,\varphi) \in R^2$ is the set of the graphics of the straight lines in the normal representation: $G_\theta = G_{(\rho,\varphi)} = \{(x,y) \in R^2: \rho = x \cdot \cos\varphi + y \cdot \sin\varphi\}$;
- parametric set $L_s = L_{(x,y)}$, $s=(x,y) \in R^2$ is the set of parameters for which correspondent lines G_θ include observation s=(x,y): $L_s = L_{(x,y)} = \{(\rho,\varphi) \in R^2: \rho = x \cdot \cos\varphi + y \cdot \sin\varphi\}$.;

Observations s=(x,y) may be with an error so without it. In the first case $y = \bar{y} + \varepsilon_{(x,y)}$, where $\varepsilon_{(x,y)}$- the error of an observation. Errors, which correspond to different observations, are supposed to be independent but not obligatory identically distributed.

One of the generalizations of that original variant may be such one, in which the spaces in the Hough-pare are any Euclidean spaces or their appropriate subsets:

- $S_o = R^m$, $S_p = R^n$
- G_θ, $\theta \in S_p$, is determined the graphic of mappings $y=g(x,\theta)$, $\theta \in R^l$ from R^n in R^m. The sample $s_1, s_2, \ldots, s_N$ consists of the pares $s_i=(x_i, y_i)$:

$$y_i = g(x_i,\theta_i) \in R^m, \ x_i \in R^n , \quad (3)$$

$$y_i = g(x_i,\theta_i) + \varepsilon_i \in R^m, \ x_i \in R^n , \ i=1,\ldots,N \quad (4)$$

Variant (4) represents the scheme of observations with an error, (3) – without it.

HT for such sample is the sequence $L_1, L_2, \ldots, L_N$ of the subsets from R^l, where

$$L_i = \{ \theta \in R^l: y_i = g(x_i,\theta)\}, \ I=1,\ldots,N.$$

Particularly, if the set of mapping is of affine-type (linear + shift) from R^n in R^m, then the matrix $\theta = A \in R^{m\times(n+1)}$ of this map may be considered as the parameter, i.e. $l = m\times(n+1)$.

HT, AF and Hough-estimator are described on the sample (x_i,y_i), i=1,N, $x \in R^n$, $y \in R^m$ points of the graphics of the affine set of the mappings $y = A \begin{pmatrix} x \\ 1 \end{pmatrix}$, - $x \in R^n$, $y \in R^m$, A –m×(n+1)) matrix, $\begin{pmatrix} x \\ 1 \end{pmatrix}$ – block vector-column from $x \in R^n$ and 1.

These observations may be observed in the scheme without the error (3) or with it (4), correspondingly:

$$y_i = A_i \begin{pmatrix} x_i \\ 1 \end{pmatrix}, \quad (5)$$

$y_i = A_i\begin{pmatrix} x_i \\ 1 \end{pmatrix} + \varepsilon_{x_i}$, $x_i \in R^n$, $y_i \in R^m$, $A_i \in R^{m\times(n+1)}$, i =1,…,N. (6)

As it was remarked earlier specific parameter $A_i \in R^{m\times(n+1)}$, i =1,…,N corresponds to each of the observations. Only scheme with the error (6) and independent errors will be considered below. The last means, that errors of the observations ε_{x_i} , i =1,…,N are independent. The distribution of ε_x will be denoted by P_x:

$P_x(B^{(m)}) = P\{\varepsilon_x \in B^{(m)}\}$, (7)

$B^{(m)}$ – Borel set from R^m.

HT $L_{(x,y)}$ of an observation (x,y) is the set of affine transforms, mapping x in the observed y, which may be disturbed:

$L_{(x,y)} = \{A \in R^{m\times(n+1)}: y = A\begin{pmatrix} x \\ 1 \end{pmatrix}\}$. (8)

$S_r(\theta)$ below denotes the r-ball with a center in the θ in the space of all m×(n+1) matrixes with the trace norm, induced by the trace scalar product:

$(A,B) = \mathrm{tr}\, A'B = \Sigma_i (A'B)_{ii} = \Sigma_{ij} a_{ij}b_{ij}$.

The trace norm, obviously, coincides with the Euclidean norm in $R^{m\times(n+1)}$.

AF in absolute or frequency variants will be defined for the balls $S_r(\theta)$ as for arguments and denoted correspondingly $A_r(B)$, $NA_r(B)$:

$A_r(B) = A(S_r(B)) = \Sigma_{i=1}^N \delta(S_r(B)\cap L_i)$, , (9)

$NA_r(B) = A_r(B)/N = N^{-1}\Sigma_{i=1}^N \delta(S_r(B)\cap L_i)$, $B \in R^{m\times(n+1)}$. (10)

Theorem 1. AF for the sample (x_i,y_i), i=1,N points of affine observations may be represented by next expression:

$A_r(B) = \Sigma_{i=1}^N \delta(S_r(B)\cap L_i) = \sum_{i=1}^N \delta(\varepsilon_{x_i} \in S_{r\sqrt{1+\|x_i\|^2}}((B - A_i)\begin{pmatrix} x_i \\ 1 \end{pmatrix})$. (11)

Proof. Accordingly with the theorem 2 [2]

$A_r(B) = \Sigma_{i=1}^N \delta(S_r(B)\cap L_i) = \Sigma_{i=1}^N \delta(\|y_i - B\begin{pmatrix} x_i \\ 1 \end{pmatrix}\|^2 \le r^2(1 + \|x_i\|^2)$, $B \in R^{m.\times(n+1)}$. (12)

As for the each of observations $y_i = A_i\begin{pmatrix} x_i \\ 1 \end{pmatrix} + \varepsilon_{x_i}$, i =1,…,N,

then condition $\|y_i - B\begin{pmatrix} x_i \\ 1 \end{pmatrix}\|^2 \le r^2(1 + \|x_i\|^2)$

in (12) is equivalent to the condition $\|A_i\begin{pmatrix} x_i \\ 1 \end{pmatrix} + \varepsilon_{x_i} - B\begin{pmatrix} x_i \\ 1 \end{pmatrix}\| \le r\sqrt{1+\|x_i\|^2}$,

that proves the theorem.

Remark 1. $S_r(B)$ in (12) is the r-ball in the trace norm in the matrix space with the center in a B, then the $S_{r\sqrt{1+\|x_i\|^2}}((B-A_i)\begin{pmatrix} x_i \\ 1 \end{pmatrix})$ is the $r\sqrt{1+\|x_i\|^2}$ -ball in R^m with the center in $(B-A_i)\begin{pmatrix} x_i \\ 1 \end{pmatrix}$, i=1,…,N.

Corollary 1. Obviously, $\delta(S_r(B)\cap L_i)$, i =1,…,N are Bernoulli-distributed random variables with the parameters, determined by the expressions

$p_i = P\{\varepsilon_{x_i} \in S_{r\sqrt{1+\|x\|^2}}((B - A_i)\begin{pmatrix} x_i \\ 1 \end{pmatrix}\}$, i=1,…,N. (13)

Proof. The result is the consequence of taking 1 for each of $\delta(S_r(B)\cap L_i)$, i =1,…,N in (12).

$\delta(\varepsilon_{x_i} \in S_{r\sqrt{1+\|x\|^2}}((B - A_0)\begin{pmatrix} x_i \\ 1 \end{pmatrix})$, i=1,…,N.

Theorem 2.(0-1 Law). The limit value of the AF when with probability 1 is finite or infinite as $n \to \infty$. It is finite iff $\sum_{n=1}^{\infty} p_i = \sum_{n=1}^{\infty} P\{\varepsilon_{x_i} \in S_{r\sqrt{1+\|x\|^2}}((B - A_i)\begin{pmatrix} x_i \\ 1 \end{pmatrix}\} < \infty$

Proof. The proof repeats that one for scalar case in [8].

Theorem 3. The next limit take place with the probability 1:

$\lim_{N\to\infty}(N^{-1}\sum_{i=1}^{N}\delta(S_r(B)\cap L_i)) - N^{-1}\sum_{i=1}^{N}p_i) = \lim_{N\to\infty}(NA_r(B)) - N^{-1}\sum_{i=1}^{N}p_i) = 0$, where p_i, i=1,…,N are determined by (13).

Proof. As in was in previous case the proof repeats that one for scalar case in [8].

Remark 2. For the case under consideration – vector case – all the consequences from [8] for scalar observations are valid.

Remark 3. The statements of the theorems earlier are free from constraints on the distribution of an error. Besides, the distribution may depends on x.

Theorem 4. AF for the sequence of K complex observations may be represented by next expression:

$$A_r(B) = \Sigma_{i=1}^{K}\delta(S_r(B)\cap Ł_i) = \Sigma_{i=1}^{K}\delta(\ ||(Y_i - BX_i)X_i^{+}|| \le r)\ ,\ B \in R^{m\times(n+1)},$$

where:

- $Ł_i$, i=1,…,K – Hough transforms for the complex observations,
- A^+- General Inverse for A,
- Y_i, i=1,...,K – block-matrixes from the y-components of the original observations the complex observation consists of,
- X_i, i=1,…,K– block-matrixes from the x-components of the original observations the complex observation consists of,

Conclusion

The subject matter of the paper is generalizing the Hough Transform that convert it in mathematical tool with wide range of application in analyzing the “uncertainty”. The abstract form of the HT is represented within the framework of Hough-pare of spaces.

The HT for observations and parameters from Euclidean spaces has been represented and investigated for affine sets of transforms. The author would like to believe that the results represented are the only one step to promote HT to be the mean for uncertainty analyzing.

Bibliography

[Albert 1977]Albert A. Regression, pseudo inverse, recurrent estimation..–M.: Nauka, 1977 г., 305 p(in Russian).

[Donchenko 1994] Donchenko V.S. General Scheme of the Hough Transform and properties of the Hough estimates in special case of Discrete Spaces.//Statist. Res. Rep.–Umea, Sweden: University of Umea,1994.–1994-6, S-901 87. 10 p.

[Donchenko 1999] Donchenko V.S. Limit theorems for Accumulator Function in Hough Transform Scheme. // Bulletin of the University of Kiev. Series: Physics \& Mathematics. –1999, Vol..№ 1, cc.191-195.(In Ukrainian)

Donchenko, Kirichenko,2001] Donchenko V.S., Kirichenko M.F. Hough Transform and Generalized Inverse. // Bulletin of the University of Kiev, Series: Physics \& Mathematics.– 2001, vol. № 4, , pp.191-196. (in Ukrainian)

[Donchenko , Kirichenko 2002] Donchenko V.S., Kirichenko M.F. Fast Hough Transform and Generalized Inverse.// Journal of Automation and Information Sciences . – 2002, vol. №2, pp..115-125. (In Russian, also translated in English)

[Hough 1962] Hough P.V.C. Method and Means for Recognizing Complex Patterns. - U.S. Patent 3069354, 1962.

[Kirichenko 1997] Kirichenko M.F. Analytical representation for Generalized Inverse disturbances // Cybernetics and System Analysis. – 1997, №2, pp.98-107. (In Russian)

[Kirichenko, Lepeha 2001]. Kirichenko M.F., Lepeha M.P. Disturbances for pseudo-inverse and projective matrixes and their applications to linear and non-linear identification.// Journal of Automation and Information Sciences. - 2001, №1, c.6-23. (In Russian, also translated in English)

[Nashed, Votruba 1976] Nashed M. Zuhair ,Votruba G.F. A Unified Operator Theory of Generalized Inverse, Proceedings of an Advanced Seminar Sponsored by the Mathematical Research Center, The University of Wisconsin, Madison, October 8-10, 1973. – New York, Academic Press, 1976.

Author information

Volodymyr S. Donchenko – Kyiv National Taras Shevchenko Universiry, Professor.
2, Akademician Glushkov prospectus, building 6, Kyiv,03680, Ukraine; e-mail: vsdon@unicyb.kiev.ua