

The reliability of pattern-match searching for the fragment on image using set of pseudo-gradient procedures

L.Sh. Biktimirov¹, A.G. Tashlinskii¹

¹Ulyanovsk State Technical University, ul. Severnyi Venets 32, 432027, Ulyanovsk, Russia

Abstract

The effectiveness of pattern-match searching for the fragment on image using set of pseudo-gradient procedures covering all initial image by their workplaces is studied. Procedures control is managed to reduce computational expenses and based on analysis of penalty function and giving priority of making next iteration to procedure that has minimum value of penalty. It is considered that required fragment belongs to the domain that has a procedure which is reached prescribed limit of iterations. If it is prior unknown if there is any required fragments on the initial image, the hypothesis of their absence should be tested. If the hypothesis is not confirmed than the scan for domains with fragments should be performed. Concerning there are fragments on the image the missing probabilities are found using penalty function and limit of iterations. Herein, the proposal is considered for single and multiple required fragments.

Keywords: digital image; fragment searching; fragment missing; pseudo-gradient procedure; probability; first and second type errors

1. Introduction

Pattern search of the single or multiple equal fragments belongs to the field of digital image processing [1-4]. There is type of search algorithms based on pseudo-gradient procedures (PGP) [5,6]. But PGP has relatively small working range [7], so it is necessary to split the high-resolution images to a multiple domains, with its own procedures. At this time appears a task to choose domains containing required fragments.

There is a missing fragment probability assigned to a search process. If all procedures are in equal conditions and make equal number of iterations than missing fragment probability depends on the preselected criteria of choice. It can be the best value of goal function of quality estimating on the last iteration. But this method provides low choice veracity cause goal function estimations calculates based on small-size local sample. Besides, number of search domains can be up to tens of thousands [8,9], so, to make all procedures achieve the prescribed limit of iterations, needs huge computational cost. To increase the probability it should be used more reliable criteria of choice, the same as using estimations of PGP on last iteration, for example, maximum of correlation index between required fragment template and its probable location on the image [10]. But this causes even more computational costs.

To reduce computational costs in [11] the algorithm for manipulating of set of PGP's is proposed where on the each step the priority to make next iteration to procedure that has best value of some penalty function (PF) X [12]. In this case step of the algorithm means complex of the operations: making ordinary iteration by procedure with best PF value, calculating new PF value and finding procedure with best PF value. The domain with procedure that firstly made prescribed number of iterations T is being considered as a required one (domain that probably is containing required fragment). To search $k > 1$ fragments it should be chosen k domains containing procedures that achieved limit of iterations faster than others. Further, probability of wrong domain choice will be considered as probability of error caused by mentioned method of manipulating ensemble of procedures.

Generally, it is unknown, if there are any required fragments in domains under investigation. Thus, search procedure must contain some kind of testing the hypothesis of absence of fragments. During testing process there can be first $P^{(1)}$ and second $P^{(2)}$ type errors. So, if prior probability of location of the fragment among considering domains, is equal to P_F than decision about presence of the fragment is accepted with probability

$$P_F(1 - P^{(2)}) + (1 - P_F)P^{(1)}$$

and about absence – with probability

$$P_F P^{(2)} + (1 - P_F)(1 - P^{(1)})$$

Work [13] considers issue of testing hypothesis of fragment's absence. If the hypothesis rejected than next step is to determine the position of the fragment. Here, the probability of choosing domain with fragment after declining hypothesis with prescribed second type error probability:

$$P = P_F P_{ER} + (1 - P_F)P^{(2)}, \quad (1)$$

where P_{ER} is relative probability of wrong fragment choice in case it really is on the image.

Let us consider the probability P_{ER} of wrong choice of image domain with fragment in case it is on the image. Also the probability of making an error selecting $k > 1$ domains when there are k identical fragments on the image, for example, images of equal objects (biological or technical) will be considered. Taking into account differences in search processes for single or multiple fragments these processes would be investigated separately.

2. Error probability in case of searching for a single fragment

Let image (or just part of it) where should be found the fragment is divided to N domains and there is just one of them to contain required fragment. Let's find error probability of that domain identification P_{ER} . Assume that goal function and PF of PGP are pre-defined. Let us call « x^+ » value of PF X for the procedure in domain with fragment and « x^- » in domain without fragment.

If there are only two domains then domain without fragment will be chosen if its procedure will be first to make T iterations, i.e. $x_T^- < x^+$, where x_T^- is equal to PF value on T th iteration. Here, second procedure may perform from 1 to $(T-1)$ iterations. Then, if value of PF on T th iteration is equal to x_0 , to assume that choice is wrong it should be two conditions simultaneously: PF of the procedure in domain with required fragment exceed x_0 and in domain without fragment PF value should be equal to x_0 . Proposing that these events are independent the probability of wrong choice will equal to:

$$P_{ER} = \int_{x_0}^{\infty} w(x_T^+) dx \int_0^{x_0} w(x_T^-) dx.$$

But value of x_0 is prior unknown and wrong choice probability generally is equal to probability that on T th iteration $x_T^+ > x_T^-$

$$P_{ER} = \int_0^{\infty} w(x_T^-) (1 - F(x_T^+)) dx = \int_0^{\infty} w(x_T^+) F(x_T^-) dx, \quad (2)$$

where $F(x_T) = \int_0^x w(x_T) dx$ is integrated distribution function.

It should be noticed that densities of distribution $w(x^+)$ and $w(x^-)$ are depending on initial approximation of search parameters [14] estimated by procedure and in this context are relative. Proposing that initial parameters' approximation for procedure in fragment's domain gets worst convergence in work range of procedure P_{ER} will be the upper limit of wrong fragment choice's probability.

If number of separated domains is equal to N , than assuming independence of procedure's PF (taking into account (2)):

$$P_{ER} = \int_0^{\infty} w(x_T^+) \left(1 - (1 - F(x_T^-))^{N-1} \right) dx. \quad (3)$$

The assumed restriction about PF independence is not strict cause samples from domains that don't have a fragment have weak correlation with samples from fragment.

3. Error probability in case of searching for multiple fragments

In previous case the required domain was assumed domain with procedure achieved limit of T iteration first of all others. Here, to provide low error probability of fragment's search with high signal/noise ratio it is necessary to specify large number of iterations. It is possible to decrease error probability with low T choosing several domains where procedures made limit of the iterations. Then probability of occurrence of domain with fragment among chosen domains increases. But it is not true for probability of right choice of the fragment.

There are criteria allowing to identify the fragment with low error probability, such as above mentioned maximum of correlation index that can be calculated on whole image, or extremes of information-theoretical measures of images similarity [15]. But using such criteria causes large computational expenses. Moreover, if image is divided into a lot of search domains and computing resources are strictly limited using of these criteria is not reasonable. But for small number of domains (for example, two) using these criteria is acceptable. On this basis, the probability of location the fragment among n domains where corresponding procedures firstly made prescribed limit of iterations.

If local samples to estimate all procedure's goal function value are independent and suppose best value of PF has domain without fragment a random event, than task may be reduced to Bernoulli scheme. So for probability $P_{ER}^{(n)}$ of missing domain with fragment during choice n domains with procedures first reached prescribed limit of iterations using binomial law it can be written:

$$P_{ER}^{(n)} = \int_0^{\infty} w(x_T^+) \sum_{i=n}^{N-1} C_i^{N-1} F(x_T^-) (1 - F(x_T^-))^{N-i-1} dx, \quad (4)$$

Where C_i^{N-1} is a number of combinations from $(N-1)$ of i elements. It should be noticed that cause number of investigating domains in general is less than general number of domains of separated image ($n \ll N$) so it is reasonable to use next expression:

$$P_{ER}^{(n)} = \int_0^{\infty} w(x_T^+) \left(1 - \sum_{i=0}^{n-1} C_i^{N-1} F(x_T^-) (1 - F(x_T^-))^{N-i-1} \right) dx. \quad (5)$$

Fig. 1 shows an example of graphs of dependency for probability of missing required fragment $P_{ER}^{(n)}$ on the number of iteration when choosing one (solid-line curve) and two (dashed-line curve) domains. Calculation carried out for relay-type PGP with working range requiring to split two different-size images on 36 (curve 1 and 2) and 625 domains (curve 3 and 4). Initial parameters of mismatch were equal to 6 steps of parallel shift and 20 degrees of turn.

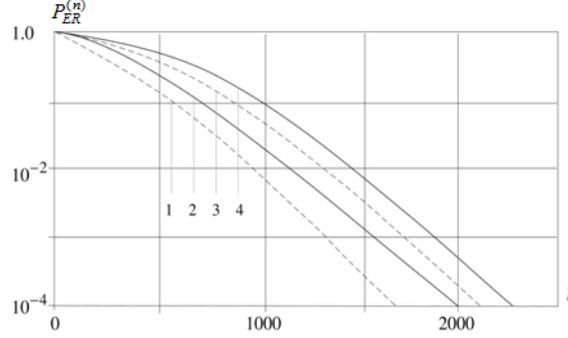


Fig. 1. Probability of missing required fragment against the number of iteration when choosing one and two domains.

During calculation was used expression (5) where meaning of expression:

$$\sum_{i=0}^{n-1} C_i^{N-1} F(x_T^-) (1 - F(x_T^-))^{N-i-1}$$

was computed as a ratio of incomplete $B_q(N-n, n) = \int_0^q x^{N-n-1} (1-x)^{n-1} dx$ and complete $B(N-n, n) = \int_0^1 x^{N-n-1} (1-x)^{n-1} dx$ beta-functions [16], where $q = 1 - F_{PT}(\psi)$. Hence while representing complete beta-function through gamma-function [17]

$$B(N-n, n) = \frac{\Gamma(N-n)\Gamma(n)}{\Gamma(N)},$$

will get

$$P_{N-1}\{0 \leq i \leq n-1\} = \frac{\Gamma(N) \int_0^q x^{N-n-1} (1-x)^{n-1} dx}{\Gamma(N-n)\Gamma(n)}$$

It is obvious from graphs that if $n=2$ probability $P_{ER}^{(n)}$ is essentially low. For example, if $N=36$ and $T=1000$ probability of error in choice depending to situation $n=1$ is decreasing by 5.5 times, and if $T=2000$ – by 9 times. In this case, of course, computational expenses are increasing too.

4. Error probability in case of searching for several equal fragments

Let's consider probability of missing at least one of domains with fragments location during search position of $k > 1$ similar fragments. In this case it should be at a minimum k procedures to make limit number of T iterations. Same as before, it will be considered that presence of the fragments is known in advance.

In particular, in case $n=k$ similar to (3) probability of all k procedures first made T iterations will correspond to domains containing fragments:

$$P^{(k)} = 1 - P_{ER}^{(k)} = 1 - \int_0^{\infty} w_k(x_T^+) \left(1 - (1 - F(x_T^-))^{N-1} \right) dx, \quad (6)$$

where $w_k(x_T^-)$ means probability density function for maximum of k PF's values of procedures from domains with fragments.

To decrease probability of missing fragments the number of domains to choose can be more than the number of required fragments ($n > k$). In this case probability $P_{ER}^{(j,k)}$ of missing j domains with fragments from k considering ratio (4) and uniqueness condition for each domain is equal to:

$$P_{ER}^{(j,k)} = \int_0^{\infty} w_k(x_T^+) \left(1 - \sum_{i=0}^n C_i^{N-k} F(x_T^-) (1 - F(x_T^-))^{N-k-i} \right) dx, \quad (7)$$

where $w_k(x_T^\pm) = C_j^K w(x_T^-) \frac{(1 - F(x_T^\pm))^{j-1} F^{k-j-1}(x_T^\pm)}{jk - j^2} (j - kF(x_T^\pm))$ is probability density function of j th ranged by maximum PF k procedures in domains with fragments $i = \overline{1, k}$.

5. Conclusion

One of the class of the algorithms to search fragment on the image by template is based on the PGP. But these procedures have relatively small working range of search that makes it necessary to split the image into array of domains each of them is containing own procedure. Here is a task about finding domains with required fragments. If all search procedures are working in same conditions and will make equal number of iterations than it require huge computational expenses. To reduce these expenses the algorithm of managing ensemble of PGP [12] can be used. In this case on the each step priority of making next iteration is giving to procedure that has best value of some penalty function. Domains with procedures made prescribed limit of iterations first are chosen as a required ones (probably containing fragments).

If it is prior unknown if there are required fragments on the image it is necessary to test the hypothesis about absence of fragments with prescribed error probabilities of first and second types. This issue and statistical criteria of hypothesis validity are considered in papers [13, 18]. If the hypothesis is rejected then choosing domains of fragments' location carried out. In this case error probability of choosing domains of fragment's location is a conditional probability and with prescribed second-type error probability in general determines by expression (1).

Probability of wrong choice assuming there are required fragments on the initial image depends on their count. If there is only one fragment and procedure then probability of wrong choice of domain with fragment determines by expression (3). To reduce error of right domain it can be chosen several domains instead one (Supposing using additional criteria to choose final domain among selected). Here, probability of presence of required domain among selected determines by expression (5).

If required fragments is more than one and each of them has corresponding one search procedure then probability of case when all procedures of domains with objects will make limited number of iterations first determines by expression (6). If number of choosing domains is greater than number of required fragments then probability of missing prescribed count of domains with fragments determines by expression (7).

Acknowledgements

The study was carried out with financial support of the RFBR grant 16-01-00276.

References

- [1] Ipatov YuA, Krevetsky AV. Modeling methods of detection and spatial localization of group point objects. Science. Technology. Production 2014; 2: 7–11.
- [2] Gerasimova NI, Verkhoturova AE. Search of the image fragment with application of Kohonen neural network. Information technologies in science, management, social sphere and medicine. Tomsk: TPU 2014; 1: 68–70.
- [3] Nikolenko AA, Babilunga OYu, Zaykovskij VN. Localization of specific image fragments based on two-dimensional wavelet filters. Herald of the National Technical University "KhPI". Subject issue: Information Science and Modelling. Kharkov: NTU KhPI 2011; 36: 122–127.
- [4] Chambon S, Crouzil A. Dense matching using correlation: new measures that are robust near occlusions. British Machine Vision Conference, Norwich, Great Britain 2003; 1: 143–152. DOI: 10.5244/C.17.15.
- [5] Tsyarkin JZ. Information theory of identification. Moscow: Nauka; Fizmatlit, 1995; 336 p.
- [6] Zitova B, Flusser J. Image registration methods: a survey. Image and vision computing 2003; 21(11): 977–1000. DOI: 10.1016/S0262-8856(03)00137-9.
- [7] Tashlinskii AG. Estimation of parameters of spatial deformations of image sequences. Ulyanovsk, UISTU, 2000; 131 p.
- [8] Szeliski R. Image alignment and stitching: A tutorial. Foundations and Trends in Computer Graphics and Vision 2006; 2(1): 1–104. DOI: 10.1561/0600000009.
- [9] Tashlinskii AG. Pseudo-gradient Estimation of Digital Images Interframe Geometrical De-formations. Vision Systems: Segmentation & Pattern Recognition. Vienna, Austria: I Tech Education and Publishing 2007: 465–494. DOI: 10.5772/4975.
- [10] Pankova TL, Reznik AL. The effectiveness of algorithms for precision alignment of digital images . Optoelectronics, Instrumentation and Data Processing (Avtometriya) 1991; 5: 39–43.
- [11] Tashlinskii AG, Muratkhanov DS. Structural optimization of algorithms of parameter estimation of geometric image deforming. The physics and technical applications of wave processes 2001: 102–110.
- [12] Tashlinskii AG, Muratkhanov DS. Structural Optimization of pseudo-gradient Algorithms for Measuring Interframe Image Deformations. Pattern Recognition and Image Analysis 2003; 13(1): 177–178. DOI: 10.1134/S1054661806020088.
- [13] Biktimirov LSh, Tashlinskii AG. Estimating the probability of absence of target fragment on image for algorithm with control of multiple search procedures. Radioengineering 2016; 9: 6–10.
- [14] Tashlinskii AG, Kaveev IN, Voronov SV. Image registration method in conditions of intensive noise. Radioengineering 2012; 9: 45–49.
- [15] Voronov SV, Tashlinskii AG. Efficiency analysis of information theoretic measures in image registration. Pattern recognition and image analysis 2016; 26(3): 502–505. DOI: 10.1134/S1054661816030226.
- [16] Levin BR. Theoretical bases of statistical radio engineering. M.: Radio and communication, 1989; 656 p.
- [17] Koroljuk VS. A Handbook on Probability Theory and Mathematical Statistics. M.: Science, 1985; 640 p.
- [18] Biktimirov LSh, Tashlinskii AG. Criteria of testing the hypothesis of absence of target fragment on image. Modern problems of design, production and operation of radio engineering systems. Ulyanovsk: UISTU, 2016; 137–140.