Stochastic Measure of Informativity and its Application to the Task of Stable Extraction of Features

Alexander Lepskiy[1]

Abstract In the paper we define a new notion of stochastic monotone measure. The application of this notion to solution of problem of finding of features on the noisy image is considered.

1 Introduction

As a rule in pattern recognition or, in particular, in image processing, we should identify images using sets of their features. Let $\Omega = \{\omega_i\}_{i=1}^n$ be a set of features that correspond to an image. To achieve the highest productivity and stable working of a pattern recognition system, it is necessary to choose a small subset of features in Ω with the highest information values. There are some very well-known approaches that can give us features with the highest information values based on the method of principal components, discriminant analysis and so on [2], [3]. But these methods fail to take into account structural (e.g. morphological) characteristics of object. In this situation, measure of informativity can be used [1]. By definition, an measure of informativity μ is a set function defined on the power set 2^Ω of Ω that for each $A \in 2^\Omega$ shows an information value of features in A. We assume that this function has monotone property: $\mu(A) \leq \mu(B)$ if $A \subseteq B$ for all $A, B \in 2^\Omega$ (i.e. additional information does not decrease the value of μ).

In certain tasks of image processing random nature of image features can be caused by some noisy effects. For example, if the pattern is a discrete plane curve that extracted on the image and features are some characteristics of curve points (e.g. feature is a estimation of curvature in given point of discrete curve [4]) then a random character of features (e.g. curvature) will be due to noise of image. In this case the expectation $\mathbf{E}[\mathrm{M}(A)]$ be characterize the level of informativeness of representation A and the variance $\sigma^2[\mathrm{M}(A)]$

Higher School of Economics, Moscow, Russia alex.lepskiy@gmail.com

be characterize the level of stability of representation to noise of pattern. Then there is the problem of finding the most stable and informative representation A of the pattern X. The complexity of solutions of this problem will be determined by the degree of dependence of random features of each other. Stochastic measure of informativity M will be additive measure if the features are independent random variables. This case was considered in [5]. In this work we will consider the case when any random feature depends from some other features. Then we get nonadditive monotone measure of informativity. More detail the problem of finding the most stable and informative representation is investigated in this paper for the most popular measure of informativity for contour image – measure of informativity by length.

2 Monotone Geometrical Measure of Informativity

Measures of informativity can be effectively used in image processing as shown in [1]. In image processing the contours of the patterns and their characteristics, for example, curvatures of smooth curves are the such features that should not depend on illumination of a scene and orthogonal transformations (such as rotation, bias, scaling). However, in reality, we have digitized curves that are given by some ordered sets of points. These curves can be corrupted by noise. This means that we can use only some statistical estimates of curvature [4] that not stable to noise. A problem of choosing an optimal polygonal representation of a contour consist in finding such a representation that preserves geometrical characteristics of contour and also that will be stable to noise. This choice can be produced by using geometrical measure of informativity that are axiomatically defined as follows.

Let X be an initial closed contour given by an ordered finite set points, i.e. $X = \{x_1, ..., x_n\}$, where $x_i \in \mathbb{R}^2$, $i = 1, ..., n$. We identify with any nonempty subset $B = \{x_{i_1}, ..., x_{i_m}\}$ a contour generated by connecting points with straight lines starting from points x_{i_1}, x_{i_2} and ending by points x_{i_m}, x_{i_1}.

Definition 1. A geometrical measure of informativity $\mu : 2^X \to [0, 1]$ is a set function that has to obey the following properties: 1) $\mu(\emptyset) = 0$, $\mu(X) = 1$; 2) $A, B \in 2^X$ and $A \subseteq B$ implies $\mu(A) \le \mu(B)$; 3) let $B = \left\{..., x_{i_{k-1}}, x_{i_k}, x_{i_{k+1}}, ...\right\} \subseteq X$ and neighbouring points $x_{i_{k-1}}, x_{i_k}, x_{i_{k+1}}$ belong to a straight line in the plane, then $\mu(B) = \mu\left(B \backslash \{x_{i_n}\}\right)$; 4) μ is invariant w.r.t. affine transformations.

Emphasize that axioms 1, 2 have been introduced by Sugeno for fuzzy measures (see [7]). Consider several ways for defining geometrical measure of informativity [1].

a) Suppose that the length of an original contour is not equal to zero and a function $L(A)$ gives us the length of subcontour $A \in 2^X$. Then a measure of informativity defined by contour length is $\mu_L(A) = \frac{L(A)}{L(X)}$.

b) Suppose that the domain limited by an original contour is convex, and a function $S(B)$ determines the area bounded by an subcontour $A \in 2^X$. Then a measure of informativity defined by contour area is $\mu_S(A) = \frac{S(A)}{S(X)}$.

c) Let $w(x, A)$ be a positive estimate of information value of the part of a contour in a neighbourhood of point $x \in A$ in a subcontour $A \in 2^X$. Then an average measure of informativity is defined by $\mu(A) = \dfrac{\sum_{x \in A} w(x, A)}{\sum_{x \in X} w(x, X)}$, where $w(x, A)$ has to be defined for any non-empty contour $A \in 2^X$ and $\mu(\emptyset) = 0$ by definition. It is easy to see that the introduced geometrical measure of informativity μ_L and μ_S can be considered as average measure of informativity. For example, for μ_L function $w(x, A) = |x - y|$, where y is a next neighbouring points in contour A; in case of μ_S function $w(x, A) = S(O, x, y)$, where O is the centroid of area, bounded by contour A, and $S(O, x, y)$ is the area of triangle with vertices in points O, x, y.

3 Stochastic Average Measure of Informativity

In real situations, values $w(x, A)$ can be considered as random values, because an original contour is corrupted by an additive probabilistic noise. To stress this, we denote these values by capital letters as $W(x, A)$. In this case we have the measure of informativity $\mathrm{M}(A) = \frac{\sum_{x \in A} W(x,A)}{\sum_{x \in X} W(x,A)}$. Then there is problem of finding the most stable and informative representation $A \in 2^X$ of the pattern X for which the expectation $\mathbf{E}[\mathrm{M}(A)]$ will be maximize and the the variance $\sigma^2[\mathrm{M}(A)]$ will be minimize. If $W(x, A) = W(x)$ and random values $W(x)$, $x \in X$, are independent random variables then the measure of informativity $\mathrm{M}(A)$ is additive and the problem of finding the most stable and informative representation was investigated in this case in [5]. We emphasize that stochastic additive measures have been already investigated in the literature (see e.g. [6]). Now we will be investigated the important case when the value $W(x, A)$ depends on two neighbouring points. For example, the geometrical information measure μ_L and μ_S are satisfied this condition.

Let $X = \{x_1, ..., x_n\}$ be an original contour and let vertices be ordered by their indices. So if we consider any subcontour $A \in 2^X$, then the order defined on A is assumed to be generated by the order on X and given by indices in the representation $A = \{x_{i_1}, ..., x_{i_m}\}$, where $i_1 < ... < i_m$. So for any $A = \{x_{i_1}, ..., x_{i_m}\} \in 2^X$ we can identify its elements by their indices and write $x_k(A) = x_{i_k}$ if $k \in \{1, ..., m\}$. We can also consider any integer index k assuming that $x_k(A) = x_l(A)$ if $l \equiv k (\mathrm{mod} m)$. To work with such indices, we use a mapping π defined by $x_k(A) = x_{\pi_A(k)}$. We suppose that $W(x_k(A), A) = W(x_k(A), x_{k+1}(A))$, $k = 1, ..., |A|$, i.e. the value $W(x_k(A), A)$ depends on two neighbouring points $x_k(A), x_{k+1}(A)$. Further, for simplicity reasons, we denote $W(x_k(A), x_{k+1}(A)) = W_{k,k+1}(A)$. Then an average monotone measure and stochastic average monotone measure have a view

$$\mu(A)=\frac{\sum_{k=1}^{|A|}w_{k,k+1}(A)}{\sum_{j=1}^{|X|}w_{k,k+1}(X)},\mathrm{M}(A)=\frac{\sum_{k=1}^{|A|}W_{k,k+1}(A)}{\sum_{j=1}^{|X|}W_{k,k+1}(X)} \tag{1}$$

correspondingly. We call M a stochastic monotone information measure if $W_{k,k+1}(A)$, $A\in 2^X$ are random variables. In this case M has random values. In this section we find estimates of numerical characteristics of M assuming that random variables $W_{k,k+1}(A)$, $W_{l,l+1}(A)$ are independent if $|l-k|>1$. This situation appears if we suppose that x_k, $k=1,...,n$, are also independent random variables.

We see that $\mathrm{M}(A)=\frac{\xi}{\eta}$, where $\xi=\sum_{k=1}^{|A|}W_{k,k+1}(A)$ and $\eta=\sum_{j=1}^{|X|}W_{k,k+1}(X)$. The following lemma is used for estimating $\mathbf{E}\left[\mathrm{M}(A)\right]$ and $\sigma^2\left[\mathrm{M}(A)\right]$.

Lemma 1. *Let ξ and η be random variables that taking values in the intervals l_ξ, l_η respectively on positive semiaxis and $l_\eta\subseteq((1-\delta)\mathbf{E}\left[\eta\right],(1+\delta)\mathbf{E}\left[\eta\right])$, $l_\xi\subseteq(\mathbf{E}\left[\xi\right]-\delta\mathbf{E}\left[\eta\right],\mathbf{E}\left[\xi\right]+\delta\mathbf{E}\left[\eta\right])$. Then it is valid the following formulas for mean and variance of distribution of $\frac{\xi}{\eta}$ respectively*

$$\mathbf{E}\left[\frac{\xi}{\eta}\right]=\frac{\mathbf{E}[\xi]}{\mathbf{E}[\eta]}+\frac{\mathbf{E}[\xi]}{\mathbf{E}^3[\eta]}\sigma^2\left[\eta\right]+\frac{1}{\mathbf{E}^2[\eta]}\mathbf{Cov}\left[\xi,\eta\right]+r_1, \tag{2}$$

$$\sigma^2\left[\frac{\xi}{\eta}\right]=\frac{1}{\mathbf{E}^2[\eta]}\sigma^2\left[\xi\right]+\frac{\mathbf{E}^2[\xi]}{\mathbf{E}^4[\eta]}\sigma^2\left[\eta\right]-\frac{2\mathbf{E}[\xi]}{\mathbf{E}^3[\eta]}\mathbf{Cov}\left[\xi,\eta\right]+r_2, \tag{3}$$

where $\mathbf{Cov}\left[\xi,\eta\right]$ is a covariation of random variables ξ and η, i.e. $\mathbf{Cov}\left[\xi,\eta\right]=\mathbf{E}\left[\left(\xi-\mathbf{E}\left[\xi\right]\right)\left(\eta-\mathbf{E}\left[\eta\right]\right)\right]$; r_1, r_2 are the residuals those depends on numerical characteristics of ξ and η. It being known that $|r_1|\le\frac{\delta}{1-\delta}\cdot\frac{\mathbf{E}[\xi]+\mathbf{E}[\eta]}{\mathbf{E}^3[\eta]}\sigma^2\left[\eta\right]\le\frac{\mathbf{E}[\xi]+\mathbf{E}[\eta]}{(1-\delta)\mathbf{E}[\eta]}\delta^3$, $|r_2|\le C\delta^3$.

Proof. We prove formula (2). The formula (3) is proved by analogy. Expand the function $\phi(x,y)=\frac{x}{y}$ into a Taylor series at the point $(\mathbf{E}\left[\xi\right],\mathbf{E}\left[\eta\right])$. We get

$$\phi(x,y)=\phi\left(\mathbf{E}\left[\xi\right],\mathbf{E}\left[\eta\right]\right)+\sum_{n=1}^{\infty}\frac{1}{n!}d^n\phi\left(\mathbf{E}\left[\xi\right],\mathbf{E}\left[\eta\right]\right)=$$

$$=\phi\left(\mathbf{E}\left[\xi\right],\mathbf{E}\left[\eta\right]\right)-\frac{\mathbf{E}[\xi](y-\mathbf{E}[\eta])-\mathbf{E}[\eta](x-\mathbf{E}[\xi])}{\mathbf{E}^2[\eta]}\sum_{n=0}^{\infty}\left(\frac{\mathbf{E}[\eta]-y}{\mathbf{E}[\eta]}\right)^n.$$

The last series converges at every point $(x,y)\in l_\xi\times l_\eta$. Then

$$\mathbf{E}\left[\frac{\xi}{\eta}\right]=\frac{\mathbf{E}[\xi]}{\mathbf{E}[\eta]}+\frac{\mathbf{E}[\xi]}{\mathbf{E}^3[\eta]}\sigma^2\left[\eta\right]-\frac{1}{\mathbf{E}^2[\eta]}\mathbf{Cov}\left[\xi,\eta\right]+r_1,$$

where

$$r_1=-\mathbf{E}\left[\frac{\mathbf{E}[\xi](\eta-\mathbf{E}[\eta])-\mathbf{E}[\eta](\xi-\mathbf{E}[\xi])}{\mathbf{E}^2[\eta]}\sum_{n=2}^{\infty}\left(\frac{\mathbf{E}[\eta]-\eta}{\mathbf{E}[\eta]}\right)^n\right]$$

and $|r_1| \le \frac{\delta}{1-\delta}\frac{\mathbf{E}[\xi]+\mathbf{E}[\eta]}{\mathbf{E}^3[\eta]}\sigma^2[\eta] \le \frac{\mathbf{E}[\xi]+\mathbf{E}[\eta]}{(1-\delta)\mathbf{E}[\eta]}\delta^3$. The last estimate is followed from inequality $\sigma[\eta] \le \delta \mathbf{E}[\eta]$. The lemma is proved.

We will use formulas (2) and (3) without their residuals. Respective values $\tilde{\mathbf{E}}[\mathrm{M}(A)] = \mathbf{E}[\mathrm{M}(A)] - r_1$, $\tilde{\sigma}^2[\mathrm{M}(A)] = \sigma^2[\mathrm{M}(A)] - r_2$ we will call by estimations of numerical characteristics.

Introduce the following notation: $S(A) = \sum_{i=1}^{|A|} \mathbf{E}[W_{i,i+1}(A)]$, $K(A,X) = \sum_{i=1}^{|A|} k_i^X(A)$, where $k_i^X(A) = \sum_{j=1}^{|X|} \mathbf{Cov}[W_{i,i+1}(A), W_{j,j+1}(X)]$, $A \in 2^X$. Then the formulas for $\tilde{\mathbf{E}}[\mathrm{M}(A)]$ and $\tilde{\sigma}^2[\mathrm{M}(A)]$ based on (2) and (3) can be written in the form

$$\tilde{\mathbf{E}}[\mathrm{M}(A)] = \frac{S(A)}{S(X)} + \frac{S(A)}{S^3(X)}K(X,X) - \frac{1}{S^2(X)}K(A,X), \tag{4}$$

$$\tilde{\sigma}^2[\mathrm{M}(A)] = \frac{1}{S^2(X)}K(A,A) + \frac{S^2(A)}{S^4(X)}K(X,X) - \frac{2S(A)}{S^3(X)}K(A,X). \tag{5}$$

In general, the random variable $\sum_{k=1}^{|A|} W_{k,k+1}(A)$ is not satisfied to conditions of Lemma 1. However the probability of large deviations of random length of noisy polygonal line from non-noisy length will be small if the variance of noise is small. Therefore we assume that the random length satisfied approximately to conditions of Lemma 1.

4 Stochastic Informational Measure by Contour Length

Assume that an original contour is corrupted by noise. In this case, $X = \{x_k + \mathbf{n}_k\}_{k=1}^m$, $x_k \in \mathbb{R}^2$ and $\mathbf{n}_k = (\xi_k, \eta_k)$ are random variables. Suppose also that ξ_k, η_k, $k = 1, ..., m$, are independent, normally distributed and such that $\mathbf{E}[\xi_k] = \mathbf{E}[\eta_k] = 0$, $\sigma^2[\xi_k] = \sigma^2[\eta_k] = \sigma^2$, $k = 1, ..., m$. In this section we consider a monotone measure μ and monotone stochastic measure M of view (1), where $W_{k,k+1}(A) = |x_{k+1}(A) + \mathbf{n}_{k+1}(A) - x_k(A) - \mathbf{n}_k(A)|$ and $w_{k,k+1}(A) = |x_{k+1}(A) - x_k(A)|$ correspondingly. We investigate its characteristics $\tilde{\mathbf{E}}[\mathrm{M}(A)]$ and $\tilde{\sigma}^2[\mathrm{M}(A)]$. Suppose that $W_{k,k+1}(X)$, $k = 1, ..., m$, are independent random variables. This requirement can be satisfied by the choice of some subcontour (basic contour) from the initial contour.

4.1 Numerical Characteristics of Random Variable $W_{k,k+1}(A)$

Let $\mathbf{l}_A(x) = x_+(A) - x$, where $x_+(A)$ is the next point w.r.t. x in the contour A, $l_A(x) = |\mathbf{l}_A(x)|$.

Proposition 1. *The following asymptotic equalities are valid*

$$\mathbf{E}\left[W_{k,k+1}(A)\right] = l\left(1+\tfrac{\sigma^2}{l^2}+\tfrac{\sigma^4}{2l^4}+O\left(\tfrac{\sigma^6}{l^6}\right)\right),$$

$$\sigma^2\left[W_{k,k+1}(A)\right] = 2\sigma^2\left(1-\tfrac{\sigma^2}{l^2}+O\left(\tfrac{\sigma^4}{l^4}\right)\right),\ l=l_A(x_k).$$

Proof. Assume that $x_{k+1}(A)-x_k(A)=(l,0)$, $\mathbf{n}_{k+1}(A)-\mathbf{n}_k(A)=(\xi_{k+1}-\xi_k,\eta_{k+1}-\eta_k)$. Denote $\xi=\xi_{k+1}-\xi_k$ and $\eta=\eta_{k+1}-\eta_k$, $\theta=W_{k,k+1}(A)=\sqrt{(\xi+l)^2+\eta^2}$. Then ξ, η are independent normally distributed random variables and such that $\mathbf{E}[\xi]=\mathbf{E}[\eta]=0$ and $\sigma^2[\xi]=\sigma^2[\eta]=2\sigma^2$. Let $u=\frac{1}{l}$. Then $\theta^2=l^2\left(1+2\xi u+\ \left(\xi^2+\eta^2\right)u^2\right)$. Let us find the representation of $\theta^*=\theta/l$ by Taylor formula at the point $u=0$: $\theta^*(0)=1$, $\theta^{*\prime}(0)=\xi$, $\theta^{*\prime\prime}(0)=\eta^2$, $\theta^{*\prime\prime\prime}(0)=-3\xi\eta^2$, $\theta^{*(4)}(0)=12\xi^2\eta^2-3\eta^4$, $\theta^{*(5)}(0)=-60\xi^3\eta^2+45\xi\eta^4$, $\theta^{*(6)}(0)=360\xi^4\eta^2-540\xi^2\eta^4+45\eta^6$. Therefore

$$\theta^*(u)=1+\xi u+\frac{\eta^2}{2}u^2-\tfrac{\xi\eta^2}{2}u^3+\tfrac{4\xi^2\eta^2-\eta^4}{8}u^4-\tfrac{4\xi^3\eta^2-3\xi\eta^4}{8}u^5+O(u^6).$$

We compute next $\mathbf{E}[\theta^*(u)]$ taking in account that $\mathbf{E}[\xi^s]=\mathbf{E}[\eta^s]=0$ if s is odd, $\mathbf{E}\left[\xi^2\right]=\mathbf{E}\left[\eta^2\right]=\sigma^2$, $E\left[\xi^4\right]=E\left[\eta^4\right]=3\sigma^4$, $E\left[\xi^6\right]=E\left[\eta^6\right]=15\sigma^6$, and that we should compute the expectation of product of independent random variables. Then we have $\mathbf{E}[\theta^*(u)]=1+\frac{1}{2}\sigma^2u^2+\frac{1}{8}\sigma^4u^4+O(u^6)$, since $E\left[\theta^{*(6)}(0)\right]=135\neq 0$. Compute the variance of θ: $\mathbf{E}\left[\theta^2\right]=\mathbf{E}\left[\xi^2+2\xi l_k+l^2+\eta^2\right]=2\sigma^2+l^2$, $\mathbf{E}^2[\theta]=l^2\left(1+\frac{\sigma^2}{l^2}+O\left(\frac{\sigma^4}{l^4}\right)\right)$, therefore, $\sigma^2[\theta]=\mathbf{E}\left[\theta^2\right]-\mathbf{E}^2[\theta]=\sigma^2\left(1-\frac{\sigma^2}{2l^2}+O\left(\frac{\sigma^4}{l^4}\right)\right)$. The general case is also true, because values $\mathbf{E}[W_{k,k+1}(A)]$, $\sigma^2[W_{k,k+1}(A)]$ do not depend on the chosen coordinate system.

Corollary 1. *It is true the equality*

$$S(A)=\sum_{k=1}^{|A|}\mathbf{E}\left[W_{k,k+1}(A)\right]=L(A)+\sigma^2\sum_{k=1}^{|A|}l_A^{-1}(x_k)+\sigma O\left(\tfrac{\sigma^3}{\underline{l}_A^3}\right),$$

where $L(A)=\sum_{k=1}^{|A|}l_A(x_k)$ *is the length of contour* A, $\underline{l}_A=\min_k l_A(x_k)$.

By analogy with Proposition 1 and Corollary 1 we compute the covariance between random variables $W_{k-1,k}(A)$, $W_{k,k+1}(A)$.

Proposition 2. *We have*

$$\mathbf{Cov}\left[W_{k-1,k}(A),W_{k,k+1}(A)\right]=$$

$$=-\sigma^2\cos\alpha_k\left(1-\left(\tfrac{1}{l_{k-1}^2}+\tfrac{\cos\alpha_k}{2l_{k-1}l_k}+\tfrac{1}{l_k^2}\right)\sigma^2+\ o\left(\tfrac{\sigma^2}{l^2}\right)\right),$$

where $\alpha_k=\alpha(x_k)=\left(\widehat{\mathbf{l}_{i-1},\mathbf{l}_i}\right)$, $l_k=l_A(x_k)$, $l=\min\{l_{k-1},l_k\}$.

Calculate the covariance $K(A,X) = \sum_i k_i^X(A)$ between the all segments of polygon A and all segments of basic polygon X with help of last proposition. Let $\alpha(x)$ ($\beta(x)$) be an inner angle of polygon A (polygon X) in vertex x, $\gamma(x)$ be an angle between the vectors $x_{+1}(A)-x$, $x_{+1}(X)-x$, where $x_{+1}(A)$ ($x_{+1}(X)$) is the next point w.r.t. x in the contour A (contour X).

Corollary 2. *To same conditions it is true equality*

$$K(A,X) = 4\sigma^2 \sum_{x\in A} \cos\tfrac{\alpha(x)}{2} \cos\tfrac{\beta(x)}{2} \cos\left(\gamma(x) + \tfrac{\alpha(x)-\beta(x)}{2}\right) + \sigma^2 o\left(\tfrac{\sigma}{\underline{l}_A}\right)$$

for $A \in 2^X$*, where* $\underline{l}_A = \min_k l_A(x_k)$.

4.2 The Numerical Characteristics of Stochastic Measure of Informativity by Length

We will find numerical characteristics of stochastic measure of informativity by length using the results of the previous item. The following theorem may be got from equality (4), Corollaries 1, 2.

Theorem 1. *The asymptotic equality*

$$\tilde{\mathbf{E}}\,[\mathrm{M}(A)] = \tfrac{L(A)}{L(X)} + C_1(A)\tfrac{\sigma^2}{L^2(X)} + o\left(\tfrac{\sigma^2}{\underline{l}_A^2}\right), A \in 2^X$$

is true, where

$$C_1(A) = -L(A)\sum_{x\in X} l_X^{-1}(x) + L(X)\sum_{x\in A} l_A^{-1}(x) + 4\tfrac{L(A)}{L(X)}\sum_{x\in X}\cos^2\tfrac{\beta(x)}{2} -$$

$$-4\sum_{x\in A}\cos\tfrac{\alpha(x)}{2}\cos\tfrac{\beta(x)}{2}\cos\left(\gamma(x) + \tfrac{1}{2}\alpha(x) - \tfrac{1}{2}\beta(x)\right).$$

Similarly we will find the asymptotic formula for variance of stochastic informational measure by length with help of formula (5), Corollaries 1, 2.

Theorem 2. *The asymptotic equality*

$$\tilde{\sigma}^2\,[\mathrm{M}(A)] = 4C_2(A)\tfrac{\sigma^2}{L^2(X)} + o\left(\tfrac{\sigma^2}{\underline{l}_A^2}\right), \quad A \in 2^X$$

is true, where

$$C_2(A) = \sum_{x\in A}\cos^2\tfrac{\alpha(x)}{2} + \tfrac{L^2(A)}{L^2(X)}\sum_{x\in X}\cos^2\tfrac{\beta(x)}{2} -$$

$$-2\tfrac{L(A)}{L(X)}\sum_{x\in A}\cos\tfrac{\alpha(x)}{2}\cos\tfrac{\beta(x)}{2}\cos\left(\gamma(x)+\tfrac{1}{2}\alpha(x)-\tfrac{1}{2}\beta(x)\right).$$

The value of random error (the variance of stochastic informational measure) characterizes the degree of stability of informational measure of curve with respect to level of curve noise. We can put the task about finding of polygonal representation of fixed cardinality $A \in 2^X$, $|A| = k$, which minimized the value of variance of stochastic informational measure by length. As can be seen from Theorem 2 the polygonal representation

$$A = \arg \min_{A \in 2^X,\ |A|=k} C_2(A)$$

is a solution of indicated task for great signal-to-noise ratio $\frac{l_A^2}{\sigma^2}$.

Example 1. Let $X = \{x_1, ..., x_6\}$ be an ordered set of vertexes of regular 6-gon. Calculate the value $C_2(A)$ for various polygonal representations A of cardinality $|A| = 3$: $A_1 = \{x_1, x_3, x_5\}$, $A_2 = \{x_1, x_2, x_4\}$, $A_3 = \{x_1, x_2, x_3\}$. We have $C_2(A_1) = 1.125$, $C_2(A_2) = 1.25$, $C_2(A_3) \approx 1.66$. Thus the contour A_1 is a most stable contour to noise w.r.t. measure of informativity by length among of contours of cardinality is equal 3.

Many other tasks of finding of informative and stable representation of noisy image may be formulated and solved with help of this approach.

Acknowledgements I would like to thank Andrew Bronevich for his helpful and stimulating comments on the manuscript of my paper. The study was implemented in the framework of The Basic Research Program of the Higher School of Economics in 2012. This work was supported by the grants 11-07-00591 and 10-07-00135 of RFBR (Russian Foundation for Basic Research).

References

1. Bronevich A, Lepskiy A (2003) Geometrical fuzzy measures in image processing and pattern recognition. In: Proc. of the 10th IFSA World Congress. Istanbul, 151–154
2. Duda RO, Hart PE, Stork DG (1998) Pattern Classification and Scene Analysis: Part I Pattern Classification. John Wiley & Sons
3. Jolliffe IT (2002) Principal Component Analysis. Series: Springer Series in Statistics. Springer, Berlin Heidelberg New York
4. Lepskii AE (2007) On Stability of the Center of Masses of the Vector Representation in One Probabilistic Model of Noiseness of an Image Contour. Automation and Remote Control. 68:75–84
5. Lepskiy AE (2011) Stable Feature Extraction with the Help of Stochastic Information Measure, Lecture Notes in Computer Science(LNCS), vol. 6744, Springer, Berlin Heidelberg New York, 54–59
6. Shiryaev AN (1995) Probability. Graduate Texts in Mathematics. Springer, New York
7. Wang Z, Klir GJ (2009) Generalized Measure Theory. IFSR International Series on Systems Science and Engineering. 25, Springer, Berlin Heidelberg New York