Applying doubly stochastic filters to evaluate the dynamics of object sizes on satellite image sequences

V E Dementyev[1], D S Kondratyev[1]

[1]Ulyanovsk State Technical University, ul. Severny Venets, 32, Ulyanovsk, Russia, 432027

e-mail: dve@ulntc.ru

Abstract. One of the important tasks facing the regional authorities is to monitor the condition of roads and power lines. In the Ulyanovsk region more than 8 thousand km of power lines and more than 9 thousand km of roads (including rural). A significant part of these facilities is located outside the settlements in places with medium and low availability. In many such places there is a problem of uncontrolled forest overgrowth. This work is devoted to solving the problem of automated satellite monitoring of such areas. For this purpose, it is proposed to use a modified convolutional neural network that processes time sequences of multispectral satellite images and allows to allocate territories occupied by forest and undergrowth with high accuracy. This approach allows us to assess the dynamics of overgrowth of the territory and perform the appropriate forecast with sufficient accuracy for practice.

1. Introduction

One of the important tasks of the satellite image processing is its thematic mapping, i.e. division of image into identifiable areas containing the similar visual, correlation or texture characteristics of pixels. The use of standard segmentation algorithms [1-3] for thematic mapping of satellite images usually leads to significant errors caused by two reasons. Firstly, these algorithms are largely incapable of taking into account the multi-zonal nature of remote sensing (RS), so each satellite image contains the results of the Earth's surface registration in different spectral ranges. Some works [4-6] suggest the possibility of processing hyperspectral images. Thus, the authors based their theory on criterion of uniformity for reception of connected areas of such hyperspectral image [4], modification and generalization of algorithm K-means [5] and use of physical properties of a satellite data [6]. Secondly, the existing approaches unable to use data on the observed territory received at previous points of time for image segmentation. But using such data can significantly improve the quality of processing at the expense of a fundamentally larger amount of information, but it is fraught with difficulties of aggregation.

It is possible to overcome the mentioned disadvantages by using neural network procedures of segmentation and classification of multidimensional data. In work [9] the variant of the U-NET neural network with full-connected layers (FCN) modification is presented. Herewith the input layer of the network consisting from spectral layers of the multizonal image has been extended by three auxiliary 2d halftone images obtained from the original using NDVI, EVI, SAVI transformations and two 2d arrays , representing the segmentation results at the previous point of time and one year ago. The use of two such reference markings allows one's to reduce the error of classification in case of rapid changes of the terrain due to the change of year time and in case of marking array absence at the

previous point of time in connection, for example, with cloudiness. The quality analysis of such an algorithm shows the processing comparable accuracy to the qualified operator results.

It should be noted that the results of the satellite images thematic processing provides information related to the peculiarities of the Earth's surface in the previous and current moments of time. However, the necessary element of the RS processing system is the tool that allow to forecast the state of certain objects and form recommendations for responsible persons. For example, let's consider the task related to the forecasting about the dangerous convergence possibility of various natural and technogenic objects. An example of such convergence may be the gradual growth of the forest area along roads or power lines or landslide processes leading to the destruction of various infrastructures.

2. Satellite Image Processing

Let us formulate the given problem as follows. Let there be an aggregate of segments describing a certain extended object $\mathbf{T_o}$ (a road, an electric power transmission line etc.). The given aggregate is usually described by a vector object having geographical reference in absolute coordinates. Let us separate on this extended object a set of points $\mathbf{T_{oi}} = (\mathbf{x_{0i}}, \mathbf{y_{oi}})$ having the distance between them equal to Δ_o. Let us assume that at a certain time instant $\mathbf{t}$ next to the object there lies a certain extended domain $\mathbf{G_R}^t$ defined by a set of points each corresponding to a pixel in the original halftone image. For each point $\mathbf{T_{oi}}$ let us construct a perpendicular to the segment $[\mathbf{T_{oi}}, \mathbf{T_{oi+1}}]$. Let us find the point $\mathbf{T_{Ei}}^t = \left(\mathbf{z_{Exi}^t}, \mathbf{z_{Eyi}^t}\right)$ of intersection of this perpendicular and the domain $\mathbf{G_R}^t$. Obviously the set of points $\{\mathbf{T_{Ei}}^t, \mathbf{i} = \mathbf{1}, .., \mathbf{N_o}\}$ describe a conditional boundary of the domain $\mathbf{G_R}^t$ as viewed from the extended object. It enables to use estimates of the points coordinates $\{\mathbf{T_{Ei}}^t, \mathbf{i} = \mathbf{1}, .., \mathbf{N_o}\}$ obtained at different times $\mathbf{t} = \mathbf{1}, .., \mathbf{T}$ as a source of information on the domain $\mathbf{G_R}^t$ dynamics for the purpose of forming a prediction about its boundaries at future times $\mathbf{t} > \boldsymbol{T}$ as well. (Figure 1).
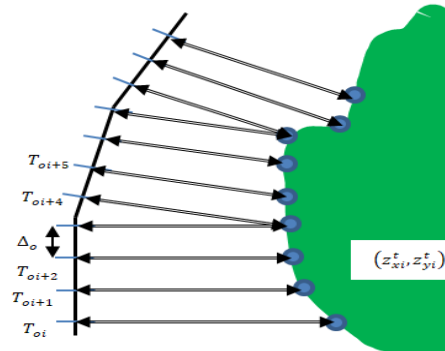
Figure 1. Geometric interpretation of the distance estimate between the stationary object and the movable area.

In view of errors arising at satellite images recording and processing we get the following relations (Eq. 1), in which $\boldsymbol{n_{Exi}^t}$ and $\boldsymbol{n_{Eyi}^t}$ – white noise samples with zero mean and variance $\boldsymbol{\sigma_n^2}$.

$$z_{Exi}^t = x_{Ei}^t + n_{Exi}^t, z_{Eyi}^t = y_{Ei}^t + n_{Eyi}^t, i = 1, .., N_o, t = 1, .., T, \qquad (1)$$

where $\boldsymbol{n_{Exi}^t}$ and $\boldsymbol{n_{Eyi}^t}$ – white noise samples with zero mean and variance $\boldsymbol{\sigma_n^2}$.

Direct measurements based on the results of satellite material image-type related mapping show that MSE $\sqrt{\sigma_n^2}$ is approximately equal to $\mathbf{1.5D_{xy}}$, where $\mathbf{D_{xy}}$ – resolution of the original images.

Let us suppose that the boundary of the domain $\boldsymbol{G_R}^t$ can move non-uniformly. Thus, for example, the area of a precipice or a ravine can increase by tens of centimeters per each year and at a certain moment its rate might increase exponentially. Then we will use doubly stochastic (DS) model [3,10] to describe unknown coordinates in the following form (Eq. 2). In this case $\mathbf{r_{ax}}, \mathbf{r_{ay}}$ – scalar parameters determining change potential for the accelerations $\mathbf{a_{Exi}^t}$ and $\mathbf{a_{Eyi}^t}$; $\boldsymbol{\xi_{axi}^t}, \boldsymbol{\xi_{ayi}^t}$ – independent normal random variables with zero mean and variance $\boldsymbol{\sigma_\xi^2}$:

$$x_{Ei}^t = 2x_{Ei}^{t-1} - x_{Ei}^{t-2} + a_{Exi}^t\left(x_{Ei}^{t-1} - x_{Ei}^{t-2}\right),$$

$$y_{Ei}^t = 2y_{Ei}^{t-1} - y_{Ei}^{t-2} + a_{Eyi}^t(y_{Ei}^{t-1} - y_{Ei}^{t-2}), \tag{2}$$
$$a_{Exi}^t = r_{ax}a_{Exi}^{t-1} + \xi_{axi}^t, \quad a_{Eyi}^t = r_{ay}a_{Eyi}^{t-1} + \xi_{ayi}^t,$$

The model (2) can be rewritten in the form:

$$\bar{X}_{Ei}^t = \varphi_{Exi}^t(\bar{X}_{Ei}^{t-1}) + \bar{\xi}_{xi}^t, \quad \bar{Y}_{Ei}^t = \varphi_{Eyi}^t(\bar{X}_{Ei}^{t-1}) + \bar{\xi}_{yi}^t,$$

where

$$\varphi_{Exi}^t(\bar{X}_{Ei}^{t-1}) = \begin{pmatrix} x_{Ei}^{t-1} & v_{Exi}^t & 0 \\ 0 & v_{Exi}^{t-1}(1+a_{Exi}^t) & 0 \\ 0 & 0 & a_{Exi}^{t-1}r_{ax} \end{pmatrix},$$

$$\varphi_{Eyi}^t(\bar{Y}_{Ei}^{t-1}) = \begin{pmatrix} y_{Ei}^{t-1} & v_{Eyi}^t & 0 \\ 0 & v_{Eyi}^{t-1}(1+a_{Eyi}^t) & 0 \\ 0 & 0 & a_{Eyi}^{t-1}r_{ay} \end{pmatrix},$$

$$\bar{\xi}_{xi}^t = \begin{pmatrix} 0 \\ 0 \\ \xi_{axi}^t \end{pmatrix}, \quad \bar{\xi}_{xi}^t = \begin{pmatrix} 0 \\ 0 \\ \xi_{ayi}^t \end{pmatrix}.$$

The assigned notations enable to apply DS nonlinear filtering to observations and for construction of predictions of behaviour of the region R_0 in reference to the object T_o. In doing so let us introduce $\bar{X}_{\Im Ei}^t = \varphi_{Exi}^t(\bar{X}_{Ei}^{t-1})$ and $\bar{Y}_{Ei}^t = \varphi_{Eyi}^t(\bar{X}_{Ei}^{t-1})$– extrapolated predictions for the point T_{Ei} coordinates at the time point t based on the preceding observations z_{Exi}^{t-1} and z_{Eyi}^{t-1}. Denote by P_{xi}^{t-1}, P_{yi}^{t-1} – filtering error covariance matrices at the time point $(t-1)$, $V_{x\xi i}^t = M\{\bar{\xi}_{xi}^t \bar{\xi}_{xi}^{t\,T}\}$, $V_{y\xi i}^t = M\{\bar{\xi}_{yi}^t \bar{\xi}_{yi}^{t\,T}\}$ – diagonal covariance matrices for random increments $\bar{\xi}_{xi}^t$. Then error covariance matrices for such extrapolation have the following form:

$$P_{\Im xi}^t = M\left\{\left(\hat{\bar{X}}_{\Im Ei}^t - \bar{X}_{\Im Ei}^t\right)\left(\hat{\bar{X}}_{\Im Ei}^t - \bar{X}_{\Im Ei}^t\right)^T\right\} = \varphi_{Exi}^{t\,\prime}(\bar{X}_{Ei}^{t-1})P_x^{t-1}\varphi_{Exi}^{t\,\prime}(\bar{X}_{Ei}^{t-1})^T + V_{x\xi i}^t,$$

$$P_{\Im yi}^t = M\left\{\left(\hat{\bar{Y}}_{\Im Ei}^t - \bar{Y}_{\Im Ei}^t\right)\left(\hat{\bar{Y}}_{\Im Ei}^t - \bar{Y}_{\Im Ei}^t\right)^T\right\} = \varphi_{Eyi}^{t\,\prime}(\bar{Y}_{Ei}^{t-1})P_{yi}^{t-1}\varphi_{Eyi}^{t\,\prime}(\bar{Y}_{Ei}^{t-1})^T + V_{y\xi i}^t,$$

If we denote by $\hat{x}_{\Im Ei}^t, \hat{y}_{\Im Ei}^t$ - the first elements of the vectors $\hat{\bar{X}}_{\Im Ei}^t$ and $\hat{\bar{Y}}_{\Im Ei}^t$; $B_{xi}^t = P_{\Im xi}^t C_x^T D_{xi}^{t\,-1}$; $B_{yi}^t = P_{\Im yi}^t C_y^T D_{yi}^{t\,-1}$; $D_{xi}^t = C_x P_{\Im xi}^t C_x^T + \sigma_n^2$; $D_{yi}^t = C_y P_{\Im yi}^t C_y^T + \sigma_n^2$; $C_x = C_y = (1 \quad 0 \quad 0)$, then we can write the following relations (Eq. 27) for DS coordinate filters:

$$\hat{\bar{X}}_{Ei}^t = \hat{\bar{X}}_{\Im Ei}^t + B_{xi}^t(z_{Exi}^t - \hat{x}_{\Im Ei}^t), \quad \hat{\bar{Y}}_{Ei}^t = \hat{\bar{Y}}_{\Im Ei}^t + B_{yi}^t(z_{Eyi}^t - \hat{y}_{\Im Ei}^t). \tag{3}$$

The filtering error variance at each step is determined by the matrices $P_{xi}^t = (E - B_{xi}^t)P_{\Im xi}^t$, $P_{yi}^t = (E - B_{yi}^t)P_{\Im yi}^t$.

Special attention must be given to the fact that due to the necessity for the distance from the extended object T_o to the domain R_0 to be surveyed, it is possible to simplify the boundaries coordinates filtering process R_0 by processing only one coordinate, namely, the distance x_{ri}^t from the point T_{oi} belonging to T_o up to the boundary R_0 at the time point t. An aggregate of similar observations $z_{Ex_r i}^t$ can be processed by a technique identical to the above-described one.

As an illustration of such a technique in figures below series satellite images fragments for the forest tract in Cherdakly district of the Ulyanovsk region for the period 2001-2017 years (figure 2) and Milanovsky opencast colliery on riverbank of the Volga in the northern part of the city of Ulyanovsk for the period 2013-2017 years (figure 3) are presented. Here for convenience of color image perception and its recovery overlapping of visible spectral bands and superposition of the segmented image fragment and normals to the object to be monitored is carried out. In the first case the number of multispectral images to be processed amounted to 42 snapshots, in the second case - 32 snapshots. The minimal time interval for satellite information production amounts to 14 days.

<p align="center">July 2003   June 2011   July 2017</p>
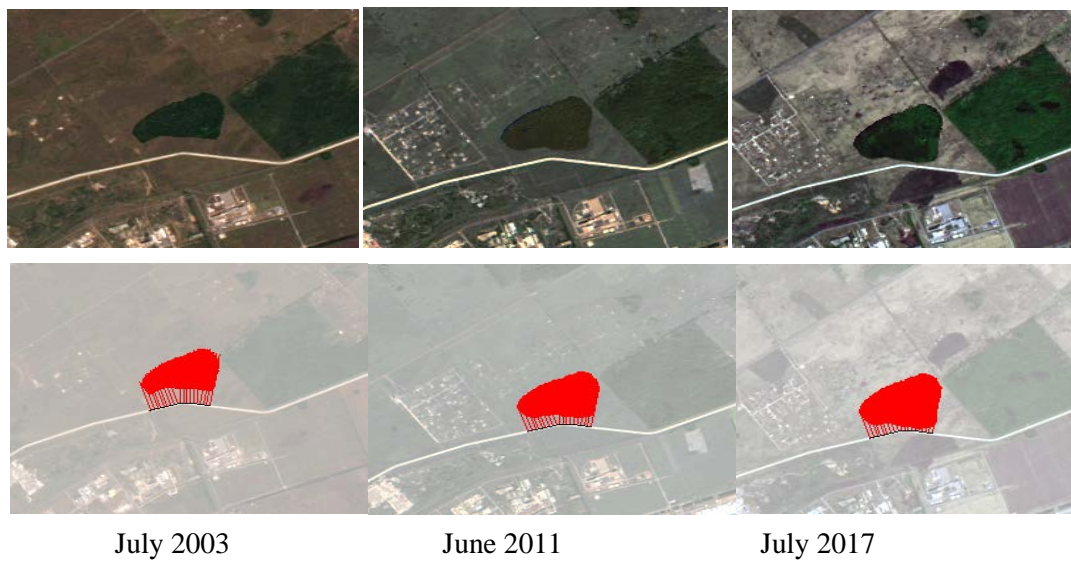
Figure 2. Forest tract satellite images fragments and the results of these images processing.

<p align="center">August 2015   April 2016   May 2016</p>

Figure 3. Milanovsky opencast colliery satellite images fragments and the results of these images processing.

The above-mentioned groups of multispectral images were separated into a training and an operating samples. The training samples were used to specify the filtering parameters, in particular, to estimate the parameters $\mathbf{r_{ax}}$ and $\boldsymbol{\sigma_\xi^2}$. The operating part of the sample was processed by three algorithms enabling to carry out the distance prediction $\mathbf{x_{ri}^{t+1}}$ basing on the preceding observations. The first algorithm (I) involves constructing a simple prediction $\mathbf{\hat{x}_{ri}^{t+1} = 2z_{ri}^t - z_{ri}^t}$ wherein only the variable $\mathbf{\hat{x}_{ri}}$ change speed is taken into account for the time interval $\mathbf{(t-1, t)}$. The second algorithm (II) assumes linear Kalman filtering of the observations $\mathbf{z_{ri}^t}$ and construction of the extrapolated prediction $\mathbf{\hat{x}_{\ni ri}^{t+1}}$ based on the results of the processing. The third algorithm (III) is in the above-described doubly stochastic filtering of an aggregate of the observations $\mathbf{z_{ri}^t}$ and construction of the vector $\mathbf{\widehat{\overline{X}}_{\ni Ei}^{t+1}}$. In the table 1 below the values of prediction average errors depending on the object kind are presented.

On average the DS filter provides prediction accuracy 6% higher than in case of using Kalman filter and 58% higher than in case of simple linear predictions. It enables to estimate coordinates and rate change dynamics for boundaries of the domain to be monitored by using the DS filter. It is

essential that the DS filter enables to quicker respond to abrupt rate change of the processes determining the object behaviour. As an illustration we provide the estimates behaviour for the distance from opencast colliery to one of the points to be monitored (figure 4a) and the parameter $\hat{a}_i^t$ estimate (figure 4b).

Table 1. Prediction average errors.

	Average error for the algorithm I	Average error for the algorithm II	Average error for the algorithm III
The forest tract snapshot. October 2014	6.7 m	2.7 m	2.6 m
The forest tract snapshot. May 2015	10.7 m	3.9 m	3.7 m
The forest tract snapshot. June 2016	6.2 m	3.6 m	3.3 m
Milanovsky opencast colliery snapshot. May 2014	7.1 m	3.8 m	3.6 m
Milanovsky opencast colliery snapshot. May 2015	7.3 m	3.9 m	3.8 m
Milanovsky opencast colliery snapshot. May 2016	6.9 m	3.9 m	3.7 m
Milanovsky opencast colliery snapshot. April 2016. Beginning of the avalanche processes.	12.4 m	8.9 m	7.8 m
Milanovsky opencast colliery snapshot. April 2016. Continuation of the avalanche processes.	30.7 m	32.8 m	12.6 m
Milanovsky opencast colliery snapshot. May 2016. Cessation of the avalanche processes.	20.3 m	18.1 m	17.3 m
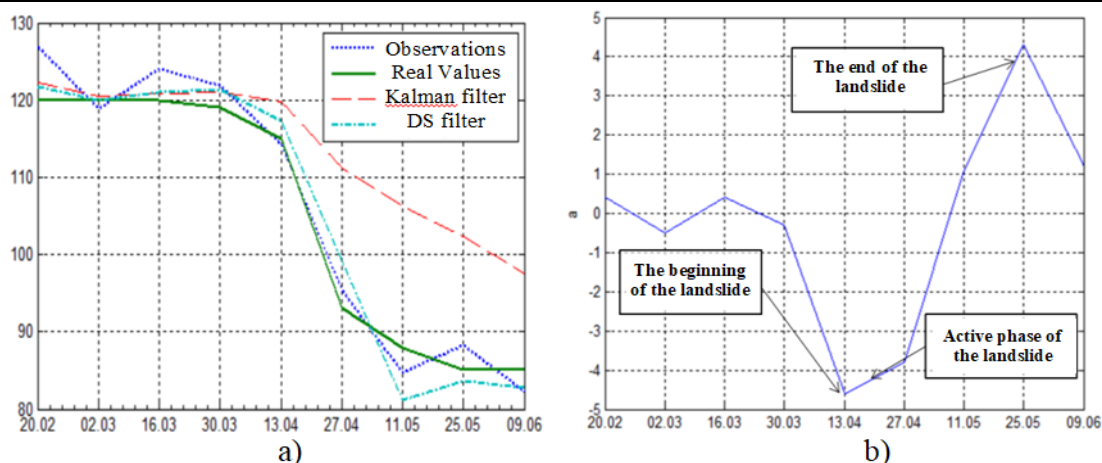
Figure 4. Dependence of the filtering results versus survey time.

3. Conclusion

Direct analysis of the given results in comparison with the data of objective monitoring (solid line) indicates superiority of the DS filter over conventional linear Kalman filter in filtering accuracy. As it takes place, this superiority makes itself evident in the most distinct manner in case of abrupt change of the rock collapsing process (and the corresponding reduction of the distance between the opencast

colliery and the point to be monitored). This change corresponds to a significant change of the parameter a_i^t estimate which enables to register considerable changes in the opencast colliery domain state basing only on this estimate dependence nature versus survey time.

4. References

[1]   Vasiliev K K 2016 *Optimum processing of signals in discrete time* (Moscow: Radiotehnika)
[2]   Krasheninnikov V R and Vasil'ev K K 2018 Multidimensional image models and processing *Intelligent Systems Reference Library* **135** 11-64
[3]   Vasiliev K K and Dementyev V E 2017 *Presentation and processing of satellite multi-zone images* (Ulyanovsk)
[4]   Fursov V A, Bibikov S A and Baida O A 2014 Thematic classification of hyperspectral images by conjugacy index *Computer Optics* **38(1)** 154-158
[5]   Zimichev E A, Kazansky N L and Serafimovich P G 2014 Spatial classification of hyperspectral images using the k-means ++ clustering method *Computer Optics* **38(2)** 281-286
[6]   Blokhinov Yu B, Gorbachev V A, Rakutin Yu O and Nikitin A D 2018 Development of an algorithm for semantic segmentation of real-time aerial photographs *Computer Optics* **42(1)** 141-148 DOI: 10.18287/2412-6179-2018-42-1-141-148
[7]   Belim S V and Larionov S V 2016 Image segmentation algorithm based on searching communities on graphs *Computer Optics* **40(6)** 904-910 DOI: 10.18287/2412-6179-2016-40-6-904-910
[8]   Dementyev V E, Kondratyev D S and Frenkel A G 2018 *Application of convolutional neural networks for thematic mapping of sequences of satellite multispectral images* (Moscow: Radiotehnika)
[9]   Dementyev V E 2017 *The use of double-stochastic autoregressive models for processing satellite images* (Moscow: Radiotehnika)
[10]   Andriyanov N A and Dement'ev V E 2018 Application of mixed models of random fields for the segmentation of satellite images *CEUR Workshop Proceedings* **2210** 219-226