

V.A. DUSHEPA, Y.A. TIAHNYRIADNO, I.V. BARYSHEV, D. Sci.

IMAGE REGISTRATION COMPARATIVE ANALYSIS: NORMALIZED CORRELATION VERSUS SIFT-BASED REGISTRATION

Introduction

Until now, a large variety of different image registration algorithms have been developed. They can usually be divided into two groups: area- (or intensity-) based and feature-based [1, 2]. It should also be noted that a new direction is actively developing now when trained neural networks are used as a measure of similarity [3].

Intensity-based algorithms are commonly used to determine the shift between images when there is little or no rotation and scale changes. If the considered model of geometric transformation between images is more complex (where, for example, the presence of rotation and change in scale cannot be neglected), then either the reliability of such algorithms drops sharply, or an excessive amount of computation is required for their normal operation (because it is necessary to iterate over all possible sets of transformation parameters, each time calculating the correlation measure of similarity). The most common among this group of methods are correlation methods.

At the same time, feature-based methods work quite stably with a variety of models of geometric transformations. However, their accuracy is inferior to correlation methods. Also, they are often inferior in speed [4]. Among the methods using the calculation of features, the most important place is occupied by the methods based on keypoints. In this case, the search (detection) of keypoints is first carried out on two images that are registered. Let's call one of them (usually larger) the reference image (RI), and the second image – the current image (CI). Then a descriptor is calculated for each keypoint. A descriptor is some array of values that describes this keypoint. As a result, we will have a set of descriptors for RI and a set of descriptors for CI. If we find correspondences between them, then we can estimate the geometric transformation between the images (using the found correspondences between the keypoints).

For comparison, from the group of intensity-based methods, the classic variation of the normalized cross-correlation method (further NCC to denote an algorithm) was chosen [5, 6], one of the most popular correlation methods. As a representative of the second group of methods, an algorithm that uses a SIFT detector (and a keypoint descriptor) was chosen [7]. SIFT was created at the turn of the 21st century. And, although since then there have been many different improvements to the original algorithm (for example, in [8]), this work uses the classic implementation available in the OpenCV package [9]. This should be enough to investigate the basic patterns and obtain approximate quantitative characteristics.

Aim and tasks of the research. The purpose of this work is to compare qualitatively and quantitatively (including computational performance) algorithms from two main classes. There are many works making comparisons within each class (for example for intensity-based – [10], feature-based – [11]) but almost no between different classes. A comparison of the two algorithms considered in the article will allow, in addition to confirming the obvious general patterns, to obtain specific indicators of quality and reliability, which are useful to understand when choosing a tool for registering images.

Also in the second part of the work, it is shown (and quantified) that the use of a combined algorithm (using the SIFT at the first stage, and the NCC at the second stage) allows to join the advantages of both approaches and achieve high estimation accuracy for complex geometric transformation models.

The novelty of this work is not the ideas of the approaches themselves, which are generally known, but specific numerical performance indicators calculated by a simulation experiment (including real terrain maps as images that are registered).

1. Simulation setup

The comparison was performed using a simulation model created in the Python programming language (with using OpenCV, SciPy [12], and NumPy [13] libraries). The structure of the computational experiment is shown in Fig. 1.

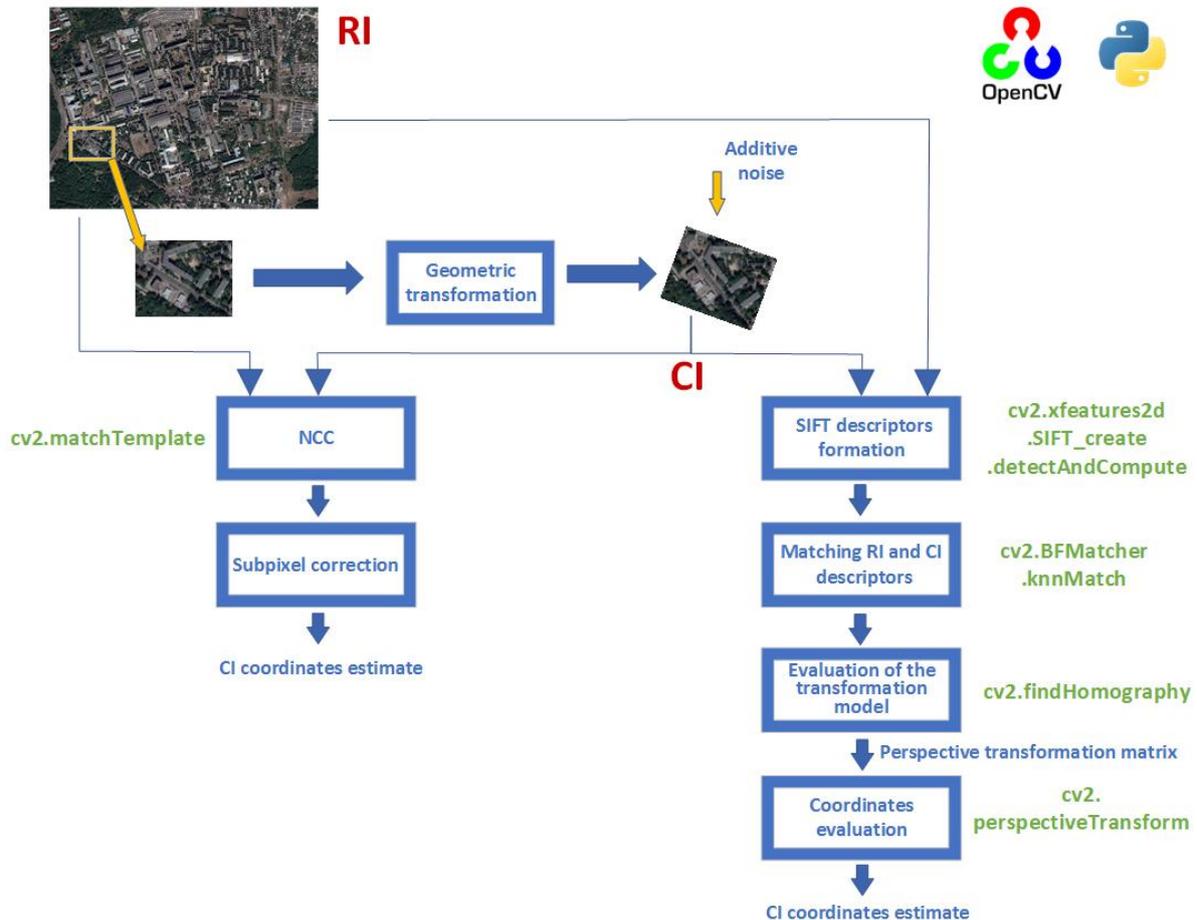


Fig. 1. The structure of the simulation experiment (the elements of the OpenCV library that implement each of the stages are indicated)

First, a reference image is selected, for example, a satellite image (or its region). Then the model randomly selects a fragment of a given size (in section 2.1 the fragment of 200 by 200 pixels was used, and 50x50 – in the 2.2 section) of this image, adds distortion (rotate, scale, add random subpixel shift and noise). So we got CI. We passed RI and CI to both registration algorithms (NCC and SIFT-based). The coordinates of the CI relative to RI were estimated. We considered the position of the CI center to be the coordinates of the CI.

In the CI formation model, we calculated new pixel positions after geometric transformation and then used bicubic interpolation to find their intensities. In all experiments, the values of the subpixel shifts along X and Y were modeled with a uniform distribution in the intervals from -0.5 to 0.5.

The `matchTemplate` function (from OpenCV) performed an NCC evaluation. Also for the NCC, the gradient method of subpixel shift correction was used [5].

Implementations of SIFT feature selection algorithms and feature matching of two images are taken also from the OpenCV library. The descriptors are first calculated using the `detectAndCompute` method of the `SIFT_create` class. The resulting descriptors of the two images (RI and CI) are then matched. OpenCV implements two matching methods, `BFMatcher` and `Flann`, and all experiments in the work were performed using the `BFMatcher` class, which showed higher performance than `Flann` in our case. After that, the `findHomography` and `getPerspectiveTransform` functions are used to calculate the coordinates of the CI relative to RI.

The process of forming a CI and searching for it in the original image (RI) was repeated 100 times for each of the parameter sets. Based on these tests, the probability of correct registration (P) and the root-mean-square error (RMSE) were then calculated. The registration was considered erroneous when the obtained estimate of the CI position differed from the true value by more than 2 pixels. All of the following results were obtained with a signal-to-noise ratio of 20 (the standard deviation of the image was 20 times greater than the noise).

2. Experiment results

2.1. NCC and SIFT for real images registration

We used two main images (as RI) in our experiments: a real terrain snapshot of the National aerospace university “Kharkov aviation institute” (downloaded using [14], see Fig. 2, *a*) and a raccoon picture (Fig. 2, *b*).



Fig. 2. Test images (RI): *a* – satellite image of the territory of NAU "KhAI", 1000x1000; *b* – raccoon image, 1280x1080

The SIFT descriptor is a vector with 128 values. For the RI from Fig. 2, *a*, the algorithm finds 17453 keypoints. While for a fragment (CI), the number will vary (depending on the fragment) from 500 to 800 keypoints.

The average runtime for NCC is 0.0264 seconds. For the SIFT-based algorithm, this value is 0.44 seconds. If we consider that the keypoints for the RI are calculated in advance, then the execution time decreases to 0.1 seconds, which is still almost four times longer than for NCC.

At first, the rotation (φ) and the scaling change (s) were not simulated. In this case, the probability of correct matching for both algorithms was equal to 1 for signal-to-noise ratios greater than 10 (modern video cameras provide very low noise levels). The RMSE was less than 0.05 pixels (slightly lower for NCC compared to SIFT). Thus, in the absence of changes in scale and rotation, we can emphasize a slight advantage of the NCC in the accuracy.

Now let's analyze what will happen if there are a rotation and a change in scale (not taken into account in the NCC algorithm).

In Fig. 3 you can see that for NCC the probability of successful registration P begins to decline sharply when the absolute value of rotation angle is more than about 2 degrees for the image in Fig. 2, *a* and 1 degree – in Fig. 2, *b*. In Fig. 4, we see that the accuracy of the NCC registration also worsens when the absolute value of the angle increases. At the same time, the probability of the SIFT-based algorithm is almost always equal to one and the accuracy only slightly deteriorates with increasing angle modulus. If P is equal to zero, then the RMSE cannot be calculated, therefore the RMSE graphs have gaps.

We see a similar situation in the presence of scale distortion (in this work, only figures for the probability of correct registration are provided). With increasing scaling, the results of the NCC drop sharply. This appears earlier for the image in Fig. 2, *a*, which is obviously less "smooth" and accordingly has a higher frequency content.

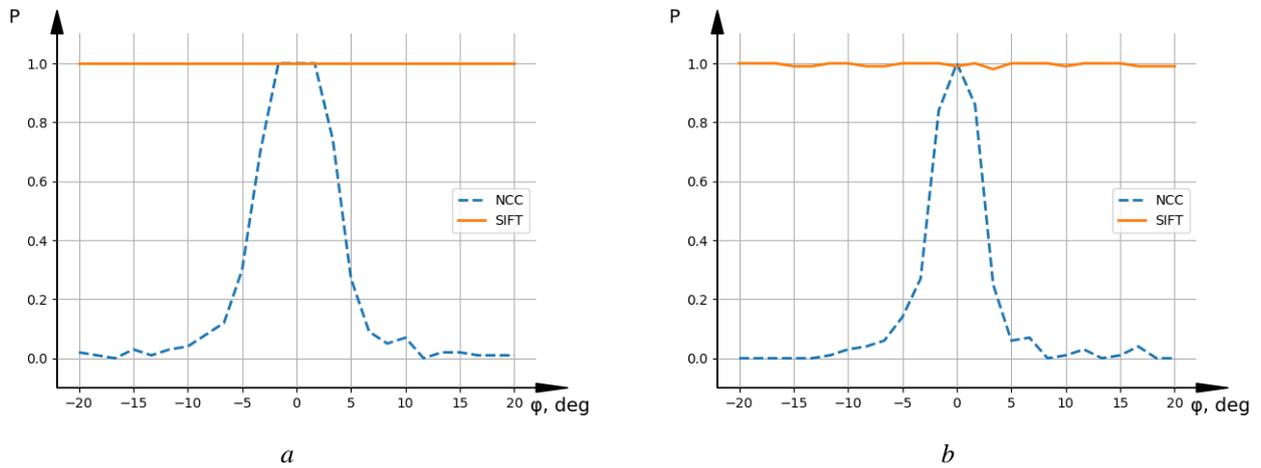


Fig. 3. Effect of rotation φ on the probability of correct registration P : a – for the image in Fig. 2, a ; b – for the image in Fig. 2, b

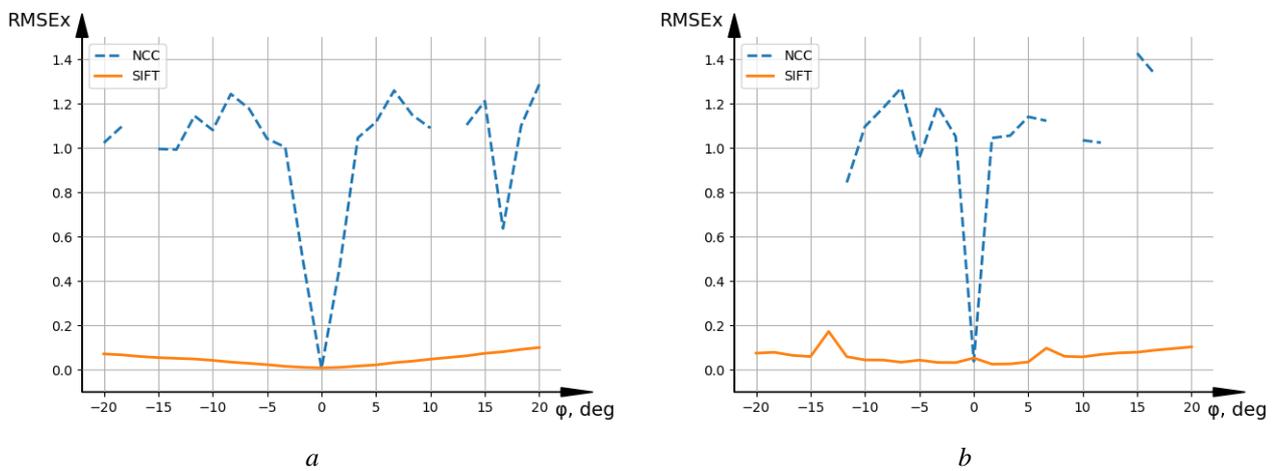


Fig. 4. Effect of rotation φ on the RMSE (on X-axis): a – for the image in Fig. 2, a ; b – for the image in Fig. 2, b

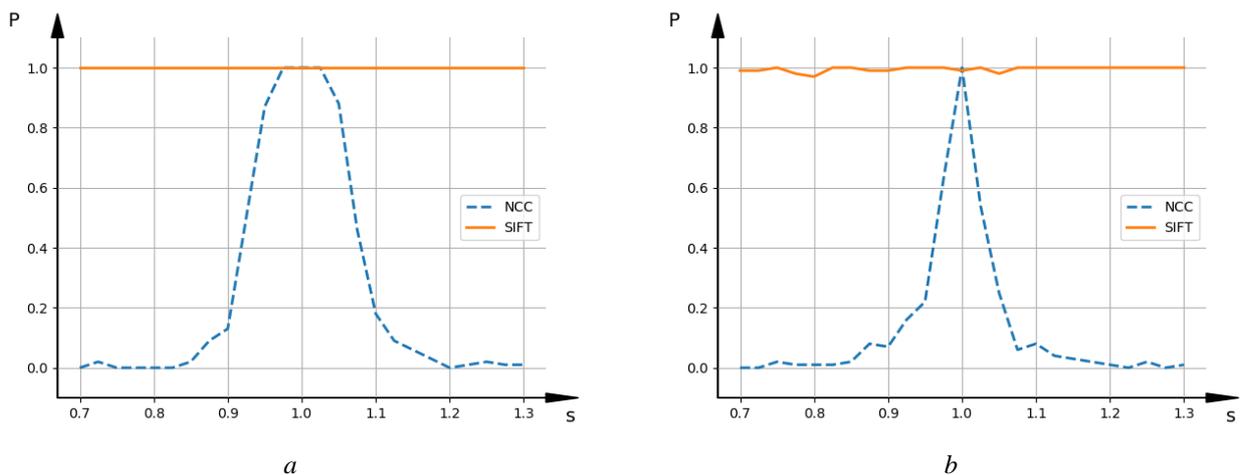


Fig. 5. Effect of scale s on the probability of correct registration P : a – for the image in Fig. 2, a ; b – for the image in Fig. 2, b

2.2. Small CI case (50x50)

If in previous experiments a CI with a size of 200x200 was used, then this checks what will happen with smaller fragments. Fig. 6 shows plots of P versus rotation and scale change (for the image in Fig. 2, a). It can be seen that the SIFT-based algorithm starts to work much worse. This is

because it is not always possible to select a sufficient number of keypoints for such a small CI. Although in such case it is possible to improve the situation somewhat by using keypoint selection settings (for example, changing the threshold for Lowe's ratio test [7], and, accordingly, choosing unreliable keypoints), the SIFT-based algorithm will have problems when working with small fragments.

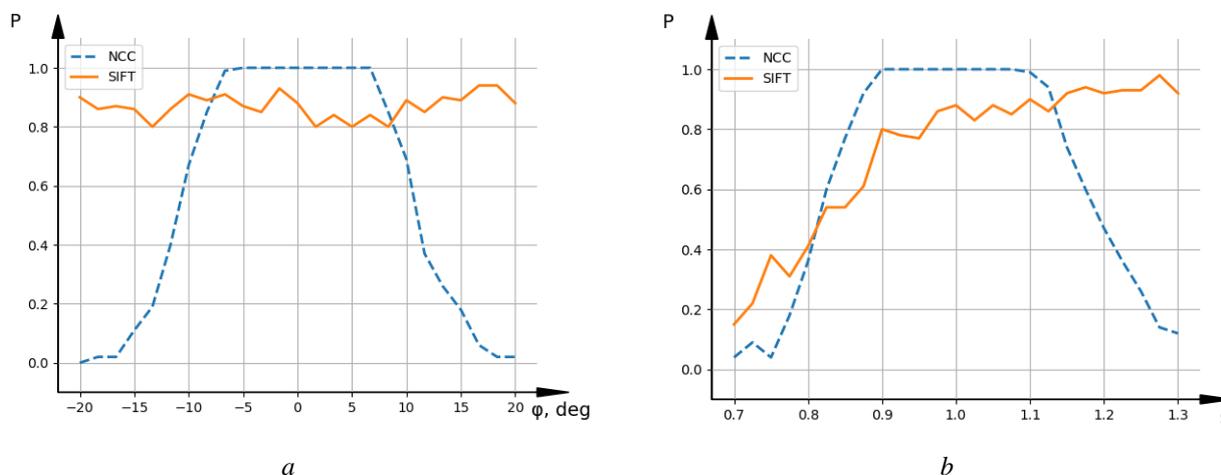


Fig. 6. Probability of correct registration P for small CI (50x50) and RI from Fig. 2, a – rotation φ ; b – scale s

2.3. Combining SIFT-based registration and NCC

In this subsection, a two-stage algorithm was investigated, when the SIFT-based algorithm is used at the first stage, and then the obtained estimates are set as initial values for the modified NCC at the second stage. The modified NCC consists in solving an optimization problem with the NCC objective function in four parameters: shifts along X and Y axes, rotation, and scale. Such an algorithm was used, for example, in the work [15]. As can be seen from Fig. 7, this two-step algorithm is rotationally resistant (it behaves similarly for the scaling) and provides high accuracy. But at the same time, it requires almost 10 times more computational costs compared to the SIFT-based algorithm (with RI descriptors precomputation).

It should also be noted that such a high accuracy (thousandths of a pixel) as in Fig. 7 became possible due to the complete coincidence of the interpolation algorithms, which were used in the CI formation model and the registration algorithm. In reality, this is not the case, which leads to noticeably higher RMSE (at least hundredths of a pixel).

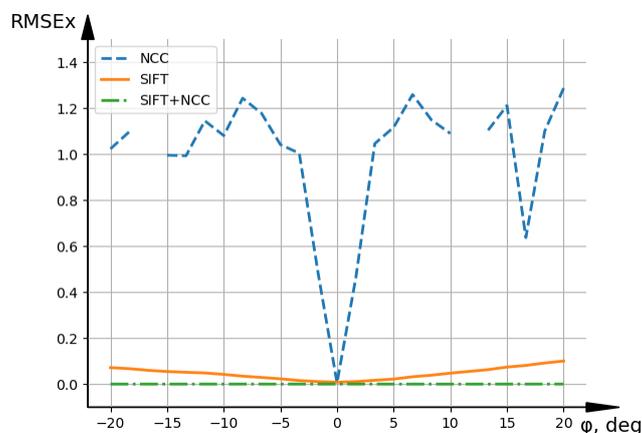


Fig. 7. Effect of rotation φ on the RMSE (on X-axis) for the image in Fig. 2, a (similar to Fig. 4, a , with a SIFT+NCC graph)

Conclusions

A comparative analysis of known registration algorithms belonging to different classes (based on pixel intensities or features) was carried out. Based on the above results, we see that SIFT-based registration is, although less accurate, but much more resistant to rotation and scale changes than the classical NCC. For the NCC, already starting from the rotation of 2 degrees (approximately, depends on the image) and a change in scale by 2 percent, the correct registration probability drops dramatically.

For a SIFT-based algorithm, it can be a problem to use small images in registration. For 50x50 fragments, the deterioration in the registration quality was already noticeable.

The two-stage algorithm considered in subsection 2.3 (SIFT+NCC) allows one to obtain an accurate algorithm that is resistant to rotation and scaling, but this will be accompanied by high computational costs (almost 10 times more than for the SIFT-based algorithm).

References:

1. Zitová B., Flusser J. Image registration methods: A survey // *Image Vis. Comput.* 2003. Vol. 21 (11). P. 977 – 1000.
2. Brown L. G. A survey of image registration techniques // *ACM Comput. Surv.* 1992. Vol. 24, No. 4. P. 325–376.
3. Uss M. L., Vozel B., Lukin V. V., Chehdi K. Efficient discrimination and localization of multimodal remote sensing images using CNN-based prediction of localization uncertainty // *Remote Sensing.* 2020. Vol. 12, No. 4., 703.
4. Conte G., Doherty P. Vision-based unmanned aerial vehicle navigation using georeferenced information // *EURASIP J. Adv. Signal Process.* 2009. Vol. 2009, Article 387308.
5. Душепа В., Усс М. Сравнительный анализ субпиксельных алгоритмов при совмещении изображений // *Радиоелектронні і комп'ютерні системи.* 2011. № 4. С. 41–51.
6. Gonzalez R. C., Woods R. E. *Digital Image Processing*, 4th ed. Pearson/Prentice Hall, 2018.
7. Lowe D. G. Distinctive image features from scale-invariant keypoints // *International Journal of Computer Vision.* 2004. Vol. 60, No. 2. P. 91–110.
8. Arandjelović R., Zisserman A. Three things everyone should know to improve object retrieval // *2012 IEEE Conference on Computer Vision and Pattern Recognition*, Providence, RI. 2012. P. 2911–2918.
9. Bradski G. *The OpenCV Library* // Dr. Dobb's J. Softw. Tools. 2000.
10. Антюфеев В. И., Быков В. Н. Сравнительный анализ алгоритмов совмещения изображений в корреляционно-экстремальных системах навигации летательных аппаратов // *Авиационно-космическая техника и технология.* 2008. № 1 (48). С. 70 – 74.
11. Karami E., Prasad S., Shehata M. Image matching using SIFT, SURF, BRIEF, and ORB: performance comparison for distorted images // *Proceedings of the 2015 Newfoundland Electrical and Computer Engineering Conference*, St. John's, Canada. 2015.
12. Virtanen P. et al. SciPy 1.0: Fundamental algorithms for scientific computing in Python // *Nature Methods.* 2020. Vol. 17(3). P. 261–272.
13. Harris C. R. et al. Array programming with NumPy // *Nature.* Vol. 585. P. 357–362.
14. SASGIS. Веб-картографія і навігація [Електронний ресурс]. Режим доступу: <http://www.sasgis.org/sasplaneta>
15. Dushpa V. A machine learning approach for image registration accuracy estimation // *2020 IEEE Ukrainian Microwave Week (UkrMW)*, Kharkiv, Ukraine. 2020. P. 368–372.

Received 05.10.2020

Відомості про авторів:

Душепа Віталій Анатолійович – ст. викладач кафедри аерокосмічних радіоелектронних систем (№ 501), Національний аерокосмічний університет ім. М. С. Жуковського «Харківський авіаційний інститут», Україна; email: v.dushpa@khai.edu, ORCID: <http://orcid.org/0000-0001-6105-3761>

Тягнирядно Єгор Андрійович – бакалавр, Національний аерокосмічний університет ім. М. С. Жуковського «Харківський авіаційний інститут», Україна; email: monkeber@gmail.com

Баришев Ігор Володимирович – д-р технічних наук, професор, професор кафедри аерокосмічних радіоелектронних систем (№ 501), Національний аерокосмічний університет ім. М. С. Жуковського «Харківський авіаційний інститут», Україна; email: biv1937@ukr.net