UDC 004.051:629.735.018.006.26:629.7.05/.06(045)

Maryna Mukhina[1]
Tetyana Yeremeyeva[2]
Aliona Kuzmenko[3]
Mykhailo Panarin[4]
Oleksandr Revchuk[5]
Olena Tkachenko[6]

ANALYSIS OF EFFICIENCY OF USE OF HARRIS AND KANADE–LUCAS–TOMASI DETECTORS FOR VISUAL NAVIGATION TASKS

[1,2,3,4,5,6]National Aviation University
1, Kosmonavta Komarova ave., Kyiv, 03680, Ukraine
E-mails: [1]m_mukhina@inbox.ru; [2]eremeeva94@mail.ua; [3]alena11_95@mail.ru;
[4]michaelpanarin@inbox.ru; [5]sasrev@yahoo.com; [6]yaelen4ik@rambler.ru

Abstract. *The article examines methods of images analysis based on computer vision. We made a comparison between the detectors of feature points determined by Harris and Kanade–Lucas–Tomasi (KLT) methods. Found points are represented by Speed-Up Robust Feature (SURF) descriptor and then used to determine homography matrix. Analyses of accuracy of visual navigation is done by estimation of a camera rotation angle via factorization of homography matrix obtained from two detector methods. Errors of visual navigation follow the normal distribution for the given sample.*

Keywords: Direct Linear Transformation (DLT) method; Harris detector; homography matrix; Kanade–Lucas–Tomasi (KLT) feature tracker; Speed-Up Robust Feature (SURF)

1. Introduction

Unmanned aerial vehicles (UAV) have gained increased interest in computer vision research. To navigate safely, this flying machine needs the ability to localize itself autonomously using its onboard sensors. One of such sensor is a camera. Its purpose is ability to estimate motion from images (including interest point detection, feature descriptors, error estimation, and iteratively closest point).

It is very useful to develop such autonomous control system, to have auto piloted aircrafts orientate in space much faster, react efficiently, by getting an information from various amount of data sources like PGS, inner sensors and data from ground center [1]. However, for developing an efficient data exchange, previous researches of efficiency needed. It is very useful to develop such autonomous control system, to have auto piloted aircrafts orientate in space much faster, react efficiently, by getting an information from various amount of data sources like GPS, inner sensors and data from ground center.

Using the transmitter of electronic jamming for GPS system, it is possible to disrupt the receiver of this navigation system. As a result, the receiver loses the ability to determinate the coordinates of the objects and navigation of UAV becomes unreliable. Accuracy of only one inertial navigation system (INS), based on gyros, for determination of UAV coordinates location is not enough, its degrades with time significantly. Therefore, it requires alternative variant for data integration and INS errors compensation. In such situation, the possible solution may be the visual navigation [2].

2. Problem statement

One of the fundamental moments of visual navigation is detection of interest points, which is further used for images comparison. It is necessary to estimate efficiency and speed of existing feature detection methods among which two are chosen: Harris and KLT. The first method – Harris affine region detector belongs to the category of feature detection. Feature detection is a preprocessing step of several algorithms that rely on identifying characteristic points or interest points to make correspondences between images, recognize textures, categorize objects or build panoramas.

$$A = \sum_{x,y} w(x,y) \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} = \begin{bmatrix} \langle I_x^2 \rangle & \langle I_x I_y \rangle \\ \langle I_x I_y \rangle & \langle I_y^2 \rangle \end{bmatrix}, \quad (1)$$

where $w(x,y)$ – weighting function;

I – image;

$I_x I_y$ – partial derivatives.

This matrix is a Harris matrix, and angle brackets denote averaging (i.e. summation over (x,y)). If a circular window (or circularly weighted window,

such as a Gaussian) is used, then the response will be isotropic.

The second method is Kanade-Lucas-Tomasi (KLT) feature tracker. It is proposed mainly for the purpose of dealing with the problem that traditional image registration techniques are generally costly. KLT makes use of spatial intensity information to direct the search for the position that yields the best match. It is faster than traditional techniques for examining far fewer potential matches between the images[3]. KLT uses functional of quality

$$F(\mathrm{h}) = \sum_{p \in P} K(p)(I(p+h) - J(p))^2 , \quad (2)$$

where P is domain of assignment of sample;

   I is landmark of feature;

   I is the next image;

   h – target shift of feature surrounding area on another shot;

   $K(p)$ – some weighting function.

Comparison of two methods is realized by using them in simple determination of angle of camera rotation algorithm from homography matrix.

Using the obtained coordinates of feature points the homography matrix is calculated. A 2D point (x,y) in an image can be represented as a 3D vector $\mathbf{x} = (x_1, x_2, x_3)$ where $\mathrm{x} = \dfrac{x_1}{x_3}$ and $y = \dfrac{x_2}{x_3}$. This is called the homogeneous representation of a point and it lies on the projective plane P^2. A homography is an invertible mapping of points and lines on the projective plane P^2. A mapping from $P^2 \to P^2$ is a projectivity if and only if there exists a non-singular 3×3 matrix H such that for any point in P^2 represented by vector x it is true that its mapped point equals $\mathbf{H}_x$ [2]. This tells us that in order to calculate the homography that maps each x_i to its corresponding x_i' it is sufficient to calculate the 3×3 homography matrix $\mathbf{H}$. Its decomposition results in two matrices: rotation and calibration.

Once camera rotation and translation have been extracted from an estimated homography matrix, this information may be used for navigation, or to insert models of 3D objects into an image or video, so that they are rendered with the correct perspective and appear to have been part of the original scene.

$$\mathbf{H} = \mathbf{K} * \mathbf{R} ,$$

where $\mathbf{K}$ – calibration matrix; $\mathbf{R}$ – rotation matrix. $\mathbf{H}$ – is considered a homogeneous matrix and has only 8 degrees of freedom even though it contains 9

elements. This means there are 8 unknowns that need to be solved for.

A rotation matrix is a matrix that is used to perform a rotation in Euclidean space. For example the matrix of the following form

$$\mathbf{R} = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} , \quad (3)$$

Rotation matrices provide a means of numerically representing an arbitrary rotation of the axes about the origin, without appealing to angular specification. These coordinate rotations are a natural way to express the orientation of a camera, or the attitude of a spacecraft, relative to a reference axes-set.

Calibration matrix consists of camera intrinsic parameters:

$$\mathbf{K} = \begin{bmatrix} \alpha_x & \gamma & u_0 \\ 0 & \alpha_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} , \quad (4)$$

The intrinsic matrix contains five intrinsic parameters. These parameters encompass focal length, image sensor format, and principal point. The parameters $\alpha_x = f + m_x$ and $\alpha_x = f + m_y$ represent focal length in terms of pixels, where m_x and m_y are the scale factors relating pixels to distance and f is the focal length in terms of distance. γ represents the skew coefficient between x and y axis, and is often 0. u_0 and v_0 represent the principal point, which would be ideally in the center of the image.

Using the rotation matrix the angle of camera rotation can be calculated. Comparing it with the true rotation angle, error can be calculated to determine the accuracy of methods.

3. Algorithm and analysis of used detectors

The first step is to calibrate the camera and getting its intrinsic parameters.

As the next step a set of pictures are taken with the same step.

These pictures are then processed by Harris detector and KLT detector.

Harris corner detector algorithm relies on a central principle: at a corner, the image intensity will change largely in multiple directions. This can alternatively be formulated by examining the changes of intensity due to shifts in a local window. Around a corner point, the image intensity will

change greatly when the window is shifted in an arbitrary direction. Following this intuition and through a clever decomposition, the Harris detector uses the second moment matrix as the basis of its corner decisions. The matrix **A** has also been called the autocorrelation matrix and has values closely related to the derivatives of image intensity.

$$\mathbf{A}(x) = \sum_{p,q} w(\mathrm{p},\mathrm{q}) \begin{bmatrix} I_x^2(x) & I_x I_y(x) \\ I_x I_y(x) & I_y^2(x) \end{bmatrix}, \qquad (5)$$

where I_x and I_y are the respective derivatives (of pixel intensity) in the x and y direction at point X and p and q are the values of the weighting function. The off-diagonal entries are the product of I_x and I_y, while the diagonal entries are squares of the respective derivatives. The weighting function $w(x,y)$ can be uniform, but is more typically isotropic, circular Gaussian,

$$w(x,y) = g(x,y,\sigma) = \frac{1}{2\pi\sigma^2} e^{\left(-\frac{x^2+y^2}{2\sigma^2}\right)}, \qquad (6)$$

where σ is a scale.

That acts to average in a local region while weighting those values near the center more heavily.

As it turns out, matrix (5)describes the shape of the autocorrelation measure as due to shifts in window location. Thus, if we let λ_1 and λ_2 be the eigenvalues of **A**, then these values will provide a quantitative description of how the autocorrelation measure changes in space. Matrix **A** centered on corner points will have two large, positive eigenvalues. Rather than extracting these eigenvalues using methods like singular value decomposition, Harris measure based on the trace and determinant is used:

$$R = \det(\mathbf{A}) - \alpha \mathrm{trace}^2(\mathbf{A}) = \lambda_1 \lambda_2 - \alpha(\lambda_1 + \lambda_2)^2, \qquad (7)$$

where α is a constant. Corner points have large, positive eigenvalues and would thus have a large Harris measure [4]. Thus, corner points are identified as local maxima of Harris measure that are above a specified threshold.

KLT feature tracker is based on Lucas and Kanade idea of a local search using gradients weighted by an approximation to the second derivative of the image.

If h is the displacement between two images $F(x)$ and $G(x) = F(x+h)$ then the approximation is made that

$$F'(x) \approx \frac{F(x+h)-F(x)}{h} = \frac{G(x)-F(x)}{h}, \qquad (8)$$

So that

$$h \approx \frac{G(x)-F(x)}{F'(x)}.$$

This approximation to the gradient of the image is only accurate if the displacement of the local area between the two images to be registered is not too large. The approximation to h depends on x. For combining the various estimates of h at various values of x, it is natural to average them:

$$h \approx \frac{\sum_x \frac{G(x)-F(x)}{F'(x)}}{\sum_x 1}, \qquad (9)$$

The average can be further improved by weighting the contribution of each term to it, which is inversely proportional to an estimate of $\left|F''(x)\right|$, where

$$F''(x) \approx \frac{G'(x)-F'(x)}{h}$$

For facilitating the expression, a weighting function is defined:

$$w(x) = \frac{1}{\left|G'(x)-F'(x)\right|}, \qquad (10).$$

The average with weighting is thereby:

$$h \approx \frac{\sum_x \frac{w(x)[G(x)-F(x)]}{F'(x)}}{\sum_x \omega(x)}, \qquad (11).$$

Upon obtaining, the estimate $F(x)$ can be moved by the estimate of h. The procedure is applied repeatedly, yielding a type of Newton-Raphson iteration. The sequence of estimates will ideally converge to the best h. The iteration can be expressed by

$$\begin{cases} h_0 = 0 \\ h_{k+1} = h_k + \dfrac{\sum_x \dfrac{w(x)[G(x)-F(x+h_k)]}{F'(x+h_k)}}{\sum_x w(x)}, \end{cases} \qquad (12)$$

Descriptor of points is determined with the help of SURF. This descriptor is based on mixing of crudely localized information and the distribution of gradient related features seems to yield good distinctive power while fending off the effects of localization errors in terms of scale or space. Using relative strengths and orientations of gradients reduces the effect of photometric changes. The first

step consists of fixing a reproducible orientation based on information from a circular region around the interest point. Then, we construct a square region aligned to the selected orientation, and extract the SURF descriptor from it [5, p. 6].

According to the detected points, the homography matrix is constructed. After decomposition of this matrix, the rotation angle of the image compares with the true one.

4. Experimental results

The study of proposed algorithms has been done on the series of images of camera from the same position with rotation around optical axis by step in 2° (Fig. 1, 2).
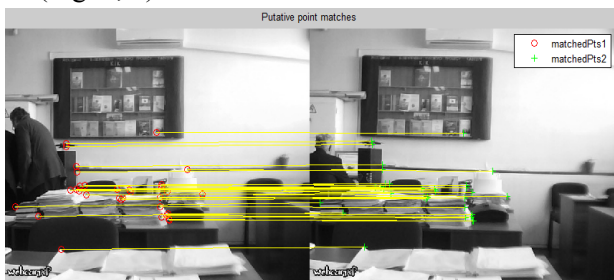
Fig. 1. Referenced image (left) and the same scene with camera rotation angle in 2°. Matched points are shown by line connections. Harris detector.

Fig. 2. Referenced image (left) and the same scene with camera rotation angle in 2°. Matched points are shown by line connections. KLT tracker.

Matching has been done between the reference image assumed to be taken with initial heading and current one, as well as comparing photos in series. Tests have been done in MATLAB R2014A.

Mean values of $\Delta\psi$:
mean1 = -0.2040(Harris detector,neighboring images);
mean2 = -0.0824(Harris detector,initial and current);
mean3 = -0.6908(KLTtracker,neighboring images);
mean4 = 0.1796(KLTtracker,initial and current);

Standard deviation of $\Delta\psi$:

std1 = 0.9123
std2 = 0.9350
std3 = 2.2236
std4 = 0.4927

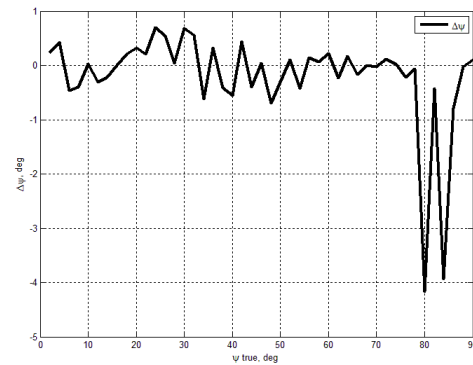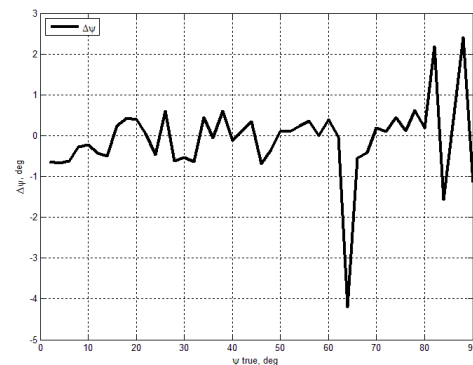**Fig. 3.** Harris detector. Error of rotation angle determination between neighboring images.

Fig. 4. Harris detector. Error of rotation angle determination between initial and current images.
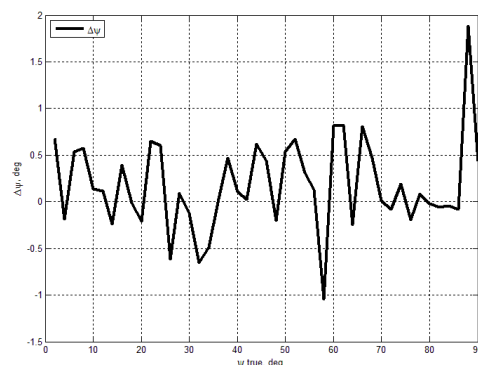
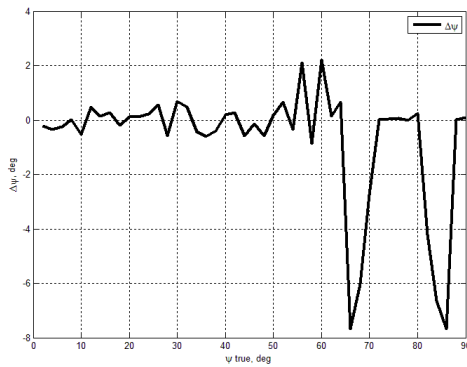**Fig. 5.** KLT tracker. Error of rotation angle determination between initial and current images.

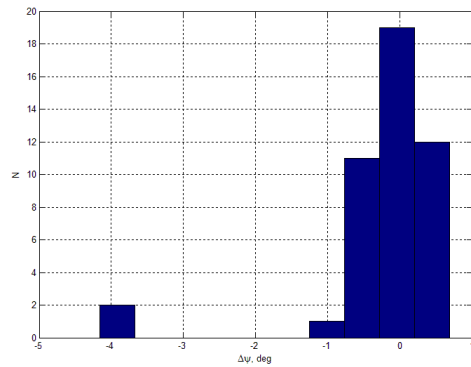**Fig. 6.** KLT tracker. Error of rotation angle determination between neighboring images

Fig. 9. Harris detector. Histogram of error of rotation angle determination between neighboring images.
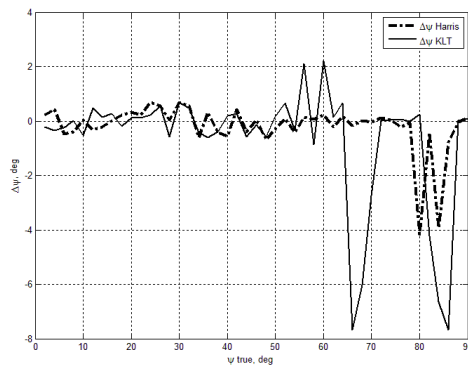

Fig. 7. Comparison of errors of KLT tracker and Harris detector in rotation angle determination of neighboring images.
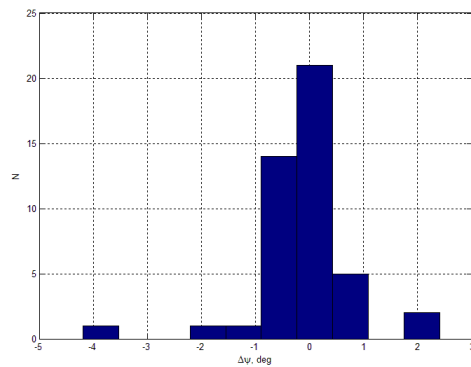

Fig. 10. Harris detector. Histogram of error of rotation angle determination between initial and current images.
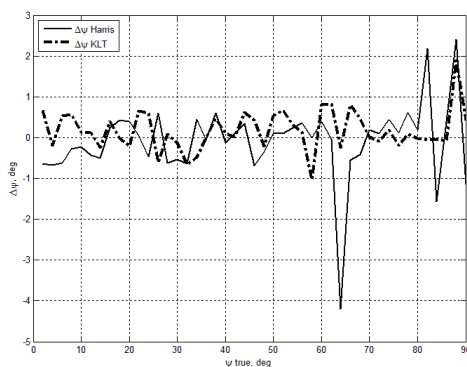

Fig. 8. Comparison of errors of KLT tracker and Harris detector in rotation angle determination of initial and current images.
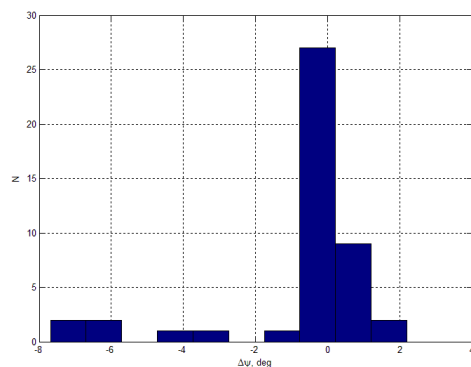

Fig. 11. KLT tracker. Histogram of error of rotation angle determination between initial and current images.
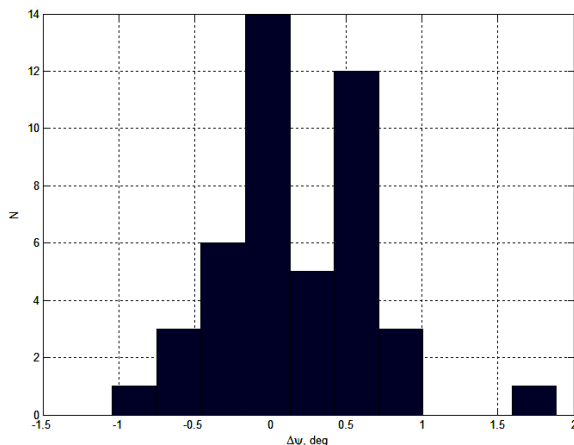
Fig. 12. KLT tracker. Histogram of error of rotation angle determination between neighboring images.

Comparison of both methods in speed of information processing was performed with MATLAB **tic toc** functions. As it varies with the used computer, both methods were tested on the same machine. KLT method spent approximately 1,17 seconds per cycle (2 images processed), while Harris method spend around 0.98 seconds. In total, KLT method used 1 minute 45,3 seconds to process 90 images, when Harris method used 1 minute 28,2 seconds. This results clearly show that Harris is a faster method. Results of accuracy comparison clearly show that Harris method is also more precise for correlation of neighboring video frames (standard deviation is about 2.5 times smaller). Nevertheless, if angular difference between compared frames is high, then better accuracy is shown by KLT tracker.

5. Conclusions

As the result of that process, for each frame recorded, aircraft receives the pair of images: expected image in a frame and the actual one; it is also necessary to calculate such performance data as velocity and the direction of flight at that point.

As the result, Harris method is very suitable to find a borders and corners in images. It is very useful, when it is necessary to track an object at long ranges. In other way, KLT tracker is based on finding active intensity points at the image, which can be spotted easily at the close range.

At larger angles KLT tracker is more accurate, at smaller - Harris method. In general, Harris method is much more usable in such circumstances to determine the images similarity.

References

[1] *Kusainov A. A.* Development of control system unmanned aerial vehicle along a predetermined path. – Novosibirsk, 2013. – 60 p.

[2] *Hartley R., Zisserman A.* Multiple view geometry in computer vision. – Australian National University, Canberra, Australia. – Cambridge University Press 2000, 2003. – 673 p.

[3] *Svoboda T.* Kanade–Lucas–Tomasi Tracking (KLT tracker). – Czech Technical University, Prague. – 2008.–29 p.

[4] *Faugeras O.* Three-Dimensional Computer Vision: A Geometric Approach, MIT Press, 1996, pp. 33- 68

[5] *Herbert Bay, Tinne Tuytelaars*, Luc Van Gool. SURF: Speeded Up Robust Features. – Katholieke Universiteit Leuven. – 2007. –13 p.

[6] *Conte G., Doherty P.* A visual navigation system for UAS based on geo-referenced imagery. – Linkoping University, SWEDEN, 2011. – 6 p.

[7] *Karpov D.P.* Cross-linking of images obtained as a result of aerial photography. – St. Petersburg National Research University, Information Technologies, Mechanics and Optics Faculty of Information Technology and Software Engineering. Department of Computer Technology. – 2010. – 25 p.

[8] *Zuliani M.* RANSAC for Dummies (With examples using the RANSAC toolbox for Matlab). Los Gatos, CA – 2012. – 101 p.

[9] *Tsai R.Y.* A Versatile Camera Calibration Technique for 3D Machine Vision, IEEE J. Robotics & Automation, RA-3, No. 4, August 1987, 334 p.

[10] *Szeliski R.* Computer vision: algorithms and applications. – 2010. – 979 p.

[11] *Zolotukhin D.A.* Using MATLAB for prototyping stage of the algorithm three-dimensional reconstruction of micro-objects. - National Research Nuclear University, Moscow Engineering Physics Institute, Moscow. – 2012. – 3 p.

М.П. Мухіна[1], Т.А. Єремеєва[2], А.О. Кузьменко[3], М.І. Панарін[4], О.О. Ревчук[5], О.Ю. Ткаченко[6]
Аналіз ефективності використання детекторів Харріса та Канаде-Лукас-Томасі для задач візуальної навігації
[1,2,3,4,5,6][2]Національний авіаційний університет, просп. Космонавта Комарова, 1, Київ, Україна, 03680
E-mails: [1]m mukhina@inbox.ru; [2]eremeeva94@mail.ua; [3]alena11_95@mail.ru; [4]michaelpanarin@inbox.ru; [5]sasrev@yahoo.com; [6]yaelen4ik@rambler.ru

Досліджено методи аналізу зображень на основі комп'ютерного зору. Проведено оцінювання ефективності детекторів характерних точок, визначених методами Харріса і Канаде-Лукас-Томасі (КЛТ). Виявлені точки представлені дескриптором прискореного виділення робастних характерних ознак, які використовуються у подальшому для визначення матриці гомографії. Аналіз точності візуальної навігації проведений шляхом оцінювання кута повороту камери за допомогою факторизації матриці гомографії, отриманої з двох детекторів. Похибки візуальної навігації відповідають нормальному розподілу для цієї виборки.
Ключові слова: детектор Харріса; матриця гомографії; метод виявлення характерних ознак Канаде-Лукас-Томасі; метод прямого лінійного перетворення; прискорене виділення робасних характерних ознак

М.П. Мухина[1], Т.А. Еремеева[2], А.О. Кузьменко[3], М.И. Панарин[4], О.О. Ревчук[5], О.Ю. Ткаченко[6]
Анализ эффективности использования детекторов Харриса та Канаде-Лукас-Томаси для задач визуальной навигации
[1,2,3,4,5,6][2]Национальный авиационный университет, просп. Космонавта Комарова, 1, Киев, Украина, 03680
E-mails: [1]m mukhina@inbox.ru; [2]eremeeva94@mail.ua; [3]alena11_95@mail.ru; [4]michaelpanarin@inbox.ru; [5]sasrev@yahoo.com; [6]yaelen4ik@rambler.ru

Исследованы методы анализа изображений на основе компьютерного зрения. Проведена оценка эффективности детекторов характерных точек, определенных методами Харриса и Канаде-Лукас-Томасе (КЛТ). Обнаруженные точки представлены дескриптором ускоренного выделения робастные характерных признаков, которые используются в дальнейшем для определения матрицы гомографии. Анализ точности визуальной навигации проведен путем оценки угла поворота камеры с помощью факторизации матрицы гомографии, полученной из двух детекторов. Погрешности визуальной навигации соответствуют нормальному распределению для данной выборки.
Ключевые слова: детектор Харриса; матрица гомографии; метод выявления характерных признаков Канаде-Лукас-Томасе; ускоренное выделение робастные характерных признаков; метод прямого линейного преобразования

Mukhina Maryna (1979). PhD, Associate Professor.
Post Doctoral, National Aviation University, Kyiv, Ukraine.
Associate Professor at the Department of Aviation Computer-Integrated Complexes, National Aviation University, Kyiv, Ukraine
Education: BSc, MSc (1996-2002) in Electrical Engineering at Aviation Equipment Faculty of National Aviation University.
Research area: correlation-extreme navigation systems, data fusion algorithms
Publications: 35.
E-mail: m_mukhina@inbox.ru

Yeremeyeva Tetyana (1994). Student.
National Aviation University, Kyiv, Ukraine
Education: BSc in Automation and Computer-Integrated Manufactures at National aviation university
Research area: correlation-extreme navigation systems, data fusion algorithms
E-mail: eremeeva94@mail.ua

Kuzmenko Aliona (1994). Student.
National Aviation University, Kyiv, Ukraine
Research area: correlation-extreme navigation systems, data fusion algorithms
E-mail: alena11_95@mail.ru

Panarin Mykhailo (1994). Student .
National Aviation University, Kyiv, Ukraine
Research area: correlation-extreme navigation systems, data fusion algorithms
E-mail: michaelpanarin@inbox.ru

Revchuk Oleksandr (1994). Student.
National Aviation University, Kyiv, Ukraine
Research area: correlation-extreme navigation systems, data fusion algorithms
E-mail: sasrev@yahoo.com

Tkachenko Olena (1994). Student.
National Aviation University, Kyiv, Ukraine
Research area: correlation-extreme navigation systems, data fusion algorithms
E-mail: yaelen4ik@rambler.ru