

CORRELATION-EXTREME VISUAL NAVIGATION OF UNMANNED AIRCRAFT SYSTEMS BASED ON SPEED-UP ROBUST FEATURES

Volodymyr Kharchenko, Maryna Mukhina

National Aviation University, 1 Kosmonavta Komarova Ave., 03680 Kiev, Ukraine
E-mail: ans@nau.edu.ua (corresponding author); eduicao@nau.edu.ua

Received 31 October 2013; accepted 15 May 2014



Volodymyr KHARCHENKO, Prof. Dr

Date of birth: 1943.

Education: Kiev Civil Aviation Engineers Institute with a Degree in Radio Engineering (1962–1967).

Affiliations and functions: Engineer at United Aircraft troops (1967–1969), Junior and Senior Researcher (1969–1984), Assistant Professor (1984–1987), DrSci. (Eng.) (1994), Professor (1987–1994) at Kiev International University of Civil Aviation. From 2001 he is the Vice-Rector for scientific-research work at the National Aviation University.

Research interests: communication, navigation, surveillance.

Publications: author of about 400 scientific papers, monographs, textbooks, training aids, articles and proceedings.

Other: member of the Ukrainian Transport Academy (1996), Russian Academy of Navigation and Control (1999), Institute of Engineers of Electrics and Electronics in USA (2006).



Maryna MUKHINA, BSc, MSc, PhD Eng.

Date of birth: 1979.

Education: BSc, MSc (1996–2002) in Electrical Engineering at the Aviation Equipment Faculty of the National Aviation University.

Affiliations and functions: Assistant, (2002–2005), Candidate of Sci (Eng) (2005) at the National Aviation University.

Research interests: correlation-extreme navigation systems, data fusion algorithms

Publications: author of 26 scientific papers, textbooks, articles and proceedings.

Present position: since 2005 Assistant Professor at the Department of Aviation Computer-Integrated Complexes of the National Aviation University.

Abstract. The peculiarities of correlation-extreme visual navigation are considered. Descriptors with 64 elements of feature points of surface images are selected on the basis of the speed-up robust feature method. An analysis of possible criteria correlation functions is carried out to find the best match between the template descriptors and current images. The use of normalized correlation function is proposed based on the matrix multiplication properties of descriptors. It allows minimizing the number of false matches in comparison with the Euclidean distance in the descriptor space. The proposed matching strategy sufficiently decreases the computation time.

Keywords: correlation-extreme navigation, speed-up robust feature method, criteria functions, normalized correlation function.

1. Introduction

Miniature unmanned aircraft systems (UAS) are increasingly used for various purposes like monitoring of ground or water surface, surveillance, reconnaissance and others. The functioning of UAS in urban areas is significantly limited by severe requirements for the accuracy of navigation and control and collision avoidance solutions. The basic navigation complex of UAS includes the integrated inertial navigation system (INS) and satellite navigation system (SNS). In the case of satellite signal loss the navigation solution is obtained from the inertial measurement unit (IMU), which, however, has its own errors that grow over time. For a small UAS with a low-cost IMU based on economic micro-electromechanical sensors (MEMS), the drift of errors in a few seconds can lead to unreliable positioning. In addition, the operation of SNS in an urban area is also limited by significant multipath errors due to multiple signal reflections from obstacles, buildings and so on. The major drawback is that the satellite signal can be easily suppressed, intentionally or accidentally by the interference of mobile frequencies, disturbance of television broadcasting, and others.

One of the alternative variants that could aid INS is the correlation-extreme navigation system (CENS) that uses a priori cartographic information of the geophysical field in the given area and finds the matches with current field realization (Baklitskiy 2009). The geophysical fields include the relief field, magnetic, gravitational fields, optical or radio contrast fields, etc. Over the last decade promising results of visual UAS navigation (Conte, Doherty 2011) have been achieved by the rapid development of computing power that makes available the use of heavy computation of image processing methods in real-time mode.

Here the modified speed-up robust feature (SURF) method (Bay *et al.* 2006) will be considered for the practical implementation aboard the UAS.

2. Problem statement

Given the current realization of a surface geophysical field (optical for this particular paper) $f(x, y, \lambda, t)$, the best (extreme) correlation with a set of a priori cartographic realizations $F(x, y, \Lambda, t)$ of the same field must be found. The surface field model describes the place of signal f in a rectangular coordinate system (Fig. 1) for the observer plane xy by the observation vector Λ . The current realization of observation vector λ determines its displacement along the x - and y -axis.

$$F(x, y, \Lambda, t) = f(x - \lambda_x, y - \lambda_y, t); \Lambda = \left\| \lambda_x, \lambda_y \right\|^T. \quad (1)$$

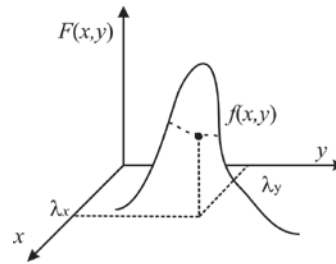


Fig. 1. Function of surface field and its realization

If the point features (anomalies of visual field) are observed, function (1) can be represented as follows:

$$F(x, y, \Lambda, t) = \sum_{i=1}^N D_i \delta(x - \lambda_{xi}, y - \lambda_{yi}, t), \quad (2)$$

where N indicates the number of observed features; D_i – the descriptors of features; $\delta(\bullet)$ – the Delta function.

In general, the difference between an a priori field and its current realization is described by a non-linear dependence $\lambda(x, y)$.

Since the realization of function (2) is random, such statistical characteristics as mean value, variance and correlation function of the following form will be used:

$$K(\lambda(x), \lambda(y), t) = M \langle f(\lambda(x), \lambda(y), t), F(x, y, t) \rangle. \quad (3)$$

The descriptors of feature points (Lowe 2004) are represented as follows:

$$D_i = \{p_1, p_2, \dots, p_k\}, \quad (4)$$

where p_1, p_2, \dots, p_k indicate properties of the feature point, which may include coordinates, average intensity, gradient orientation, Hessian value, Laplacian value, etc.

The aim is to find the optimal correlation function (3) for use in the speed-up robust feature method with the possibility to work in real-time mode.

3. Variants of criteria correlation functions

According to (Baklitskiy 2009), the following classification of correlation functions is used in CENS:

- 1) a classic correlation function based on statistical moments – the cross-correlation function:

$$K_1(\lambda) = M \{ [F(x, y, t) - M[F(x, y, t)]] [f(\lambda(x), \lambda(y), t) - M[f(\lambda(x), \lambda(y), t)]] \},$$

where $F(x, y, t)$ is the template image (hereinafter – F), and $f(\lambda(x), \lambda(y), t)$ is the current realization of the field (hereinafter – f);

and the normalized correlation function:

$$K_2(\lambda) = \frac{K_1(\lambda)}{\sigma[F]\sigma[f]}, \quad (5)$$

where σ is the operator of the root mean square error.

The normalized cross-correlation function is highly recommended for template matching (Briechele, Hanebeck 2001);

- 2) difference criteria functions, represented as follows:

$$K_3(\lambda) = M[|F - f|^p], \quad (6)$$

where p takes values 1, 2 and so on.

There are several variants of difference criteria functions: the function of mean square of difference – $K_4(\lambda) = M[F - f]^2$; the function of mean absolute value of difference – $K_5(\lambda) = M[|F - f|]$, etc.

The advantage of difference functions is the absence of multiplication operations that decreases the computational costs sufficiently. However, if similar images match only partly, the difference functions have the high values;

- 3) spectral criteria functions are used to match the spectra of images. The following functions can be related to them:

- the correlation spectral function:

$$F^{-1} \left\{ g_F(u) g_f^*(u) \right\} = F^{-1} \{ G(u) \},$$

where F^{-1} indicates the inverse Fourier transform; $g_F(u), g_f(u)$ – spectra of a priori and current realization of images; $G(u)$ – the cross-spectrum; * – complex conjugation;

- Rott's function:

$$K_6(\lambda) = F^{-1} \left\{ \frac{G(u)}{g_f(u)} \right\};$$

- the coherence function:

$$K_7^2(\lambda) = F^{-1} \left\{ \frac{G^2(u)}{\sqrt{g_F(u) g_f(u)}} \right\}.$$

The use of spectral correlation functions is advisable when it is necessary to increase the components of spatial frequencies of a valid signal to a maximum and to suppress the components distorted by noise;

- 4) the even criteria functions are used for quantized images with quantization levels of two or more. If the element of the first image has a quantization level i , and the element of the second image has a level j , then the even function $f_{ij}(\lambda), 0 \leq i, j \leq 2^n - 1$ is increased by one. Here 2^n is the number of quantization levels. Correspondingly, function $f_{ij}(\lambda)$ with $i = j$ equals the number of image elements the intensities of which are the same, and function $f_{ij}(\lambda)$ with $i \neq j$ equals the number of elements with different intensities. If the two images with size $N \times N$ are identical, then

$$K_8 = \sum_{j=0}^{2^n-1} \sum_{i=0}^{2^n-1} F_{ij}(\lambda) = \begin{cases} N^2, & i = j; \\ 0, & i \neq j. \end{cases}$$

- 5) rank criterial functions assume the ordering of image elements according to intensity and indexing them by ranks. A good example of such a function is the Spearman function:

$$K_9(\lambda) = 1 - \frac{6 \sum_{i=1}^N (R_{F_i} - R_{f_i})}{N(N^2 - 1)},$$

where R_{F_i}, R_{f_i} indicate the ranks of i -th elements in template and current images.

The proposed solution. In a general case the descriptor of feature point (4) includes the following information: coordinates $P = \{x, y\}$, the scale of Gaussian filter $M = \{\sigma\}$, gradient orientation $R = \{\varphi\}$, Laplacian $L = \{0, 1\}$ (either a white spot on a black background or a black spot on a white), and the gradients of quadrants $D = \{D_1, D_2, \dots, D_{64(128)}\}$ which surround the point.

To calculate the descriptor a rectangular area is formed around the feature point with a size of 20σ , where σ indicates the filter scale that was used to find the point. For the first octave the size of the area is 40×40 pixels. The quadrant is oriented along the major direction calculated for the feature point.

The descriptor is calculated as the gradients for $4 \times 4 = 16$ quadrants around the feature point. Then each quadrant is divided further into 16 smaller quadrants, as shown in figure 2.

For each quadrant the responses of Haar wavelets (Porwik, Lisowska 2004) of a size 2σ are computed on the regular grid of $5 \times 5 = 25$. The responses for directions x and y are designated as dx and dy , respectively, and then for each quadrant the following vector is found:

$$D_{quadrant} = \left[\sum dx, \sum dy, \sum |dx|, \sum |dy| \right].$$

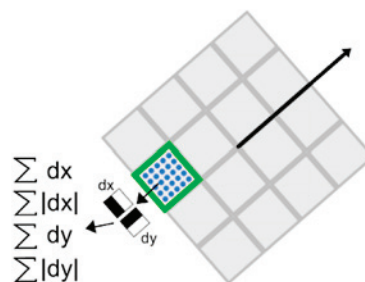


Fig. 2. Descriptor of feature point (Bay et al. 2006)

With Haar wavelet calculation the image is not rotated and the filter is computed into image coordinates. Afterwards, the gradient coordinates (dx, dy) are rotated by an angle corresponding to the orientation of the quadrant. Four components on each quadrant must be computed, which results in 64 descriptor components of the area around the feature point. By the formation of the descriptor array the values are weighted by a Gaussian 3.3σ and are centered in the feature point to minimize the possible noise components.

It is necessary to select the best criteria function to match the pairs of descriptors D_E, D_f of the template and current images. In (Elesina *et al.* 2010), the normalized correlation function (5) and the function of mean square difference were proposed for the same purposes in the following expression:

$$K_{ij} = \frac{D_i(64) \cdot D_j(64)}{\sqrt{D_i^2(64) \cdot D_j^2(64)}} \quad (7)$$

and

$$K_{ij} = [D_i(64) - D_j(64)]^2, \quad i = \overline{1, N}, \quad j = \overline{1, n}, \quad (8)$$

where indexes i, j correspond to the feature points on the template and current images, and respectively N, n – to numbers of points found on these images.

Function (8) is actually used in most realizations of the SURF method, in particular, in the “Code of SURF listing in MATLAB”. But a significant disadvantage of such strategy is the greater number of calculations that may become crucial for real-time application aboard a UAS. The use of function (7) is more preferable due to its natural properties of the matrix multiplication and the possibility to use advanced CUDA technologies of parallel processing as in (Sheta *et al.* 2012).

The realization of SURF method in the “Code of SURF listing in MATLAB” was used for practical experiments. The descriptors are formulated as matrix D with the size of $64 \times N$, or $128 \times N$, where N indicates the number of feature points. Function (7) can be found by the *single* multiplication of two matrixes of template and current image descriptors:

$$K = D_{template}^T \cdot D_{current} = \begin{bmatrix} D_1(1) & \dots & D_1(N) \\ \vdots & \ddots & \vdots \\ D_{64}(1) & \dots & D_{64}(N) \end{bmatrix}^T \cdot \begin{bmatrix} D'_1(1) & \dots & D'_1(n) \\ \vdots & \ddots & \vdots \\ D'_{64}(1) & \dots & D'_{64}(n) \end{bmatrix} \quad (9)$$

Since the descriptor matrix is already normalized due to peculiarities of the SURF method, it is possible to state that each component K_{ij} of matrix (9) corresponds to (7). Obviously, the maximum values in matrix K provide the best matches between feature points. The search of maximum elements is carried out by finding the maximum values of matrix (9) in each row and checking whether this value is greater than the threshold. Only one maximum element in each row is selected since it is supposed that one point on the template will be matched to a single point on the current image.

Experimental results and comparison. The realization of the SURF method is carried out based on the available code listing in MATLAB by (Code of SURF listing in MATLAB). A set of testing images as video frames was selected from a UAS flight taken in Borodyanka at an altitude of 185 m. The threshold value of the normalized cross-correlation was selected as 0.95 to satisfy the standard of confidence probability of reliable matching (Baumberg 2000). Actually, the value of the correlation threshold must be selected carefully depending on the image structure. For highly-detailed images this coefficient may be decreased since the probability of having the same descriptor for a different point will be lower than for images with uniform or repeated patterns. The results and comparison with the method used in (Code of SURF listing in MATLAB) are represented in figures 3–6.



Fig. 3. The matching of two video frames from aboard a UAS with a difference of 10 frames (0.4 sec, 25 fps) with the correlation threshold $K = 0.95$. The number of false matching is 73 from a total of 189 matching features (on very similar areas of road). The average time for computing the matches is 0.0067 sec

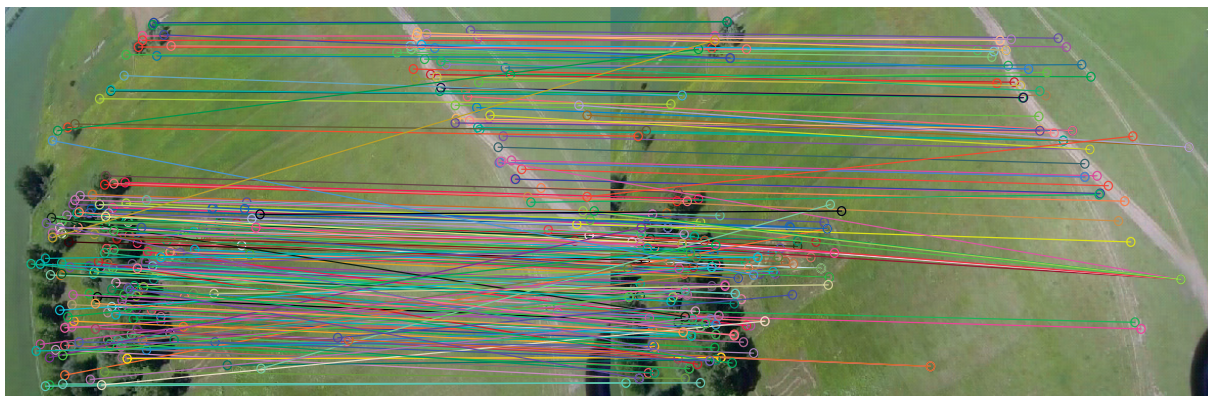


Fig. 4. The matching of two video frames from aboard a UAS with a difference of 10 frames (0.4 sec, 25 fps) done in (Code... 2013). The number of false matching is 126 from a total of 189 matching features. The average time for computing the matches is 0.0499 sec



Fig. 5. The matching of two video frames from aboard a UAS with a difference of 40 frames (1.6 sec, 25 fps) with the correlation threshold $K = 0.95$. The number of false matching is 4 from a total of 31 matching features. The average time for computing the matches is 0.0039 sec

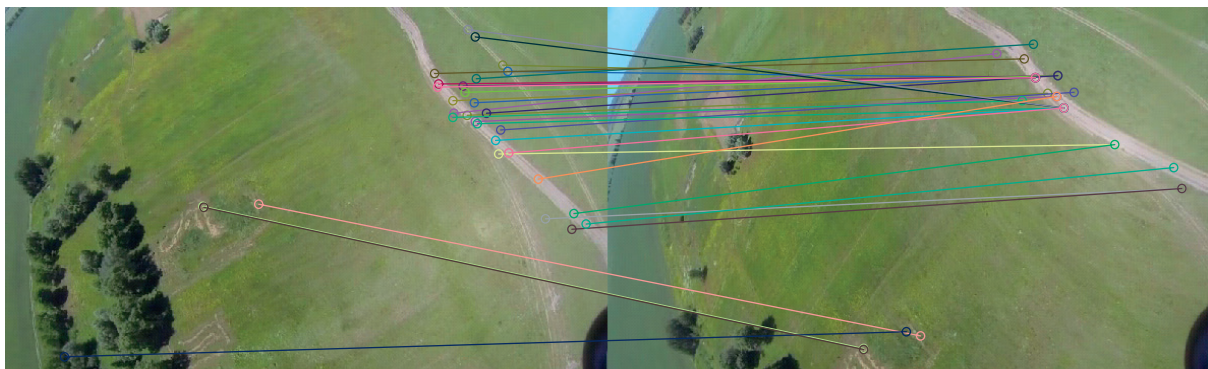


Fig. 6. The matching of two video frames from aboard a UAS with a difference of 40 frames (1.6 sec, 25 fps) done in (Code of SURF listing in MATLAB). The number of false matching is 9 from a total of 31 matching features. The average time for computing the matches is 0.0362 sec

4. Conclusions

The proposed variant of the normalized correlation function provides the possibility to correct the number of matched features depending on the degree of image change in order to minimize the false matching. The threshold value of the normalized cross-correlation was selected as 0.95 to satisfy the standard of confi-

ence probability of reliable matching. Variation of the threshold value K of the correlation function results in the minimization of false matching in comparison with “Code of SURF listing in MATLAB” from 1.7 to 2.25 times. The computational time of matching varies in a range 7.45–9.3 times smaller than for the correlation function (8).

References

- Baklitskiy, V. K. 2009. *Correlation-Extreme Methods of Navigation and Guidance*. Tver: TO “Knizhnyi klub”. 360 p.
- Baumberg, A. 2000. Reliable feature matching across widely separated views, in *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, 13–15 June, 2000, Hilton Head Island, 1: 774 – 781.
- Bay, H.; Tuytelaars, T.; Van Gool, L. 2006. SURF: speeded up robust features, in *European Conference on Computer Vision*, 7–13 May, 2006, Graz, Austria, 1: 404–417.
- Briechle, K.; Hanebeck, D. 2001. Template matching using fast normalized cross correlation, in *Proc. SPIE 4387, Optical Pattern Recognition*, 20 March, 2001, vol. 95.
<http://dx.doi.org/10.1117/12.421129>
- Code of SURF listing in MATLAB. 2013 [online], [cited 15 August 2013]. Available from Internet: <http://www.mathworks.com/matlabcentral/fileexchange/28300-opensurf-including-image-warp>
- Conte, G.; Doherty, P. 2011. A visual navigation system for UAS based on geo-referenced imagery, in *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences: Proceedings of Conference on Unmanned Aerial Vehicle in Geomatics*, 14–16 September, 2011, Zurich, Switzerland, 38–1/C22: 101–106.
- Elesina, S. I.; Kostrov, B. V.; Loginov, A. A. 2010. Search methods in correlation-extreme navigation systems, in Pylkina, A. N. (Ed.). *Proc. of Programnye informatsionnye sistemy*. Ryasan: RGRTU, 85–90.
- Lowe, D. 2004. Distinctive image features from scale-invariant keypoints, *International Journal of Computer Vision* 60(2): 91–110.
<http://dx.doi.org/10.1023/B:VISI.0000029664.99615.94>
- Porwik, P.; Lisowska, A. 2004. The haar-wavelet transform in digital image processing: its status and achievements, *Machine Graphics & Vision* 13(1–2): 79–98.
- Sheta, B.; Elhabiby, M.; El-Sheimy, N. 2012. Assessments of different speeded up robust features (SURF) algorithm resolution for pose estimation of UAV, *International Journal of Computer Science and Engineering Survey* 3: 15–41.
<http://dx.doi.org/10.5121/ijcses.2012.3502>