UDC 528.715:004.932(045)

Maryna Mukhina

COMPARISON OF ERROR METRICS IN MATCHING ALGORITHMS OF IMAGES BY SURF DETECTOR

National Aviation University, Kyiv, Ukraine
1, Kosmonavta Komarova avenue, Kyiv, 03680, Ukraine
E-mail: m_mukhina@inbox.ru

Abstract. *Speed-Up Robust Feature (SURF) method is used to detect feature points of images. The analyses of matching algorithms of feature points is done. The comparison of different error metrics by their accuracy and computing efficiency is provided on the series of test images for basic transforms like scaling, rotation and shifting.*

Keywords: error metric; feature point; homography matrix; normalized cross-correlation; Speed-Up Robust Feature.

Introduction

Matching of images is used in variety of practical applications like photomapping by unmanned aerial vehicles (UAV), visual correlation-extreme navigation, target detection, etc. Most of these application require robust real-time algorithms to detect the feature points and then to provide their reliable matching. In some situations, e.g., for video sequences or for stereo pairs that have been rectified the local motion around each feature point may be mostly translational. In this case, simple error metrics, such as the sum of squared differences or normalized cross-correlation can be used to directly compare the intensities in small patches around each feature point. Because feature points may not be exactly located, a more accurate matching degree can be computed by performing incremental motion refinement but this can be time consuming and can sometimes even decrease performance.

Among method of feature detection the scale invariant feature transform (SIFT) [3] is at present a very popular basis for image stitching. SIFT delivers point-wise correspondences between distinctive, non-repetitive local features in the two images. The number of detected features is significantly smaller than the number of pixels in the image. Other methods for identifying features include local image descriptors like intensity patterns [5] and the Kanade-Lucas-Tomasi Feature Tracker (KLT) [6].

But for real-time application Speed-Up Robust Feature (SURF) method is used [1], since it approximates or even outperforms previously proposed schemes with respect to repeatability, distinctiveness, and robustness, yet can be computed and compared much faster.

2. Problem statement

Much of the performance increase in SURF can be attributed to the use of an intermediate image representation known as Integral Image. The integral image is computed rapidly from an input image and is used to speed up the calculation of any upright rectangular area. Given an input image I and a point (x;y) the integral image is calculated by the sum of the values between the point and the origin. Formally this can be defined by the formula:

$$I_{\Sigma}(x,y)=\sum_{i=0}^{i\le x}\sum_{j=0}^{j\le y} I(i,j) \tag{1}$$

Using the integral image, the task of calculating the area of an upright rectangular region is reduced to four operations.

Since computation time is invariant to change in size this approach is particularly useful when large areas are required. SURF makes good use of this property to perform fast convolutions of varying size box filters at near constant time.

The SURF detector is based on the determinant of the Hessian matrix:

$$\mathbf{H}(x,y,\sigma)=\begin{vmatrix} L_{xx}(x,y,\sigma) & L_{yx}(x,y,\sigma) \\ L_{xy}(x,y,\sigma) & L_{yy}(x,y,\sigma) \end{vmatrix}, \tag{2}$$

where $L_{xx}(x,y,\sigma)$ – convolution of second order partial of Gaussian $\frac{\partial^2}{\partial x^2}g(\sigma)$ from the image in the point (x, y). The same is for $L_{xy}(x,y,\sigma)$ and $L_{yy}(x,y,\sigma)$.

But for SURF the fast Hessian is found that is the approximation of matrix (2) by box filters. Dimension of filters is selected as 9×9 with scale $\sigma = 1.2$ (minimal). The approximations are designated as D_{xx}, D_{yy}, D_{xy}. The weights are selected from Frobenius norm:

$$\det\left(\mathbf{H}_{approx}\right)=D_{xx}D_{yy}-\left(0.9D_{xy}\right)^2.$$

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

To calculate the descriptor the rectangular area is formed around the feature point. It has the size 20σ, where σ – filter scale, that was used to find the point. For the first octave the size of area is 40x40 pixels. The quadrant is oriented along the major direction calculated for feature point.

The descriptor is calculated as the gradients for 4×4=16 quadrants around the feature point. Then each quadrant is divided further by 16 smaller quadrants as it is shown in Fig. 1.

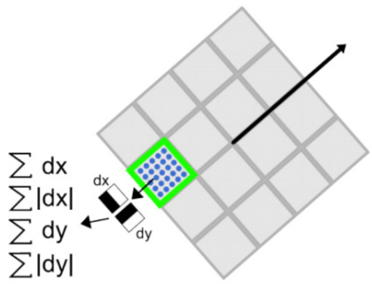

Fig. 1. Descriptor of feature point [1]

For each quadrant the responses of Haar wavelets of size 2σ are computed on the regular grid 5×5=25. Responses by directions *x* and *y* are designated as *dx* and *dy*, respectively, and then for each quadrant the following vector is found:

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

With Haar wavelets calculation the image is not rotated, the filter is computed in image coordinates. But after the gradients coordinates (*dx*, *dy*) are rotated in angle corresponding to orientation of quadrant.

Four components on each quadrant must be computed that gives totally the 64 components of descriptor of area around the feature point. By the forming of descriptor array the values are weighted by Gaussian 3.3σ and centered in the feature point to minimize the possible noise components.

After detecting all feature points on the pair of compared images it is necessary to find matches between these points.

3. Related works

Matching of feature points found by any descriptors is performed usually by well known method as random sampling (RANSAC) researched e.g. for SURF algorithm in [2]. Transformation between the pair of images can be described by homograhy matrix 3×3:

$$\mathbf{H}=\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}, \tag{3}$$

which describes the transformation of image point (*x, y*) into the point (*x*', *y*') on the other image by the following relationships:

$$\begin{aligned} x' &= \frac{a_{11}x+a_{12}y+a_{13}}{a_{31}x+a_{32}y+1}, \\ y' &= \frac{a_{21}x+a_{22}y+a_{23}}{a_{31}x+a_{32}y+1}. \end{aligned} \tag{4}$$

The calculation of image renovation needs to work out at least eight parameters, and at least eight equations in theory are needed, so it needs at least 4 non-collinear feature points. False matching points will influence the result of the least squares estimates.

RANSAC algorithm steps:

1) From the equation of matching feature points are selected, randomly selected from 4 points to set up equations, solve the eight unknown parameters of matrix **H**.

2) Calculate rest of the feature points after matrix **H** transformation, and calculate the distance between candidate matching points.

3) If the distance is less than a certain value, the candidate point is looked as interior point or the outside point.

4) Make the statistics of the quantity of interior point under the homography matrix.

5) Choose another four match points, carry out steps 1 to 4 again, and repeat several times, choose the collection with largest number of interior points.

Other approach is based on the fast approximate nearest neighbors (FLANN) algorithm investigated in [4]. It uses the nearest distance as Euclidean distance, defined as following:

$$D=\sqrt{(x_1-x'_1)^2+(x_2-x'_2)^2+...+(x_{64}-x'_{64})^2},$$

where (x_1, x_2, ..., x_{64}), (x'_1, x'_2, ..., x'_{64}) are detected features of points to be matched.

Any of these methods require the error metrics to be used in approximation or sampling. Calculation of error metrics must be fast and stable, with good repeatability. Geometrical correctness refers to the point being localized on the same image structure or region and can be measured in a number of ways. A simple approach is to evaluate the correspondences by eye, but this is laborious and lacks robustness. Another method is used that relates the points by a homography. Homography matrix maps points **X** in one image to the points **X'** in another. Using this function it is possible to determine whether a point has a correspondence by performing the mapping operation and checking for detected points within a neighborhood of the target location.

4. Error metrics comparison

Three main error metrics are selected to compare their efficiency by SURF method: sum of absolute differences (SAD)

$$\mathbf{E}_{SAD} = \sum \left| \mathbf{D}_i - \mathbf{D}_j \right|, \tag{5}$$

sum of squared differences (SSD)

$$\mathbf{E}_{SSD} = \sum \left(\mathbf{D}_i - \mathbf{D}_j \right)^T \cdot \left(\mathbf{D}_i - \mathbf{D}_j \right), \tag{6}$$

and normalized cross-correlation (NCC)

$$\mathbf{E}_{NCC} = \sum \mathbf{D}^T{}_i \cdot \mathbf{D}_j , \tag{7}$$

where $\mathbf{D}_i$, $\mathbf{D}_j$ - matrices of detected feature points *i* and *j* on the pair of images. It is obvious that in all of the cases the error metric will be the matrix of dimension *n*-by-*m*, where *n* and *m* - are numbers of feature pints detected on both compared images.

Realization of SURF method in (Code of SURF listing in MATLAB) was used in practice for experiments. The descriptors are formed as matrix **D** by size 64×*N*, or 128×*N*, where *N* – number of feature points. The function (7) can be found by *single* multiplication of two matrixes of compared image descriptors:

$$\mathbf{E}_{NCC} = \mathbf{D}_i^T \cdot \mathbf{D}_j = = \begin{bmatrix} D_1(1) & \cdots & D_1(N) \\ \vdots & \ddots & \vdots \\ D_{64}(1) & \cdots & D_{64}(N) \end{bmatrix}^T \cdot \begin{bmatrix} D_1'(1) & \cdots & D_1'(n) \\ \vdots & \ddots & \vdots \\ D_{64}'(1) & \cdots & D_{64}'(n) \end{bmatrix} \cdot \tag{8}$$

Since descriptor matrix is already normalized due to peculiarities of SURF method, then it is possible to state that each component of matrix $\mathbf{E}_{NCC}$ is normalized. Search of maximal elements is done by the finding the maximal values of matrix (8) in each row and checking whether this value is greater than the threshold. Only one maximal element in each row is selected since it is supposed that one point on the template will be matched to the single point on the current image.

Error metrics (5), (6) are realized as vector difference and multiplication in loops by indexes n and m in order to form the same matrices $\mathbf{E}_{SAD}$ and $\mathbf{E}_{SSD}$.

5. Experimental results of error metrics comparison

Set of images with known homography matrices was used [7]. Testing of correctness of detected feature points on the pair of images was done by all type of error metrics (5), (6), (7). Tests were done using known homography matrices of images. Results are represented in Tables 1, 2, 3 as the ratios between the general number of detected features *N*, number of true matched points N_{true}, number of false matched points N_{false}. time of calculation of each type of error metrics. Tests were done in MATLAB 7.8.0. Examples of matching are shown in Fig. 2, Fig. 3.

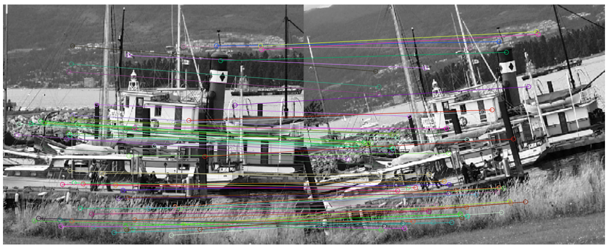

Fig. 2. Matching of two images (a-2) with homography matrix H=[8.5828552e-01, 2.1564369e-01, 9.9101418e+00;-2.1158440e-01, 8.5876360e-01, 1.3047838e+02; 2.0702435e-06, 1.2886110e-06, 1.0000000e+00]

Fig. 3. Matching of two images (b-2) with homography matrix H=[8.5828552e-01, 2.1564369e-01, 9.9101418e+00;-2.1158440e-01, 8.5876360e-01, 1.3047838e+02; 2.0702435e-06, 1.2886110e-06, 1.0000000e+00]

Table 1. Results for NCC metrics (threshold value is 0.97)

Pair of images	N_{true} / N	N_{false} / N	Time, sec
a-2	46/46	0/46	0.2454
a-3	35/139	104/139	0.2081
a-4	3/28	25/28	0.1009
a-5	223/245	22/245	0.3702
a-6	20/61	41/61	0.1037
b-2	125/138	13/138	0.2807
b-3	59/96	37/96	0.2230
b-4	12/33	21/33	0.1554
b-5	3/5	2/5	0.1309
b-6	1/17	16/17	0.1287

Table 2. Results for SSD metrics (threshold value is 0.03)

Pair of images	N_{true} / N	N_{false} / N	Time, sec
a-2	74/74	0/74	4.1528
a-3	30/100	70/100	3.6245
a-4	3/28	25/28	2.0733
a-5	223/245	22/245	2.1784
a-6	20/61	41/61	1.6149
b-2	92/102	10/102	2.9622
b-3	39/67	28/67	3.0226
b-4	4/15	11/15	3.0305
b-5	3/5	2/5	2.4859
b-6	1/17	16/17	2.6992

Table 3. Results for SAD metrics (threshold value is 0.03)

Pair of images	N_{true} / N	N_{false} / N	Time, sec
a-2	68/68	0/68	3.6381
a-3	20/105	85/105	2.9929
a-4	4/26	22/26	1.7089
a-5	230/252	22/252	1.8438
a-6	109/537	428/537	1.4983
b-2	122/136	14/136	2.5629
b-3	54/97	43/97	2.5954
b-4	10/25	15/25	2.5311
b-5	4/6	2/6	2.2979
b-6	0/7	7/7	2.1109

Results with small number of detected points and small number of true matching can be explained by the fact that in such pairs of images there are significant distortions of images (rotation more than 90 degrees for pair a-4, and perspective transformation for pair b-5 and b-6). And as known SURF method is not affine invariant. Graphical comparison of given error metrics is represented in Fig. 4.

As can be seen from Fig. 4 the accuracy of matching is approximately the same for all three error metrics but time of calculation is significantly smaller for NCC error metrics.

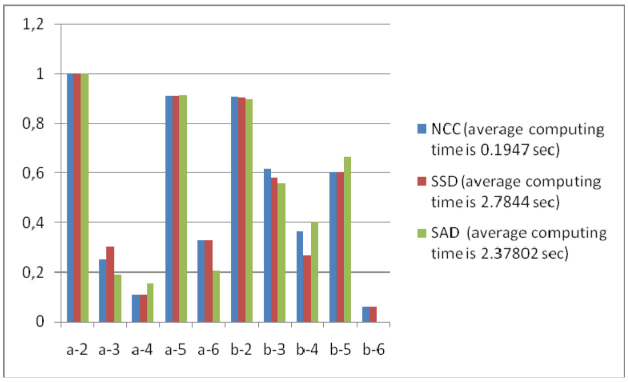

Fig. 4. Comparison of error metrics for tested pairs of images

6. Conclusions

NCC error metrics demonstrates the same results as widely used SSD with significantly being more effective in computing. The time efficiency is about 14 times greater than for both SSD and SAD error metrics that is easily explained by decreasing the number of calculation and excluding the loop iterative operations of vector difference and multiplication and replacing them by single matrix multiplication.

References

[1] *H. Bay, T. Tuytelaars, L. Van Gool*. SURF: Speeded up robust features. European Conference on Computer Vision, Vol. 1, pp.404–417, 2006.

[2] *Y. X. Chen, Q. S. Sun, H. Y. Xu, et al.* "Matching method based on SURF algorithm and RANSAC algorithm" // Journal of Frontiers of Computer Science and Technology, vol. 6, 2012. - p. 1-9.

[3] D. G. Lowe “Distinctive Image Features from Scale-Invariant Keypoints” // International Journal of Computer Vision, vol. 60, no. 2, pp. 91–110, 2004.

[4] M. Marius, G. L. David "Fast approximate nearest neighbors with automatic algorithm conjugation" // Proc. Computer Vision Theory and Applications' 09, 2009, pp. 331-340.

[5] K. Mikolajczyk, C. Schmid “A Performance Evaluation of Local Descriptors” // Proceedings of Computer Vision and Pattern Recognition, June, 2003.

[6] *J. Shi, C. Tomasi* “Good features to track” // IEEE Conference on Computer Vision and Pattern Recognition (CVPR94), 1994.

[7] http://www.robots.ox.ac.uk/_vgg/research/affine/

Received 10 June 2014

М. П. Мухіна. Порівняння метрик похибок в алгоритмі співставлення зображень детектором surf

Національний авіаційний університет, Київ , Україна, проспект Космонавта Комарова, 1
E-mail: m_mukhina@inbox.ru

Метод Speed- Up Robust Feature (SURF) використовується для робастного виявлення характерних точок зображень. Проаналізовано алгоритми співставлення характерних точок. Проведено порівняння різних метрик похибок за їх точністю та обчислювальною ефективністю на серії тестових зображень для базових перетворень, таких як масштабування, поворот і зсув.

Ключові слова: матриця гомографії; метрика похибки; нормалізована взаємна кореляція; прискорений робасний детектор характерний ознак; характерна точка.

М. П. Мухина. Сравнение метрик погрешностей в алгоритмах сопоставления изображений детектором surf

Национальный авиационный университет, Киев, Украина, проспект Космонавта Комарова, 1
E-mail: m_mukhina@inbox.ru

Метод Speed-Up Robust Feature (SURF) используется для робастного обнаружения характерных точек изображений. Проанализированы алгоритмы сопоставления характерных точек. Проведено сравнение различных метрик погрешностей по их точности и вычислительной эффективности на серии тестовых изображений для базовых преобразований, таких как масштабирование, поворот и смещение.

Ключевые слова: метрика ошибки, характерная точка, матрица гомографии, нормализированная взаимная корреляция, ускоренный робастный детектор характерных признаков.

Mukhina Maryna
Candidate of Engineering. Associate Professor.
Department of Aviation Computer-Integrated Complexes, National Aviation University, Kyiv, Ukraine.
Education: National Aviation University, Kyiv, Ukraine.
Research area: image processing, correlation-extremal navigation, data fusion.
Publications: 35.
E-mail: m_mukhina@inbox.ru