

Real-time multispectral video panorama construction

I A Kudinov¹, O V Pavlov¹, I S Kholopov^{1,2} and M Yu Khramov¹

¹Scientific and Design Center of Video and Computer technologies, Ryazan State Instrument-making Enterprise, Seminarskaya Street 32, Ryazan, Russia, 390005

²Ryazan State Radio Engineering University, Gagarina Street 59/1, Ryazan, Russia, 390005

Abstract. An algorithm for the video panorama construction from distributed multispectral cameras data is described. It is shown that operations of vision enhancement (modified Multiscale Retinex algorithm) and multispectral image fusion are implemented for two independently chosen regions of interest with a frame size of 1024x768 pixels and 30 fps using CUDA technology.

1. Introduction

The real-time automatic generation of high-resolution video panoramas from information of several cameras with partially overlapping fields of view (FoV) is one of the modern trends in the vision systems development. Generally, panorama navigation implies the presence of a user-controlled region of interest (RoI). This approach is an alternative for mechanical drive-based vision systems, as it ensures simultaneous operation of several users with an independent choice of a personal RoI without mechanically moving the camera system. Another advantage of distributed panoramic systems (DPS) is the integration of video cameras into the body / fuselage of the carrier object, what positively affects its aerodynamic properties.

2. Methods of panorama construction

There are two basic approaches for panorama stitching:

- an approach based on finding the correspondence of homogeneous pixel coordinates on the frames of cameras with i and j numbers by detecting and matching keypoints using their descriptors and estimating the homography matrix \mathbf{H}_{ij} [1, 2];
- an approach based on finding the correspondence of homogeneous pixel coordinates by preliminarily calibrating the cameras of the DPS by the test object with an auxiliary wide-angle camera [3] if the fields of view of cameras have small angular sizes of the intersection zone or do not intersect at all.

The advantage of the first approach is operability even in the absence of a priori information about the mutual placement of DPS cameras; the advantage of the second approach is operability in difficult observation conditions and in low-contrast observable scenes.

3. Multispectral images superimposition

It is known [4] that one of the main approaches for increase of situational awareness in poor visibility conditions is the simultaneous use of different spectral ranges cameras: visible TV and infrared (IR) – near IR (NIR), short wave IR (SWIR), medium wave IR (MWIR) and long wave IR (LWIR). Panorama stitching from different spectral frames for each of the methods considered above is

hindered by the different physical nature of the images formed by TV and IR cameras: TV, NIR and SWIR cameras see the light reflected by the object in the wavelength ranges 0.38...0.76, 0.7...1.1 and 0.9...1.7 μm respectively whereas MWIR and LWIR cameras see only the thermal radiation of the object at wavelengths 3...5 and 8...12 μm respectively. Therefore, in order to construct a video panorama in the DPS with different spectral ranges cameras (depending on the selected method of the panorama stitching), either a solution of the problem of keypoint matching on TV and IR frames or the production of a universal (having high contrast in the all operating spectral ranges) pattern for camera calibration is necessary.

3.1. Multispectral images keypoints matching

The analysis of publications [5-9] allows distinguishing of four basic approaches for automatic superimposition of multispectral images:

- based on the transition to the decorrelated colour space and the use the SIFT method: this approach is applicable only for combining RGB frames of the visible range and NIR [5];
- based on the mutual morphological correlation of pre-segmented images [6];
- based on the results of the contour analysis [7],
- based on estimating the homography matrix by manually matches search [8, 9].

Restrictions on the use of these methods are situations where the image in one of the video channels (usually in TV) does not allow selecting of the important details of the scene for matches searching (for example, in near to zero illumination, as well as in dense smoke or fog).

The solution of the superimposition problem can also be achieved by mechanical alignment, which is to ensure the parallelism of the sighting lines and the same angular dimensions of the field of view of the each spectral range cameras, as well as their mutual placement minimizing parallax, but this approach is not applicable for the DPS.

3.2. Multispectral camera calibration with universal pattern

Photogrammetric calibration of DPS cameras is the most universal approach for combining images in the far zone, but requires the use of a test pattern that have high contrast simultaneously in several spectral ranges. Examples of such patterns with typical "chessboard" image are considered in [10]. The results of the calibration allow us to evaluate the matrix of internal and external parameters of the multispectral DPS cameras.

4. Algorithm of the video panorama construction according to information from multispectral cameras

As the images of DPS cameras are characterized by geometric distortions caused by different shooting angles, in order to minimize them it is expedient to construct a panoramic frame on a virtual surface of uniform curvature: a sphere or a cylinder of unit radius [11]. Implemented by the authors the algorithm of the moving along the spherical video panorama RoI filling [12] for the DPS with pre-calibrated multispectral cameras contains the following steps.

1. Initialization: calculating quaternions \mathbf{q}_{uv0} , that specify the initial angular direction to the pixels of RoI [12]. If it is necessary to dynamically change the field of view the RoI, the quaternion \mathbf{q}_{uv0} is recalculated in the body of the main operation cycle.

Main operation cycle

2. Estimation of the current angular position of the DPS reference camera by pitch θ and roll φ (for example, according to the data from the inertial measurement unit) and corresponding rotation matrix:

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

3. Calculation of the rotation quaternion for a given angular position of the line of sight (center of RoI) \mathbf{q}_{vis} and quaternions \mathbf{q}_{uv} , determining the current angular position of the RoI pixels (u, v):

$$\mathbf{q}_{uv} = \mathbf{q}_{vis}\mathbf{q}_{uv0}. \quad (2)$$

4. Filling RoI for each spectral range with pixels from cameras by re-projecting points from the surface of a unit virtual sphere to their matrices (with lens distortion compensation [13]).

5. The implementation of the blending procedure [1, 14] for the RoI of each spectral range.

6. Pixel-level image fusion [15] according to selected fusion algorithm.

Since for each pixel of the RoI the processing according to the above algorithm is homogeneous, this allows us to apply the procedure of parallelizing the computations, for example, using the resources of the GPU.

5. Description of the DPS layout

The layout of the DPS is a development of previous authors' work [12] and in addition to grayscale TV cameras contains a LWIR thermal camera with a field of view of $50^\circ \times 40^\circ$ (figure 1). The mutual angular position of the fields of view of the TV cameras and the thermal camera in the sector $200^\circ \times 120^\circ$ is shown in figure 2. To synchronize frames in time, external sync pulses are applied.



Figure 1. RPM layout cameras: LWIR (bottom center) and five TV.

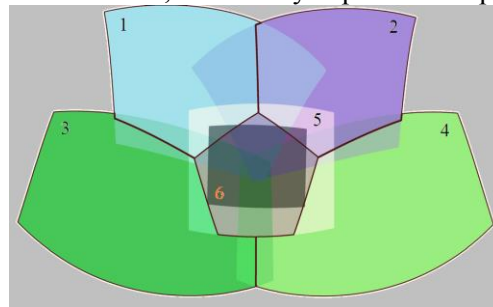


Figure 2. The intersection of the fields of view of the cameras from figure 1.

For filling of the RoI the computations are divided into parallel blocks (64 horizontally and 48 vertically) with 256 threads in each (16 threads horizontally and 16 vertically) using CUDA and CUDA C. As copying of the data of the CPU memory into the GPU memory and back is relatively slow, the number of such operations is minimized in our implementation of the video panorama.

In the DPS layout are implemented:

- angular position control of the line of sight of the operator: according to the data from the head tracking system or (if not) from the joystick;
- independent display of two RoIs with a dynamic change of its field of view from $80^\circ \times 60^\circ$ (wide angle) to $10^\circ \times 7.5^\circ$ (telephoto angle);
- blending according to the algorithm [14] (figure 3);
- increasing the contrast of the TV image (figure 4) using the modified Multiscale Retinex algorithm [16]: in order to accelerate calculations to estimate the brightness of the background, instead of smoothing by a Gaussian filter a box filter is used;
- RoI display mode selection: TV, grayscale thermal image, false-color thermal image (Jet [17] and Cold-to-Warm [18] color maps are realized), contrasted TV (Multiscale Retinex), image fusion mode – grayscale [19] or false-color [19-21]); results are shown in figures 5 and 6.
- mapping of the mutual angular position of the RoIs of the first and second operators.

As the DPS layout currently contains a single thermal camera, the fusion mode is realized only at angular positions of the RoI containing a part of area 6 (see figure 2). Otherwise, the user in this mode receives information only from the TV cameras. This is illustrated in figure 5, where in the thermal camera and fusion modes the lower rows of the RoI are filled with information from the TV cameras, because at the current position of the line of sight their angular coordinates do not fall in the field of view of the LWIR camera and therefore do not contain data in infrared spectral band.

On the NVIDIA GeForce GTX 560 Ti GPU (384 cores) with the maximum amount of calculations (blending, false-color fusion of contrasted TV and IR channels) and the RoI size 1024×768 pixels for

each of the two operators, the rate of updating the information is 32 Hz. The calculation speed for other modes is given in the table, where for comparison, information on the processing speed is also provided when implementing a video panorama on an Intel Core i5 CPU. All values in the table are rounded to the nearest whole number.

Table 1. The maximum rate of updating information in the RoIs for two users in different operating modes, Hz.

	Mode 1: TV without blending	Mode 2: TV with blending	Mode 3: Mode 2+ Multiscale Retinex	Mode 4: Mode 3 + false-color fusion
GPU	77	39	35	34
CPU	4	2	2	2

As can be seen from the table, the use of parallel computations increases the processing speed by an average of 16 times.



Figure 3. Blending results (observation conditions – cloud cover, 3x digital zoom): left – original frame in RoI, right – after blending.



Figure 4. Multiscale Retinex results (observation conditions – sunny day): left – original frame, right – after Multiscale Retinex.



Figure 5. Grayscale fusion (observation conditions – after rain, twilight): left – LWIR RoI layer, center – TV RoI layer, right – fusion frame.

6. Conclusion

With the use of CUDA technology the algorithm for creating a video panorama based on information from multispectral cameras for two windows of interest with a resolution of 0.7 Mp implements an

independent display of video information and vision enhancement functions (Multiscale Retinex and pixel-level fusion) with a frequency of at least 30 fps.



Figure 6. Fusion results for false-color technique (observation conditions –after rain, partly cloudy): top left – LWIR RoI layer, top right – TV RoI layer, bottom left – false-color fusion frame, bottom right – frame from external RGB camera for comparison.

7. References

- [1] Brown M and Lowe D 2007 *Int. J. Comput. Vision* **74**(1) 59-73
- [2] Fischler M and Bolles R 1981 *Commun. ACM* **24**(6) 381-395
- [3] Shirokov R I and Alekhnovich V I 2014 *Contentant* **4** 10-23
- [4] Knyaz V A, Vygolov O V, Vizilter Y V, Zheltov S Y and Vishnyakov B V 2016 *Proc. SPIE* **22** 984022
- [5] Brown M and Susstrunk S 2011 *Proc. IEEE CVPR* (Washington DC: IEEE Comput. Soc.) 177-184
- [6] Vizilter Y V, Zheltov S Y, Rubis A Y and Vygolov O V 2016 *J. Comput. Syst. Sci. Int.* **55** 598-608
- [7] Efimov A I, Novikov A I and Sablina V A 2016 *Proc. 5th Mediterr. Conf. Embedded Comput. (Bar)* 132-137
- [8] Bhosle U, Roy S D and Chaudhuri S 2005 *Pattern Recognit. Lett.* **26**(4) 471-482
- [9] Efimov A I and Novikov A I 2016 *Computer Optics* **40**(2) 258-265 DOI: 10.18287/2412-6179-2016-40-2-258-265
- [10] St-Laurent L, Mikhnevich M, Bubel A and Prévost D 2017 *Quant. Infrared Thermography J.* **14**(2) 193-205
- [11] Szeliski R 2006 *Found. Trends Comput. Graphics Vision* **2**(1) 1-104
- [12] Kudinov I A, Pavlov O V, Kholopov I S and Khranov M Yu *CEUR Workshop Proc.* **1902** 37-42
- [13] Brown D C 1971 *Photogramm. Eng.* **37**(8) 855-866
- [14] Burt P and Adelson E 1983 *ACM Trans. Graphics* **2**(4) 217-236
- [15] Li S, Kang X, Fang L, Hu J and Yin H 2017 *Inf. Fusion* **33** 100-112
- [16] Jobson D J, Rahman Z and Woodell G A 1997 *IEEE Trans. Image Proc.* **6**(7) 965-976
- [17] *MATLAB Jet Array* (Access mode: <https://www.mathworks.com/help/matlab/ref/jet.html>)
- [18] Moreland K 2009 *Proc. 5th Int. Symp. Adv. Visual Comput. (Las Vegas)* **II** 92-103
- [19] Zheng Y 2011 *Image Fusion and its Applications* (Rijeka: inTech)
- [20] Li G, Xu S and Zhao X 2010 *Proc. SPIE* **7710** 77100S
- [21] Kholopov I S 2016 *Computer Optics* **40**(2) 266-274 DOI: 10.18287/2412-6179-2016-40-2-266-274