Engineering, **2013, 5, 199-202**
http://dx.doi.org/10.4236/eng.2013.510B042 Published Online October 2013 (http://www.scirp.org/journal/eng)

Matching DSIFT Descriptors Extracted from CSLM Images

Stefan G. Stanciu[1,2*], Dinu Coltuc[3], Denis E. Tranca[1], George A. Stanciu[1]
[1]Center for Microscopy-Microanalysis and Information Processing, University Politehnica of Bucharest, Bucureşti, Romania
[2]Light Microscopy and Screening Center, Swiss Federal Institute of Technology, Zurich, Switzerland
[3]Electric Engineering Department, Valahia University of Târgovişte, Târgovişte, Romania
Email: *stefan.stanciu@cmmip-upb.org

Received May 2013

ABSTRACT

The matching of local descriptors represents at this moment a key tool in computer vision, with a wide variety of methods designed for tasks such as image classification, object recognition and tracking, image stitching, or data mining relying on it. Local feature description techniques are usually developed so as to provide invariance to photometric variations specific to the acquisition of natural images, but are nonetheless used in association with biomedical imaging as well. It has been previously shown that the matching of gradient based descriptors is affected by image modifications specific to Confocal Scanning Laser Microscopy (CSLM). In this paper we extend our previous work in this direction and show how specific acquisition or post-processing methods alleviate or accentuate this problem.

Keywords: Local Features; Local Descriptors; Feature Matching; SIFT; CSLM

1. Introduction

The detection and description of affine-invariant regions have been regarded as high interest topics during the past decade. Image matching using local invariant features represents a key method used in many computer vision tasks such as image retrieval [1], recognition [2,3], wide baseline matching [4], building panoramas [5], microscopy image stitching [6], image based localization [7,8] or medical image classification [9,10]. In these applications, local invariant features are detected independently in each image and then the features of one image are matched against the features of other images by direct or indirect comparisons of their respective feature descriptors. The matched features can subsequently be used to indicate presence of a particular object, to vote for a particular image, to establish correspondences for epipolar geometry estimation, or to classify an image as belonging to a specific class. For all the above tasks, the core of the application is based on interest point correspondences between individual image pairs or between an image and a class of images. Among various methods reported in the literature, the Scale-Invariant Feature Transform (SIFT) [11] became one of the most preferred choices for local feature detection/description because of its high accuracy, relatively low computation time and the availability of open-source implementations.

Confocal scanning laser microscopy (CSLM) represents an essential imaging tool for many research fields. It provides the possibility to acquire in-focus images from selected depths (optical sections) from both living and fixed specimens in a non-invasive manner. The optical sectioning capability is given by the presence of a pinhole aperture which acts as a spatial filter at the conjugate image plane, rejecting out of focus light [12]. The dimension of the pinhole aperture is responsible for the thickness of the imaged optical section. A stack of optical sections, imaging 2D confocal planes collected at different volume depths can be used to create 3D reconstructions of the imaged specimen.

In CSLM the illumination light is scanned onto the specimen point by point by a mirror on galvano-motor-driven scanner and the light that is emitted from the specimen is likewise collected and de-scanned. The in-focus light that passes the pinhole reaches a photomultiplier tube (PMT), which detects light and converts photon hits into an analogue electron flow. Raising gain (voltage) on the PMT can amplify a weak signal but also amplifies the noise. It is usual that pinhole changes are accompanied by PMT Gain adjustments for reaching a balance between the signal intensity and the background noise. Narrowing the pinhole aperture leads to a reduced volume contributing to the image, resulting in lower image intensity and the need for higher signal amplification. Reciprocally, increasing the pinhole aperture leads to

*Corresponding author.

higher signal and the PMT gain is modified in order to avoid pixel saturation.

It was previously shown that image modifications associated with pinhole aperture or PMT gain adjustments pose problems to gradient based techniques designed for the detection and description of affine-invariant regions [6,13]. The experiment presented in this paper extends our previous investigations in this direction, showing how three usual CSLM image enhancement methods alleviate or accentuate this problem. These three techniques are line averaging, spatial filtering and deconvolution.

2. Methods

2.1. Image Acquisition

The image set that we use has been collected on a mouse kidney section, labeled by Alexa Fluor 488 WGA (Invitrogen, Molecular Probes) by using a Zeiss LSM 510 CSLM system. We have imaged the same field of view under five combinations of the pinhole aperture and PMT gain, resulted from concomitantly decreasing the PMT gain when increasing the dimension of the pinhole aperture. The pinhole aperture was varied between 1 and 2 Airy Units (AU) in steps of 0.2 AU, while the PMT gain was varied between 450 and 400 Zeiss LSM 510 Units (ZU). For each of the six "pinhole-PMT" gain combinations, we have imaged 20 optical sections of 450 μm × 450 μm, collected at 0.750 μm steps along the z axis by using a 20x − 0.8 NA objective. Higher pinhole aperture corresponds to higher optical section thickness. The presented results have been achieved by using as support a reference image of the stack automatically detected by using the reference frame estimator introduced [14].

For excitation we have used a 488nm Ar laser line. The fluorescence signal was collected by passing the emitted light through a 530 - 595 nm band pass filter. In **Figure 1**, we present the brightest image of the stack collected at highest pinhole aperture/lowest pmt gain combination.

2.2. Descriptor Extraction

The SIFT keypoint descriptor is a histogram representation that combines local gradient orientations and magnitudes from a certain neighborhood around a keypoint. More precisely, the descriptor is in fact a 3D histogram of gradient location and orientation, where location is quantized into a 4 × 4 location grid and the gradient angle is quantized into 8 orientations, one for each of the cardinal directions. The resulting descriptor is a normalized vector with the dimension of 128 elements [11].

The SIFT technique provides solutions for both keypoint detection and description. In this experiment we concentrate our attention to the description capabilities of SIFT, extracting descriptors from fixed locations corresponding to a grid. In this purpose we employ the "vl_dsift" function of the VL-Feat library [15] for calculating DSIFT descriptors at fixed grid locations, which according to the authors is "roughly equivalent to running SIFT on a dense grid of locations at a fixed scale and orientation".

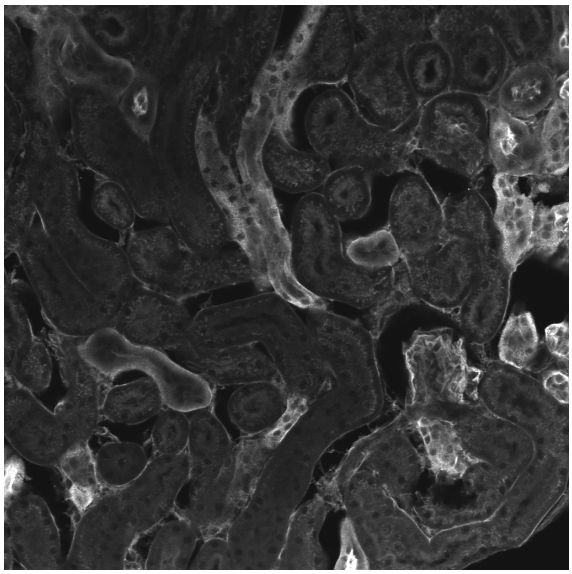

Figure 1. Confocal optical section of mouse kidney tissue collected at 1 AU pinhole aperture/450 ZU PMT gain.

We use a10 pixel grid spacing, resulting in 10,404 features per image. The evaluated sizes for the SIFT bins, are 4, 6 and 8 pixels.

2.3. Evaluated Methods

Line averaging is a usual CSLM acquisition method that is used for compensating low SNR at the expense of bleaching. It consists in scanning the same line for a specified number of times before adding an averaged instance to the image and moving on to the next line. The averaged instance that is added to the image is the arithmetic mean of the summed pixel values from a specified number of scans. By averaging, persistent image content is preserved while fluctuated image content (usually noise) is attenuated.

Median Filtering is a common nonlinear digital filtering technique that is used to remove noise while preserving edges [16]. It evaluates in turns each image pixel and decides whether it is representative for its surroundings or not. The pixel values are replaced by the median of the pixels lying in a specified neighborhood. If the specified neighborhood contains an even number of pixels, the average of the two middle pixel values is used. Median filtering is demonstrably better than Gaussian blur at removing noise whilst preserving edges for a

given, fixed window size.

Deconvolution techniques are routinely used in microscopy imaging for compensating the effect of the unavoidable convolution with the Point Spread Function (PSF) of the optical signal generated by the sample [17]. This process can be mathematically expressed by the following equation:

$$g = f \times h$$

where g represents the collected image which generated through the convolution of the real optical signal (f) object) and the system's PSF (h). Deconvolution consists in solving Equation (2) in order to find out f, knowing both g and h. For deconvolving the image we have used a used a Classic Maximum Likelihood Estimation (CMLE) method available in the Huygens Professional (SVI, Netherlands) software platform.

3. Results

We consider all two-fold pairs of images in the set. The first image of a pair is always the image collected at a higher pinhole aperture and lower PMT gain. Each of the descriptors extracted from the first image in the pair are matched against the descriptors extracted from the other image by using a nearest-neighbour approach. The distance that we use is Euclidean. If the matched nearest-neighbor is the descriptor extracted from the same x, y coordinates we consider to have found a "true positive", otherwise a "false positive". The performance of the nearest-neighbor matching of the descriptors is evaluated in terms of precision (Equation (1)):

$$\text{Precision} = \text{True positives} / (\text{True positives} + \text{False Positives}) \quad (1)$$

In **Table 1** we show the calculated precision in case of the nearest-neighbor matching of DSIFT descriptors extracted from the image set collected without line averaging and not post-processed—"RAW". In **Table 2** we refer to the precision associated to the three other evaluated image sets: image set collected without line averaging and post-processed by median filtering (3 × 3 median filter)—"MF"; image set collected without line averaging and deconvolved by a CMLE approach available in Huygens Professional—"DEC"; image set collected with line averaging (4 time averaging)—"AV4".

In the case of the "RAW" image set we observed a precision increase with higher bin size. Median filtering provides a slight improvement ranging from 4% to 7% depending on the considered bin size. The image set resulted after deconvolution is associated a massive decrease of precision when compared to the RAW image set. The precision decrease varies with bin size and the lowest value is observed in the case of the lowest considered bin size 4, going as low as 48% in this case. In the case of the image set collected under lie averaging we can observe increased precision when compared to the RAW image set. This increase is more consistent in the case of lower bin sizes, going as high as 15% for the lowest considered bin size. It should be noted that in the case of this image set the increase comes at the cost of light exposure, since each image is scanned four times before being added to the image.

Table 1. Precision of nearest-neighbor matching calculated for the image set collected without averaging and not post-processed ("RAW").

Image set	Precision		
	Bin size		
	4	6	8
RAW	0.43	0.56	0.63

Table 2. Nearest-neighbor matching precision difference for image sets "MF", "DEC", "AV4" in respect to the "RAW" image set.

Image set	Precision difference		
	Bin size		
	4	6	8
MF	104%	106%	107%
DEC	48%	56%	61%
AV4	115%	108%	105%

4. Conclusion

Image modifications associated with combined pinhole aperture dimension—PMT gain changes raise problems to gradient based local feature description. These problems can be alleviated or accentuated by specific CSLM image acquisition or image post-processing methods. By the experiment that we present in this paper we place a first step in the direction of identifying the methods that affect feature description and the ones that could be used to increase the performance of gradient based description techniques. We have evaluated three usual techniques that are commonly used for CSLM image enhancement. We have observed that median filtering and line averaging are associated with an increase in the precision of DSIFT descriptor based matching, while deconvolution yields negative effects in this regard. We consider that research efforts placed in this direction are important as a wide variety of biomedical computer vision applications rely on local feature description and matching and their efficient optimization cannot be achieved without identifying specific methods that need to be avoided and ones that need to exploited for enhancing the results.

5. Acknowledgements

The presented work was supported by the UEFISCDI

PN-II-PT-PCCA-2011-3.2-1162 Research Grant and the CRUS SCIEX NMS-CH Fellowship nr. 12.135. The corresponding author thanks Dr. Gábor Csúcs, Dr. Tobias Schwarz and Dr. Joachim Hehl, of the Light Microscopy and Screening Center of ETH Zurich for their support and advice.

REFERENCES

[1] L. J. Zhi, S. M. Zhang, D. Z. Zhao, H. Zhao, S. K. Lin, D. Z. Zhao and H. Zhao, "Medical Image Retrieval Using SIFT Feature," *Proceedings of the* 2009 2*nd International Congress on Image and Signal Processing*, Vol. 1-9, 2009, pp. 2252-2255. http://dx.doi.org/10.1109/CISP.2009.5304112

[2] G. Kordelas and P. Daras, "Viewpoint Independent Object Recognition in Cluttered Scenes Exploiting Ray-Triangle Intersection and SIFT Algorithms," *Pattern Recognition*, Vol. 43, 2010, pp. 3833-3845. http://dx.doi.org/10.1016/j.patcog.2010.05.030

[3] M. Brown and S. Susstrunk, "Multi-Spectral SIFT for Scene Category Recognition," 2011 *IEEE Conference on Computer Vision and Pattern Recognition* (Cvpr), 2011, pp. 177-184.

[4] J. Matas, O. Chum, M. Urban and T. Pajdla, "Robust Wide-Baseline Stereo from Maximally Stable Extremal Regions," *Image and Vision Computing*, Vol. 22, 2004, pp. 761-767. http://dx.doi.org/10.1016/j.imavis.2004.02.006

[5] M. Brown, and D. G. Lowe, "Automatic Panoramic Image Stitching Using Invariant Features," *International Journal of Computer Vision*, Vol. 74, 2007, pp. 59-73. http://dx.doi.org/10.1007/s11263-006-0002-3

[6] S. G. Stanciu, R. Hristu and G. A. Stanciu, "Influence of Confocal Scanning Laser Microscopy Specific Acquisition Parameters on the Detection and Matching of Speeded-Up Robust Features," *Ultramicroscopy*, Vol. 111, 2011, pp. 364-374. http://dx.doi.org/10.1016/j.ultramic.2011.01.014

[7] P. Piccinini, A. Prati and R. Cucchiara, "Real-Time Object Detection and Localization with SIFT-Based Clustering," *Image and Vision Computing*, Vol. 30, 2012, pp. 573-587. http://dx.doi.org/10.1016/j.imavis.2012.06.004

[8] M. Dawood, C. Cappelle, M. E. El Najjar, M. Khalil and D. Pomorski, "Harris, SIFT and SURF Features Comparison for Vehicle Localization Based on Virtual 3D Model And Camera," 2012 3*rd International Conference on Image Processing Theory*, *Tools and Applications*, 2012, pp. 307-312.

[9] J. C. Caicedo, A. Cruz and F. A. Gonzalez, "Histopathology Image Classification Using Bag of Features and Kernel Functions," *Artificial Intelligence in Medicine*, *Proceedings*, Vol. 5651, 2009, pp. 126-135.

[10] T. Tamaki, J. Yoshimuta, M. Kawakami, B. Raytchev, K. Kaneda, S. Yoshida, Y. Takemura, K. Onji, R. Miyaki and S. Tanaka, "Computer-Aided Colorectal Tumor Classification in NBI Endoscopy Using Local Features," *Medical Image Analysis*, Vol. 17, 2013, pp. 78-100. http://dx.doi.org/10.1016/j.media.2012.08.003

[11] D. G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," *International Journal of Computer Vision*, Vol. 60, 2004, pp. 91-110. http://dx.doi.org/10.1023/B:VISI.0000029664.99615.94

[12] J. B. Pawley, "Handbook of Biological Confocal Microscopy," Springer, New York, 2006. http://dx.doi.org/10.1007/978-0-387-45524-2

[13] S. G. Stanciu, R. Hristu, R. Boriga and G. A. Stanciu, "On the Suitability of SIFT Technique to Deal with Image Modifications Specific to Confocal Scanning Laser Microscopy," *Microscopy and Microanalysis*, Vol. 16, 2010, pp. 515-530. http://dx.doi.org/10.1017/S1431927610000371

[14] S. G. Stanciu, G. A. Stanciu and D. Coltuc, "Automated Compensation of Light Attenuation in Confocal Microscopy by Exact Histogram Specification," *Microscopy Research and Technique*, Vol. 73, 2010, pp. 165-175. http://dx.doi.org/10.1002/jemt.20767

[15] A. Vedaldi and B. Fulkerson, "VLFeat: An open and Portable Library of Computer Vision Algorithms," 2008.

[16] R. C. Gonzalez and R. E. Woods, "Digital Image Processing," Addison-Wesley Longman Publishing Co., Inc., Boston, 2001.

[17] W. Wallace, L. H. Schaefer and J. R. Swedlow, "A Workingperson's Guide to Deconvolution in Light Microscopy," *Biotechniques*, Vol. 31, 2001, p. 1076.

