33rd Annual International Conference of the IEEE EMBS
Boston, Massachusetts USA, August 30 - September 3, 2011

Mosaicing of Optical Microscope Imagery Based on Visual Information

Ludovico Carozza[1] and Alessandro Bevilacqua[1,2] and Filippo Piccinini[1]

***Abstract*— Tools for *high-throughput high-content* image analysis can simplify and expedite different stages of biological experiments, by processing and combining different information taken at different time and in different areas of the culture. Among the most important in this field, image mosaicing methods provide the researcher with a global view of the biological sample in a unique image. Current approaches rely on known motorized x-y stage offsets and work in batch mode, thus jeopardizing the *interaction* between the microscopic system and the researcher *during* the investigation of the cell culture. In this work we present an approach for mosaicing of optical microscope imagery, based on local image registration and exploiting visual information only. To our knowledge, this is the first approach suitable to work on-line with non-motorized microscopes. To assess our method, the quality of resulting mosaics is quantitatively evaluated through on-purpose image metrics. Experimental results show the importance of model selection issues and confirm the soundness of our approach.**

I. Introduction

Microscopic imaging of *in vitro* live cells represents an important tool for researchers to study spatial and temporal evolution patterns of cell cultures. Tools for *high-throughput high-content* image analysis, besides saving the microscopist from tedious as well as repetitive and time expensive tasks, can enhance the range of functionalities offered traditionally by the microscope, providing *virtual microscopy* [1] capabilities. For example, a global view of the whole cell culture can be useful to identify special spatial patterns, like cell colonies, or to build a global reference pattern, for subsequent registration or cell tracking. In addition, it can provide the microscopist with a more complete scene understanding of some particular features of the biological sample under investigation. The full-resolution image of the whole cell culture can be then used for subsequent image analysis steps, like cell segmentation, cell counting, multi-modality image fusion, to cite some of them. To these purposes, *image mosaicing* techniques permit to build a wide field-of-view image of the whole cell culture area during the microscopic investigation, while fully preserving the spatial resolution of each single image. Geometric and photometric properties of the scene must be preserved with a high accuracy, since they can affect the subsequent image analysis stages.

Methods and systems currently employed for image mosaicing perform in *batch* mode, building the mosaic at a separate stage at the end of the acquisition of the whole sequence of images. Moreover, geometric registration of the images makes often use of human intervention and/or relies on known motorized x-y stage offsets of the microscope holder [1], [2] to align the images, then requiring the application of a subsequent *global* refinement stage. This approach, besides not working for the most common non-motorized microscopes, even utilizing current-generation motorized stage microscopes however requires the microscopist to wait the end of the acquisition process to achieve the mosaic.

In this work, we describe an effective *on-line* approach used to design an automatic mosaicing method for optical microscopy imagery, exploiting visual information only. It relies on an efficient image registration method, robust to the presence of outliers and partial photometric artifacts (vignetting and shading). As a consequence, it does not need automated equipment, and preserves photometric and geometric consistency during the *manual* motion of the microscope holder. Moreover, its limited computational complexity could make our approach suitable for real-time performance. Accordingly, our approach could permit to browse the cell culture through regions of interest *interactively* within the acquisition process, providing the microscopist with an immediate visual feedback of the explored area.

The paper is organized as follows. In Sect. II, the approaches utilized in this research field are illustrated. In Sect. III, the mosaicing algorithm we devised is discussed. In Sect. IV, we describe the experimental testbed and the resulting mosaics. Their quality is quantitatively assessed and measured through on-purpose metric indexes. Moreover, these measures permit to evaluate the geometrical registration models to be adopted. Finally, Sect. V draws some conclusions regarding the current achievements and proposes future developments.

II. Previous Works

In image mosaicing, different views of overlapping regions have to be *matched* and their geometrical and photometric relations estimated, in order to align them in a common reference frame, relying on image registration methods. Image registration and mosaicing represent an important issue in the computer vision research community and, accordingly, a substantial number of papers has been published in this field in the last two decades.

Images to be registered can be taken from video shots [3] or can present a wide baseline [4], thus affecting the requirements of the matching stage in terms of robustness. Both featureless dense correlation-based methods, working on pixel intensity, and sparse feature-based approaches have been employed in this stage, depending on the required

[1] ARCES (Advanced Research Center on Electronic Systems), University of Bologna, Italy; {lcarozza,abevilacqua,fpiccinini}@arces.unibo.it
[2] DEIS (Department of Electronics, Computer Science and Systems), University of Bologna, Italy; URL:http://cvg.deis.unibo.it

accuracy and computational performance. The work in [5] is representative of the *bundle adjustment* approaches, based on iterative optimization of some non linear cost function on the whole sequence of images. Work in [3] also follows this approach. Accordingly, the high computational payload associated to this optimization stage requires off-line processing.

As far as mosaicing of light microscopy imagery is concerned, the algorithms employed in this context have different hardware requirements and degrees of automation. A first class of algorithms follow a *dense* featureless registration approach with likelihood error functions based on pixel image intensities and accordingly are computational intensive. They have been often focused on *post processing* the whole sequence of images. The work described in [6] concerns a semiautomated method, based on pairwise registration through image cross correlation. Its implementation performs off-line on a desktop computer, without requiring the holder to be motorized. In [7], a semiautomated method, which requires the user to manually align the *tiles* for a subsequent fine registration stage, is presented. This work being focused on accuracy performance (up to sub-pixel level), images are pixel-wise registered using a dense featureless approach, thus resulting in a high computational burden that prevents this method to be used on-line. The methods proposed in [8] and [9] are conceived to be used necessarily with high-precision motorized x-y stages. Metadata provided from motorized stage controllers [8] and mosaic initialization through manual alignment [9] are used for a coarse geometric registration, while global tonal and geometric alignments are performed by minimizing a cost function over the pixel intensities of all the images. Accordingly, these methods work in batch mode at the end of the acquisition.

A second class of algorithms relies on sparse feature-based registration approaches, *detecting* and *matching* salient regions in consecutive images. The method proposed in [10] uses wavelet-based edge correlation to detect feature points and normalized cross correlation for their matching. This method is not conceived for on-line mosaicing since it needs global registration to achieve an accurate mosaic.

III. Methods

The algorithm we devised has been tailored to work on-line and to be compliant with real time performance, while at the same time preserving both photometric and geometrical consistency.

First, inhomogeneous spatial distribution of the microscope light field should be taken into account. With respect to natural images, in microscopy laboratories we can assume that the illumination conditions are relatively well controlled. We have seen from our experience that the limited extent of the area analyzed is more sensitive to the optical layout of the microscopy system rather than to external light variation. This results in evident vignetting and shading effects that must be compensated before warping the frames into a common geometrical reference frame. To this purpose, the microscope luminous field has to be modeled and the acquired images *normalized* accordingly. Light distribution within the microscope's Field Of View (FOV) can be estimated from images of an *empty field*, before positioning the specimen on the holder. Alternatively, the method described in [11], [12] to estimate the background (i.e., the culture medium free of cells) can be used directly during image acquisition.

Geometric alignment has been performed using sequential image registration (Frame-to-Frame, F2F), applied to consecutive frames. A *coarse-to-fine* strategy, which represents a good trade-off between accuracy and computational payload, has been employed to this purpose. Initially, the Shi-Tomasi [13] corner detector is used to extract salient points in the first image, since it is associated with local gradient information, retaining good robustness to noise and illumination changes. At a coarse level, Phase Correlation [14] is applied to estimate an *average* translation vector, common to all the pixels of the previous image, at pixel level accuracy. At a finer level, the translated coordinates of the detected salient points are used as guess locations to feed the Lukas-Kanade-Tomasi (LKT) tracking stage [15], thus yielding sub-pixel accuracy. This approach is more robust to large translations since it overcomes the small-signal approximation upon which the LKT linearization is based.

Once sparse features correspondences are established, the geometrical transform linking the two images can be estimated by Least Square regression on a given model. Since the thickness of the specimen (some microns) is negligible with respect to the working distance of the microscope objective (several centimeters), the scene can be considered as being planar. Under these conditions, corresponding features $(X_i,\ X_j)$ of two consecutive views $(I_i(x,y),\ I_j(x,y))$ are related by a planar homography H_i^j, with $X_j = H_i^j \cdot X_i$. However, taking into account the physical constraints of the microscopy system (small extent of the scene depth, narrow FOV, negligible holder mechanical play, etc.), this general model can be relaxed to nested sub-models, the affine and the translative models, respectively. The Direct Linear Transform algorithm [16] has been employed, jointly with the RANSAC [17] stage, for the robust estimation of these models, being imaged live structures. We have chosen the mosaic reference frame to be coincident with the first frame, so that the mosaic warping matrix for the n^{th} frame, M^n, can be obtained by incrementally chaining the estimated pairwise transform matrices H_i^j, according to:

$$M^n = \prod_{i=0}^{n-1} H_i^{i+1} \tag{1}$$

The n^{th} image is then warped into the mosaic reference frame according to the matrix M^n using bilinear interpolation and merged into the mosaic using a *stitching* approach. This method has been preferred to traditional *blending* approaches in order to avoid *ghosting* effects due to the motion of particles in temporally adjacent frames.

IV. Experimental Results

In order to test our algorithm, image sequences of biological samples have been acquired in phase contrast mode,

using a standard, non-motorized, optical microscopy instrumentation. An inverted microscope Nikon Eclipse TE2000-U, widely used in research labs, has been equipped with a digital camera (Nikon DXM1200) able to perform live acquisition at 640×512 pixel resolution. Our algorithm has been tested on different sequences of several biological samples processed by a consumer PC. Due to the lack of space, here we just report results related to one sequence of 60 images (59 image pairs) of a histological sample (HS, hereinafter) of altered bone tissue. Incremental pairwise registration has been performed using the three warping models in Sect. III. In Fig. 1 (a), the resulting mosaic for

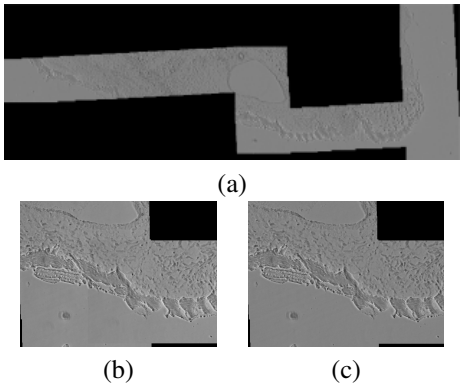

Fig. 1. **(a)** Mosaic obtained from the sequence HS using F2F registration with flat field correction for the translative model. (b) Details of the mosaic achieved without applying the flat field compensation compared with **(c)** the correspondent tonally-aligned region

the translative model only is shown, being the differences with the affine model negligible. Fig. 1 refers to a detail of the mosaic obtained with translative model, without (b) and with (c) the application of the flat field correction stage. In the first case, it can be clearly noticed the presence of seams due to shading and vignetting effects. These artifacts disappear when the flat field correction is applied.

In Fig. 2 (a), the resulting tonally-aligned mosaic for the

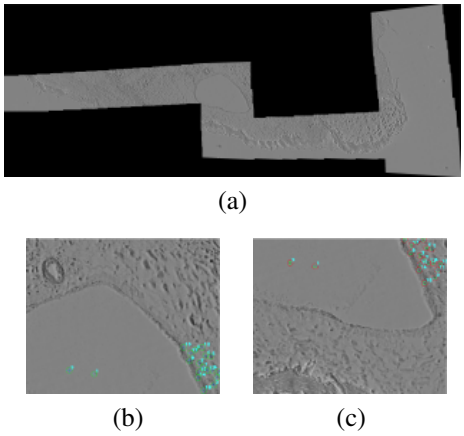

Fig. 2. **(a)** Mosaic obtained from the tonally aligned sequence HS using the projective model. **(b), (c)** An image pair with a small number of features concentrated in one region (ill-conditioned problem for complex models).

projective model is shown. Fig. 2 (b), (c) shows how the error propagation leads to large, unnatural, image deformations, due to the ill-conditioned nature of the estimation of the projective model, being the data used for model estimation small in number and concentrated in a small region [18].

The quality of the resulting mosaic can be assessed quantitatively using proper quality indexes. Since the experimental equipment is not automated, ground-truth data are not available, not even as far as the holder motion is concerned. The only ground-truth data available are the single frames acquired during the holder motion, that is the *reference images*. Accordingly, the information contained in the single frame can be compared with its corresponding area in the mosaic according to some *likelihood* metric. Given the sequence of N reference images I^i to be mosaicked, let $R_I^i(x,y)$ be the i^{th} warped reference image, achieved by *warping* $I^i(x,y)$ according to the matrix M^i, and $R_M^i(x,y)$ be the mosaic region having the same support. The area $R_M^i(x,y)$ of the mosaic is generally the result of the contributions of the $R_I^i(x,y)$ and also of other different warped reference images. Accordingly, $R_M^i(x,y)$ is generally partitioned into $m_i + 1$ regions $A_i^i(x,y) \dots A_{i+m_i}^i(x,y)$, each partition containing only the mosaic pixels derived from one of the reference images that are warped into $R_M^i(x,y)$ (see Fig. 3). For the

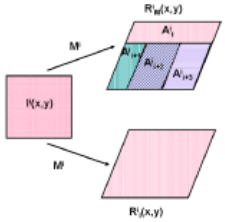

Fig. 3. Explicative figure for the symbols used compute quality metric indexes (m_i=3 in this case).

generic i^{th} reference image, the support S^i used to compute metric indexes is defined as the union of the partition subsets of $R_M^i(x,y)$, except for the A_i^i region which contains information related to the i^{th} reference frame only. This region is simply obtained through the estimated warping, and its pixels must not be taken into account in the comparison between $R_M^i(x,y)$ and $R_I^i(x,y)$, since they would cancel out (up to an interpolation effect) in this comparison. Accordingly, the Root Mean Squared Error (RMSE^i) and the Signal to Noise Ratio (SNR^i) can be defined by considering the image intensities on the support $S^i(x,y)$, with cardinality $P(S^i)$, as follows:

$$RMSE^i = \sqrt{\frac{\sum_{(x,y)\in S^i} (R_I^i(x,y) - R_M^i(x,y))^2}{P(S^i)}} \quad (2)$$

$$SNR^i = \left[\frac{\sum_{(x,y)\in S^i} (R_M^i(x,y))^2}{\sum_{(x,y)\in S^i} (R_I^i(x,y) - R_M^i(x,y))^2} \right] \quad (3)$$

usually expressing SNR in [dB]. These indexes can be extended considering all the supports $S^i(x, y)$, so that *global* RMSE_M and SNR_M can be computed for the resulting mosaic. Table I reports values of these indexes relative to the sequence HS, without (HS-RAW) and with (HS-FF) the application of the flat field compensation, for the three warping models. To assess consistently the effect of the flat field compensation only, the geometric model must be fixed, in order to separate geometric and photometric effects. Fixed a model, by applying the flat field correction the RMSE value decreases, while conversely, as expected, the SNR increases, for all the three models. This is consistent with the improved visual quality of the mosaics due to the flat field correction stage, for all the adopted models.

Sequence	$RMSE_M$			SNR_M[dB]		
	Trasl	Affine	Proj	Trasl	Affine	Proj
HS-RAW	4.77	4.36	10.21	29.13	29.91	20.51
HS-FF	2.72	1.99	8.62	33.52	36.22	24.78

TABLE I

QUALITY METRIC INDEXES FOR THE 60-FRAME HS SEQUENCE.

In order to evaluate consistently the *soundness* of the selected model with respect to their geometric distortions, the values of the metric indexes for the three models must be compared *once* the flat field correction has been applied, that is comparing values on the HS-FF row only. Considering the RMSE in the HS-FF row, for both translative and affine models the corresponding values are quite small, both in absolute terms and in comparison with the gray level ranges of the images (0-255). For the projective model, the RMSE values are bigger, due to the large image distortion introduced by this model. The SNR values show, as expected, an opposite trend, decreasing for the projective model. Accordingly, in absolute terms the translative and the affine model performances are very similar (the latter being slightly better), and worsen, as hypothesized, for the projective model. For these reasons, the affine model can be considered a good compromise between robustness and complexity.

V. CONCLUSIONS AND FUTURE WORKS

This work presents the algorithm we have conceived for mosaicing of optical microscope images. It is based on incremental image registration and exploits visual information only. Through the proposed method, microscope capabilities can be extended and researchers provided with a global view of the whole biological sample directly *during* its investigation. Since our algorithm is based on visual information only, thus not relying on external automated equipment, several issues had to be faced. Photometric effects, like vignetting and shading artifacts, must be compensated. Furthermore, geometrical registration has to comply with sub-pixel accuracy, preserving robustness even in the presence of outliers. An incremental coarse-to-fine registration approach has permitted to face these problems at an acceptable computational payload, being suitable to work on-line. Experimental results have confirmed the efficacy of our approach, emphasizing how a correct choice of warping model can improve the quality of geometric registration.

While going on with system integration, these considerations are prompting us to deepen the study of some robust model selection criteria. Also, strategies to compensate, at an acceptable time cost, for error drift effects intrinsic to our frame-to-frame approach, are currently under investigation. Finally, we are extending our approach to color images, although the mosaic still undergoes false color artifacts.

REFERENCES

[1] D. J. Romer, K. H. Yearsley, and L. W. Ayers, "Using a modified standard microscope to generate virtual slides," *Anat. Rec. (Part B: New Anat.)*, vol. 272, no. 3, pp. 91–97, 2003.

[2] M. Gagnon, R. Taylor, J. Lange, T. Lee, C. Collins, R. Draut, and E. Kujawsky, "Automated microscopic image acquisition, compositing and display - us patent 7027628b1," 2006.

[3] R. Marzotto, A. Fusiello, and V. Murino, "High resolution video mosaicing with global alignment," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2004*, vol. 1, 2004, pp. 692–698.

[4] M. Brown and D. Lowe, "Automatic panoramic image stitching using invariant features," *International Journal of Computer Vision*, vol. 74, no. 1, pp. 59–73, 2007.

[5] P. F. Mclauchlan and A. Jaenicke, "Image mosaicing using sequential bundle adjustment," *Image and Vision Computing*, vol. 20, no. 1-2, pp. 751–759, August 2002.

[6] S. R. Ott, "Acquisition of high resolution digital images in video microscopy: automated image mosaicking on a desktop microcomputer," *Microscopy research and techniques*, vol. 38, no. 3, pp. 335–339, 1997.

[7] P. Thevenaz and M. Unser, "User-friendly semiautomated assembly of accurate image mosaics in microscopy," *Microscopy research and techniques*, no. 70, pp. 131–146, 2007.

[8] C. Sun, R. Beare, V. Hilsenstein, and P. Jackway, "Mosaicing of microscope images," in *Proceedings of IEEE Digital Image Computing: Techniques and Applications 2005 (DICTA '05)*, 2005, pp. 343–348.

[9] S. K. Chow, H. Hakozaki, D. L. Price, N. A. B. MaLlean, T. J. Deernick, J. C. Bouwer, M. E. Martone, S. T. Peltier, and M. H. Ellisman, "Automated microscopy system for mosaic acquisition and processing," *Journal of Microscopy*, vol. 222, no. 2, pp. 76–84, 2006.

[10] W. J. Hsu, W. F. P. Poon, and Y. N. Sun, "Automatic seamless mosaicing of microscopic images: enhancing appearance with color degradation compensation and wavelet based blending," *Journal of Microscopy*, vol. 231, no. 3, pp. 408–418, 2008.

[11] A. Gherardi, A. Bevilacqua, and F. Piccinini, "Illumination field estimation through background detection in optical microscopy," in *IEEE CIBCB*, 2011, pp. 49–54.

[12] A. Bevilacqua, F. Piccinini, and A. Gherardi, "Vignetting correction by exploiting an optical microscopy image sequence," in *Proc. of EMBC 2011 - 33rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (accepted for publication)*, 2011.

[13] J. Shi and C. Tomasi, "Good features to track," in *IEEE International Conference on Computer Vision and Pattern Recognition (CVPR 1994), Seattle, WA, USA, June 21-23*, 1994, pp. 593–600.

[14] H. Foroosh, J. B.Zerubia, and M. Berthod, "Extension of phase correlation to subpixel registration," *IEEE Transactions on Image Processing*, vol. 14, no. 1, pp. 12–22, 2002.

[15] J.-Y. Bouguet, "Pyramidal implementation of the Lukas Kanade feature tracker: Description of the algorithm," Intel Research Laboratory, Technical Report, Tech. Rep., 1999.

[16] R.Hartley and A.Zisserman, *Multiple view geometry in computer vision*. Cambridge Academic Press, 2003.

[17] M. A. Fischler and R. C. Bolles, "Random sample and consensus: A paradigm for model fitting with application to image analysis and automated cartography," *Comm. of the ACM*, vol. 24, no. 6, pp. 381–395, 1981.

[18] Y. Kanazawa and K. Kanatani, "Stabilizing image mosaicing by model selection," *Lecture Notes in Computer Science*, vol. 2018, pp. 35–51, 2000.