Geosci. Instrum. Method. Data Syst. Discuss., 3, C97–C103, 2013
www.geosci-instrum-method-data-syst-discuss.net/3/C97/2013/

Geoscientific Instrumentation Methods and Data Systems Discussions

Interactive comment on "LAPM: a tool for underwater Large-Area Photo-Mosaicking" *by* Y. Marcon et al.

Y. Marcon et al.

ymarcon@marum.de

Received and published: 27 June 2013

We are grateful to the anonymous Referee #2 for his comments and more general questions and suggestions. Please see below our answers to all comments and questions from the anonymous Referee #2.

General comments: The manuscript by Marcon et al. deals with a new software tool which introduces state-of-art methodologies from computer science in a "turn-key" software package for scientists in other disciplines. This topic is very relevant for a wider audience and as such is definitely worth publication. As my appointment as a reviewer came at a later stage, I have also had the opportunity to read the authors' response to another referee's comments. I concur that the manuscript would benefit from reorga-

nization and making it more concise. Also, I think it would be useful to mention in the text that there is a user manual that potential end-users could download for perusal. Response: changed accordingly. We mention in the introduction that a user manual is also provided.

Nevertheless, besides the technical corrections, which are currently being worked, I have a few general observations and questions that I consider worth addressing at least in the sections "Performance and limitation" (Q2 and Q5) and/or "Conclusions" (Q3, Q4 and Q5).

Question #1: There is little about the actual use of the produced mosaicks. Besides being able to create mosaicks, what are the qualities that marine scientists desire in the final images? Is geometric accuracy the most important factor? Or does the further (automatic?) analysis prefer smooth changes in intensities etc.? This should be discussed already in the introduction and maybe elaborated later.

Response: As suggested by Referee #2, we added more information to the introduction about the use of the mosaics and the needs of marine scientists. The deep-sea is a very difficult environment to explore, and it is often hard to get a global view over a large area. Marine scientists need mosaics mainly for spatial analyses (area calculations, habitat descriptions, fauna distribution, spatial correlations, etc.) as well as for navigation purposes. In that regard, geometric accuracy is by far more important than the quality of the visual rendering of the mosaic. Although 'blended' mosaics with smooth intensity changes are visually more appealing, blending options may also gum out some registration errors (duplicated features, slight shifts, etc.) in areas where the quality of image matching is poor (strong relief, low overlap, etc.). In such case, the absence of blending may allow the user to more accurately interpret the observations. The authors do not have experience with automatic image analyses (such as faunal detection), but we would expect that sharp intensity changes may alter the quality of the analyses.

Question #2: How do you quantify errors in mosaicking? In my opinion, this question is very relevant to visualisations of data. As the development proceeds you need error analysis to guide the selection of algorithms (to be added). Can you use test image sets already benchmarked in computer science? Have you extracted overlapping tiles from a high-resolution underwater image and then compared the constructed mosaick with the original? How does the blurring and other degradiations affect the final mosaick? How much overlap is needed? Could this be something that is provided as a convenient demonstration case that the user manual uses as an example?

Response: This is a very good point. We added some discussion in section 6 of the manuscript. We do agree that quantifying errors in mosaicking would be necessary to allow marine scientists to better estimate the error of the spatial analyses. In the case of the LAPM tool, we think that error could be easily estimated by computing and plotting the final registration residuals of every pair of overlapping images. This could allow the user (1) to identify areas of the mosaic where registration errors are the largest and (2) to quantify such errors. We added this to the list of future improvements for the LAPM tool. However, we do not think that quantifying the error would help much in guiding the choice of algorithm. In the case of 2D underwater mosaicking, most errors do not come from the mosaicking technique that is used but from (1) the non-respect of the planar scene assumption (i.e. the relief) and (2) the lens distortions (known as barrel and pincushion distortions). Indeed, for the feature detection and matching, the LAPM tool is based on well-tried techniques, which have been tested and benchmarked by their developers and users (Lowe, 2004, 1999; Vedaldi and Fulkerson, 2010, 2008). From there, the global registration is done by minimizing a cost function, i.e. the global error. Therefore, under ideal conditions (planar scene, no lens distortion), the motion between overlapping images can be entirely explained with a homography transform. For these reasons, the causes of mosaicking errors are mostly extrinsic to the techniques used in the LAPM tool. Unfortunately in 2D mosaicking, little can be done to reduce errors caused by the relief without using image warping techniques (which require extraction of 3-dimensional information from the

images and, hence, are closer to 3D mosaicking techniques). Relief inevitably introduces errors, and causes duplicated features or blurring (if blending is used), no matter how well matched the images are. We explain in section 6 that this issue is common to all homography-based photo-mosaicking techniques. Conversely, errors caused by lens distortions can be reduced by applying some pre-processing to the image dataset prior to mosaicking. Such method has been successfully applied by Pizarro and Singh (2003). We mention in the conclusion that this will be the subject of a future improvement for the LAPM tool. Creating a mosaic with overlapping tiles (that were extracted from a high-resolution underwater image) amounts to the same thing as constructing a mosaic under ideal conditions. Indeed, overlapping areas between tiles are free of distortions and of intensity discrepancies (both from the relief and from the lens) and mosaicking errors or blurring effects are, therefore, negligible. Furthermore, we do not know of any benchmarked test set of underwater images that could be used to test the LAPM tool. Hence, for demonstration we chose to use a real set of underwater images as example in the user manual (and also on Figures 4, 8, 10 of the manuscript). The chosen dataset represents downslope mud flows in the Black Sea, and contains many characteristic linear features, which can be used to assess the correctness of the final mosaic. Figure 8 shows that linear features were well reconstructed during the mosaicking process, which indicates that mosaicking errors are low, despite the slope. The Figure 8 also shows that blurring or degradations of the blended mosaic are also low. We added this information into section 6, and mention that this sample dataset can be provided upon request for testing purposes. About the blurring: blurring occurs only if the blending rendering method is used and in areas where overlapping features are not perfectly aligned (due to relief or lens distortions). In such case it may be preferable not to use the blending option in order to keep the ability to identify areas with large registration errors, and to analyze and interpret them knowledgeably. About the minimum overlap that is needed: in our experience, the automatic feature detection and matching requires at least about 25% overlap. However, manual feature matching can be done with overlaps of any size. Therefore, low overlap is a limitation for the automatic feature matching only. Therefore, we added this information into section 4.3 instead of section 6.

Question #3: Reflecting the needs of marine scientists, have you considered "user-guided adaptive" processing where one could trade speed for accuracy in areas where accuracy is not needed? Or is the workflow such that realtime guidance is not practical? You already mention that low-resolution mosaicks are often used and a multi-scale approach might be very useful. One mouse-click on the area of interest would increase the mosaick resolution locally. (This could also be used for the browsing of very large mosaicks via net, c.f. existing web applications)

Response: Unfortunately, in its present state the LAPM tool does not allow such realtime guidance, and the entire mosaic must be processed at the same resolution. Constructing low resolution mosaics is useful to visualize potential crossovers and improve the global registration but, if high-resolution analyses are needed, then the mosaic must be entirely constructed in high-resolution. However, the mosaic tiles are geo-referenced to be visualized into ArcGIS. ArcGIS is very efficient to browse quickly through very large datasets (by computing multi-resolution pyramids from raster and image datasets).

Question #4: What is your view on and/or have you considered fusing multiple image sources? Is it possible to combine images from cameras and sonars? Again, would marine scientists need this for their research?

Response: We did originally consider combining images from various camera systems to enable the construction of mosaics using datasets from various research cruises. Therefore, some parts of the LAPM tool, such as the feature detection, do accept multi-resolution image datasets. However, we encountered numerous technical difficulties (mainly display issues in the graphical interfaces, homography computation, image resampling, and mosaic construction), and we abandoned the implementation of this capability for now. Fusing images from camera with data from ROV-borne multibeam

echosounders (micro-bathymetry data) would, however, be very interesting for marine scientists, in that micro-bathymetry data can provide the relief information, which is often lost on 2D mosaics. It is indeed very difficult to visualize the relief on images of the deep-sea. However, such fusion can be done in 3D visualization programs (such as Fledermaus, CARIS HIPS and SIPS, or ArcGIS 3D Scene), for instance by using geo-referenced mosaics as ‘textures’ for the micro-bathymetry. Therefore, we did not consider implementing such capability in the LAPM tool.

Question #5: How do you handle multidimensional pixel data in images? This is related to Question #4. The off-the shelf SIFT in VLFeat requires greyscale images as input. I would assume that using the full information available in each image pixel would result in more reliable identification of matching points. Of course, registering mosaicks from spectral cameras with mosaicks from sonars is outside the scope of this manuscript. However, as the fusion of images from multiple cameras, especially with differing fields-of-views, is an active research topic in computer science, I encourage the authors to take a look at what other state-of-art algorithms could be brought to marine scientists. The authors may also want to take a look at the paper by Koen et al., where numerous SIFT variants using the full colour information are evaluated. Koen et al. also provide software for computing color-SIFT on their website (http://koen.me/research/colordescriptors/) Koen et al., "Evaluating color descriptors for object and scene recognition", IEEE Transactions on PAMI, 2010. (doi: 10.1109/TPAMI.2009.154) Response: This is an interesting point. The LAPM tool only accepts greyscale (1 pixel dimension) and RGB colour (3 pixel dimensions) images as input. In the case of RGB images, they are converted to greyscale before the SIFT detection step. However, the construction of the final mosaic is done independently for every pixel dimension. Therefore, in its current state, the construction step of the LAPM tool could handle image data with any dimension (for instance images from hyperspectral cameras). As suggested by Referee #2, some of the SIFT variants presented by (van de Sande et al., 2010) could indeed allow the computation of SIFT features from colour or hyperspectral images and improve the feature detection and matching. Furthermore, it would be interesting to test if such variants work reliably with underwater imagery. Indeed, spectral absorption in seawater varies for the different wavelengths of light and seafloor images do not render the actual colours of the seafloor. For instance, RGB seafloor images are often depleted in red colour. Presently we do not know if, without colour calibration, such artifact would impair the feature matching. Therefore, we added this to our list of future improvements for the LAPM tools.

References:

Lowe, D.G., 1999. Object recognition from local scale-invariant features, in: The Proceedings of the Seventh IEEE International Conference on Computer Vision, 1999. Presented at the The Proceedings of the Seventh IEEE International Conference on Computer Vision, 1999, Corfu, Greece, pp. 1150–1157.

Lowe, D.G., 2004. Distinctive image features from scale-invariant keypoints. Int J Comput Vision 60, 91–110.

Pizarro, O., Singh, H., 2003. Toward large-area mosaicing for underwater scientific applications. IEEE J. Oceanic Eng. 28, 651–672.

Van de Sande, K.E.A., Gevers, T., Snoek, C.G.M., 2010. Evaluating Color Descriptors for Object and Scene Recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence 32, 1582–1596.

Vedaldi, A., Fulkerson, B., 2008. VLFeat: An open and portable library of computer vision algorithms [WWW Document]. www.vlfeat.org. URL http://www.vlfeat.org/

Vedaldi, A., Fulkerson, B., 2010. VLFeat: An open and portable library of computer vision algorithms, in: Proceedings of the International Conference on Multimedia. Firenze, Italy, pp. 1469–1472.

Interactive comment on Geosci. Instrum. Method. Data Syst. Discuss., 3, 127, 2013.