Technical note: probabilistic visual and electromagnetic data fusion for robust drift-free sequential mosaicking: application to fetoscopy

Marcel Tella-Amo, Loic Peter, Dzhoshkun I. Shakir, Jan Deprest, Danail Stoyanov, et al.

Marcel Tella-Amo, Loic Peter, Dzhoshkun I. Shakir, Jan Deprest, Danail Stoyanov, Juan Eugenio Iglesias, Tom Vercauteren, Sebastien Ourselin, "Technical note: probabilistic visual and electromagnetic data fusion for robust drift-free sequential mosaicking: application to fetoscopy," Proc. SPIE 10576, Medical Imaging 2018: Image-Guided Procedures, Robotic Interventions, and Modeling, 105760I (13 March 2018); doi: 10.1117/12.2319839

SPIE.

Event: SPIE Medical Imaging, 2018, Houston, Texas, United States

Technical note: Probabilistic visual and electromagnetic data fusion for robust drift-free sequential mosaicking. Application to fetoscopy.

Marcel Tella-Amo[a, *], Loic Peter[a], Dzhoshkun I. Shakir[a], Jan Deprest[a, c], Danail Stoyanov[a], Juan Eugenio Iglesias[b], Tom Vercauteren[a, c], and Sebastien Ourselin[a]

[a] Wellcome / EPSRC Centre for Interventional and Surgical Sciences, University College London, 66-72 Gower Street, London, UK, WC1E 6AA
[b] University College London, Translational Imaging Group, CMIC, Medical Physics, 8.18, Malet Place Engineering Building, London, UK, WC1E 6B
[c]KU Leuven, Centre for Surgical Technologies, Faculty of Medicine, Herestraat 49, Leuven, Belgium, 3000

ABSTRACT

The most effective treatment for Twin-to-Twin Transfusion Syndrome is laser photocoagulation of the shared vascular anastomoses in the placenta. Vascular connections are extremely challenging to locate due to their caliber and the reduced field of view of the fetoscope. Therefore, mosaicking techniques are beneficial to expand the scene, facilitate navigation and allow vessel photocoagulation decision-making. Local vision-based mosaicking algorithms inherently drift over time due to the use of pairwise transformations. We propose the use of an electromagnetic tracker (EMT) sensor mounted at the tip of the fetoscope to obtain camera pose measurements, which we incorporate into a probabilistic framework with frame-to-frame visual information to achieve globally consistent sequential mosaics. We parametrize the problem in terms of plane and camera poses constrained by EMT measurements to enforce global consistency while leveraging pairwise image relationships in a sequential fashion through the use of Local Bundle Adjustment. We show that our approach is drift-free and performs similarly to state-of-the-art global alignment techniques like Bundle Adjustment albeit with much less computational burden. Additionally, we propose a version of Bundle Adjustment that uses EMT information. We demonstrate the robustness to EMT noise and loss of visual information and evaluate mosaics for synthetic, phantom-based and ex vivo datasets.

Keywords: Fetoscopy, laser coagulation, TTTS, sequential mosaicking, globally consistent mosaicking, electromagnetic tracking, sensor fusion

1. INTRODUCTION

Twin-to-Twin Transfusion Syndrome (TTTS) complicates 10-15% of monochorionic diamniotic (MCDA) pregnancies.[1] Monochorionic twins share a single placenta and their circulation due to the presence of inter-twin anastomes. A certain unfavorable pattern may result in an imbalance of inter-twin blood flow, leading to acute, mid-trimester TTTS. This causes overproduction of urine (hence polyhydramnios) in the recipient, while the other fetus will have oligohydramnios. TTTS can also lead to cardiac dysfunction in one or both fetuses, worsening the prognosis. If this condition is not treated, the outcome is nearly always fatal.[2]

The standard of care today is fetoscopic laser photocoagulation.[3] The procedure consists of the insertion of a fetoscope, identification, and coagulation of all visible anastomoses, and functionally disconnect the two circulations. The most limiting factor for keeping an overview of the vascular anatomy is the small field of view of the fetoscope. To address this limitation, the creation of a 2D mosaic of the placenta has been proposed[4–6] as a means of expanding the field of view by stitching the fetoscopic images to a common reference frame.

Local vision-based mosaicking algorithms make use of pairwise transformations between images to compose a mosaic. This has a fundamental limitation; since the transformation of each frame to the reference space is made

Medical Imaging 2018: Image-Guided Procedures, Robotic Interventions, and Modeling, edited by Baowei Fei, Robert J. Webster III, Proc. of SPIE Vol. 10576, 105760I · © 2018 SPIE
CCC code: 1605-7422/18/$18 · doi: 10.1117/12.2319839

dependent on all the previous pairwise registrations, any new pairwise registration error is propagated through a chain of transformations. As a result, an inevitable, progressive drift in the reference space occurs.

To address this problem, we propose the use of an Electromagnetic Tracking system (EMT) by attaching an EMT sensor[7–12] to the tip of the fetoscope. The EMT system provides measurements of the 3D pose of the sensor, which then relate to the 3D pose of the fetoscope through a pre-computed rigid Hand-eye calibration matrix.

We present a probabilistic model that uses the complementarity between the EMT and visual information to drive robustly the estimation towards globally consistent mosaics, independently of the number of frames. Additionally, we compare our algorithm with the state-of-the-art in global alignment[13,14] i.e. Bundle Adjustment, and show that we achieve a similar performance to the state-of-the-art with a much lower computational burden.

2. RELATED WORK

Mosaicking algorithms have been extensively explored in computer vision for the last two decades. Vision-based mosaicking generally relies on estimating transformations with respect to the mosaic space by chaining transformations between adjacent images.[13–15] Therefore, it inherently accumulates drift.

The state-of-the-art in terms of drift-free alignment is the well-known Bundle Adjustment.[13,14] This is a batch non-linear optimization that minimizes the reprojection residuals in all images. As an example close to our application, Atasoy et al.[16] proposed a vision-based version of Bundle Adjustment for fibroscopic video mosaicking that weights the images with the number of matched features found in each pair.

The use of an external sensor in our framework aims to constrain the global pose estimates in order to bound the drift. Similar to ours is the work of Agrawal et al.,[17] who explored the idea of using an inexpensive Global Positioning System (GPS) combined with a stereo vision system. They used a Kalman filter limiting the drift in translation for long robot trajectories. In our scenario, the fetoscope is close to the tissue and its movement follows a hand-held pattern, which implies additional complications.

In our previous work,[6] we proposed a preliminary model to reduce the drift using the EMT and visual information jointly. In the current work, we extend our framework and validate the complete elimination of the drift.

3. METHODS

3.1 The Probabilistic Model

Let $\mathcal{X}$ be the set of true camera motions and $\mathcal{Z}$ their respective EMT measurements. Additionally, let $\mathcal{P}^A$ and $\mathcal{P}^B$ be the corresponding points detected from image A to image B. We assume:

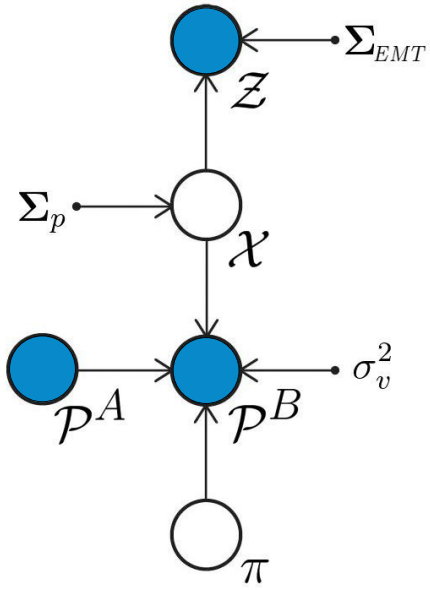
1. The imaged object to be a plane[6] π.
2. Every EMT measurement $\mathbf{z}_k$ is modeled as a Gaussian random variable centered on the true camera pose $\mathbf{x}_k$ with diagonal covariance $\mathbf{\Sigma}_{EMT}$, that is $\mathbf{z}_k|\mathbf{x}_k \sim \mathcal{N}_{\mathbf{z}_k}(\mathbf{x}_k, \mathbf{\Sigma}_{EMT})$.
3. A 2D point in image B is a Gaussian measurement generated from a true 2D point in image A, that is $\mathbf{p}^{B,i}_{l,m}|\mathbf{x}_l, \mathbf{x}_m, \pi, \mathbf{p}^{A,i}_{l,m} \sim \mathcal{N}_{\mathbf{p}^{B,i}_{l,m}}(\mu^i_v, \sigma^2_v\mathbf{I})$ in which the mean $\mu^i_v(\mathbf{x}_l, \mathbf{x}_m, \pi, \mathbf{p}^{A,i}_{l,m})$ is the projected location of the point in image A.
4. We assume that the camera is moving smoothly and therefore model the relation between camera poses with a constant velocity motion model[18] with mean $\mu_p(\mathbf{x}_{k-1}, \mathbf{x}_{k-2})$ and diagonal covariance matrix $\mathbf{\Sigma}_p$. This is $\mathbf{x}_k|\mathbf{x}_{k-1}, \mathbf{x}_{k-2} \sim \mathcal{N}_{\mathbf{x}_k}(\mu_p, \mathbf{\Sigma}_p)$. This model expresses that the velocity at time k must be the same as the velocity at time $k-1$ plus a perturbation.

The graphical model of the proposed probabilistic framework is presented in Fig. 1. Circles represent random variables, which can be either latent (white background) or observed (shaded). Parameters of the model are depicted as a point. Within this probabilistic framework, the estimation of the mosaic can be cast as a Bayesian inference problem, in which the posterior $P(\mathcal{X}, \pi|\mathcal{Z}, \mathcal{P}^A, \mathcal{P}^B)$ is maximized with respect to the camera poses $\mathcal{X}$ and plane π

$$(\hat{\mathcal{X}}, \hat{\pi}) = \underset{(\mathcal{X},\pi)}{\operatorname{argmax}}\ P(\mathcal{X}, \pi|\mathcal{Z}, \mathcal{P}^A, \mathcal{P}^B) \tag{1}$$

in which the posterior probability factorizes as

$$\begin{aligned} P(\mathcal{X}, \pi|\mathcal{Z}, \mathcal{P}^A, \mathcal{P}^B) &\propto P(\mathcal{Z}, \mathcal{P}^B|\mathcal{X}, \pi, \mathcal{P}^A) P(\mathcal{X}, \pi|\mathcal{P}^A) \\ &\propto \underbrace{P(\mathcal{Z}|\mathcal{X})}_{(EMT)} \underbrace{P(\mathcal{P}^B|\mathcal{X}, \pi, \mathcal{P}^A)}_{(visual)} \underbrace{P(\mathcal{X})P(\pi)}_{(priors)}. \end{aligned} \tag{2}$$



$\pi \triangleq$ Plane

$\mathcal{X} = \{\mathbf{x}_k\}_{k=1}^N \triangleq$ True camera poses

$\mathcal{Z} = \{\mathbf{z}_k\}_{k=1}^N \triangleq$ EMT measurements

$\mathcal{P}^A = \bigcup_{l,m\in\mathcal{L}} \mathcal{P}_{l,m}^A \triangleq$ All sets of corresponding points in images A

$\mathcal{P}^B = \bigcup_{l,m\in\mathcal{L}} \mathcal{P}_{l,m}^B \triangleq$ All sets of corresponding points in images B

$\mathbf{z}_k|\mathbf{x}_k \sim \mathcal{N}_{\mathbf{z}_k}(\mathbf{x}_k, \boldsymbol{\Sigma}_{EMT})$

$\mathbf{p}_{l,m}^{B,i}|\mathbf{x}_l, \mathbf{x}_m, \pi, \mathbf{p}_{l,m}^{A,i} \sim \mathcal{N}_{\mathbf{p}_{l,m}^{B,i}}(\mu_v, \sigma_v^2\mathbf{I})$

$\mathbf{x}_k|\mathbf{x}_{k-1}, \mathbf{x}_{k-2} \sim \mathcal{N}_{\mathbf{x}_k}(\mu_p, \boldsymbol{\Sigma}_p)$

Fig. 1. Graphical model of the proposed probabilistic framework. Circles represent random variables, which can be either latent (white background) or observed (shaded). Parameters of the model are depicted as a point.

By applying a negative logarithm to the posterior probability distribution, we can express the proposed model as the minimization of a cost which contains three terms, the visual cost C_v, the EMT cost C_{EMT} and the cost associated with the temporal model C_p as

$$(\hat{\mathcal{X}}, \hat{\pi}) = \underset{(\mathcal{X},\pi)}{\operatorname{argmin}}(C_v + C_{EMT} + C_p) \tag{3}$$

where

$$C_v = \sum_{l,m\in\mathcal{L}} \sum_{i=1}^{N_{l,m}} \frac{1}{\sigma_v^2} ||\mathbf{p}^{B,i} - \mu_v^i(\mathbf{x}_l, \mathbf{x}_m, \pi, \mathbf{p}_{l,m}^{A,i})||_2^2 \tag{4}$$

$$C_{EMT} = \sum_{k=1}^{N} (\mathbf{z}_k - \mathbf{x}_k)^T \mathbf{\Sigma}_{EMT}^{-1} (\mathbf{z}_k - \mathbf{x}_k) \tag{5}$$

$$C_p = \sum_{k=3}^{N} (\mathbf{x}_k - \mu_p(\mathbf{x}_{k-1}, \mathbf{x}_{k-2}))^T \mathbf{\Sigma}_p^{-1} (\mathbf{x}_k - \mu_p(\mathbf{x}_{k-1}, \mathbf{x}_{k-2})). \tag{6}$$

This is a large scale non-linear least squares problem, which can be solved using a Gauss-Newton method, for which the EMT measurements can be used for initialization.

The proposed algorithm is an adapted version of Bundle Adjustment that also incorporates the EMT measurements and temporal consistency of the camera motions. These algorithms require to have all the information beforehand, i.e. they are offline. This is prohibitive in our case since we aim for a sequential estimate. Consequently, we move towards local methods.

We use Local Bundle Adjustment (LBA), an approximation of Bundle Adjustment in which only the components within a temporal window of size W are considered. This reduces drastically the computational burden of the algorithm and allows for sequential estimation. We slightly modify this approach following these two main assumptions: (i) The cameras far from the current window provide little information about the new cameras to be estimated, yet they provide information about the plane given visual measurements. (ii) The cameras that have already been estimated are considered fixed in the next iteration.

4. EXPERIMENTS AND RESULTS

4.1 Algorithms, Datasets and Evaluation Metric

Algorithms: We name our proposed algorithm LBAVis+EMT. We compare it against the pairwise solution (PairVis) of the mosaicking pipeline that Brown et al.[13] proposed as initialization for a further global refinement step. We also compare it against the algorithm established as the state-of-the-art in global alignment, so-called Bundle Adjustment[15] (BAVis). However, rather than using homographies to parametrize the problem, we used the same parametrization as in LBAVis+EMT to avoid over-parametrization. We also compare LBAVis+EMT against the proposed version (BAVis+EMT) of Bundle Adjustment that incorporates the EMT information.

Datasets: We introduce a synthetic (SYN, 3370 frames), a phantom-based (PHB, 902 frames), and an ex vivo human placenta (EX, 366 frames) datasets, which are composed of a set of EMT measurements as well as a sequence of images from which correspondences have been obtained using SIFT and RANSAC.[13] The SYN dataset contains only translations and was generated by selecting image regions (368x378) of a large image, representing the imaged plane, observed by the ground truth cameras. The PHB (783x782) and EX (806x779) are hand-held datasets. The PHB is recorded by imaging a printed version of a placenta taped onto a planar surface. Example images for all datasets are shown in Fig. 2.

The PHB and EX datasets were recorded using the following setup: A camera head IMAGE1 H3-Z SPIES mounted on a 3 mm straight scope 26007 AA 0 (Karl Storz Endoskope, Tuttlingen, Germany), an EMT system NDI Aurora with a planar field generator and a Mini 6 DoF sensor.

Metrics: We parametrize a mosaic as a collection of homographies. Therefore, comparing two mosaics becomes equivalent to comparing two collections of homographies. Starting by comparing individual homographies, we define the error between any homography $\mathbf{H}$ and the ground truth homography $\mathbf{H}_{GT}$ as the mean residual error of a projected grid of points $\{\rho_i\}_{i=1}^{N_g} \in \mathbf{\Omega}_I$ from the image space $\mathbf{\Omega}_I \subset \mathbb{R}^2$ to the mosaic space $\mathbf{\Omega}_M$

$$e_j = \frac{1}{N_g} \sum_{i=1}^{N_g} ||w(\mathbf{H}_j^{-1}, \rho_i) - w(\mathbf{H}_{GT}^{-1}, \rho_i)||_2 \tag{7}$$

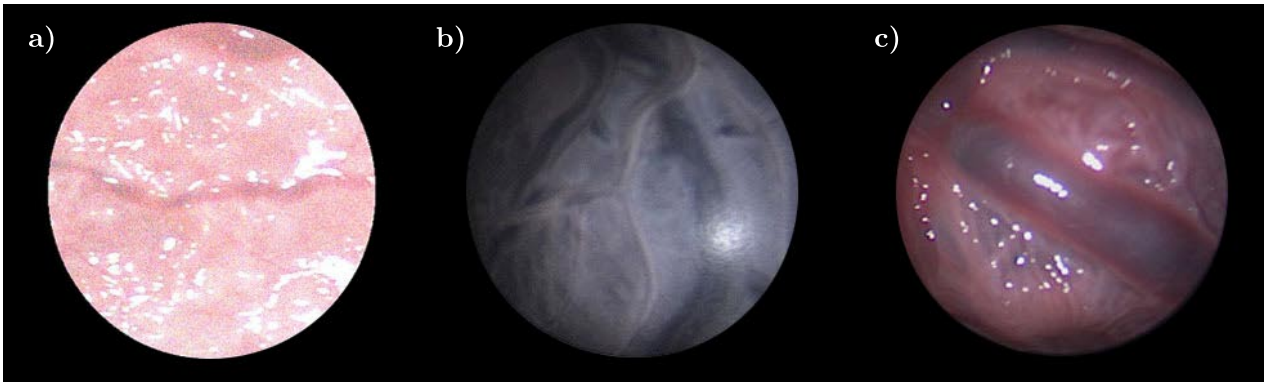


Fig. 2. a) The synthetic (SYN), b) phantom-based (PHB), and c) ex vivo human placenta (EX) datasets.

where $w(\mathbf{H}, \rho)$ projects the point ρ from the image space $\mathbf{\Omega}_I$ to the mosaic space $\mathbf{\Omega}_M$ through $\mathbf{H}$ by propagating the point and converting it to Cartesian coordinates.

To further compare two collections of homographies, we take the mean of the error associated to each of the homographies with respect to the ground truth and define the error e_M that represents the average reprojection error in pixels

$$e_M = \frac{1}{N} \sum_{j=1}^{N} e_j. \tag{8}$$

4.2 Experimental Suite

4.2.1 Drift-free mosaicking

The goal of this experiment is to show that LBAVis+EMT does not drift over time. Figure 3 shows the curves for the PairVis (blue dashed), LBAVis+EMT (green), BAVis (yellow dotted) and BAVis+EMT (green dotted) in the SYN (Fig. 3a) and PHB (Fig. 3b) datasets. The x-axis is the number of frames and the y-axis is the error e_j in pixels corresponding to a homography $\mathbf{H}_j$. Note that in the case of PairVis, this homography is created through a composition of pairwise homographies and therefore a point in the curve shows the cumulative error up to the corresponding frame. Both experiments show a similar trend; the growing tendency in the

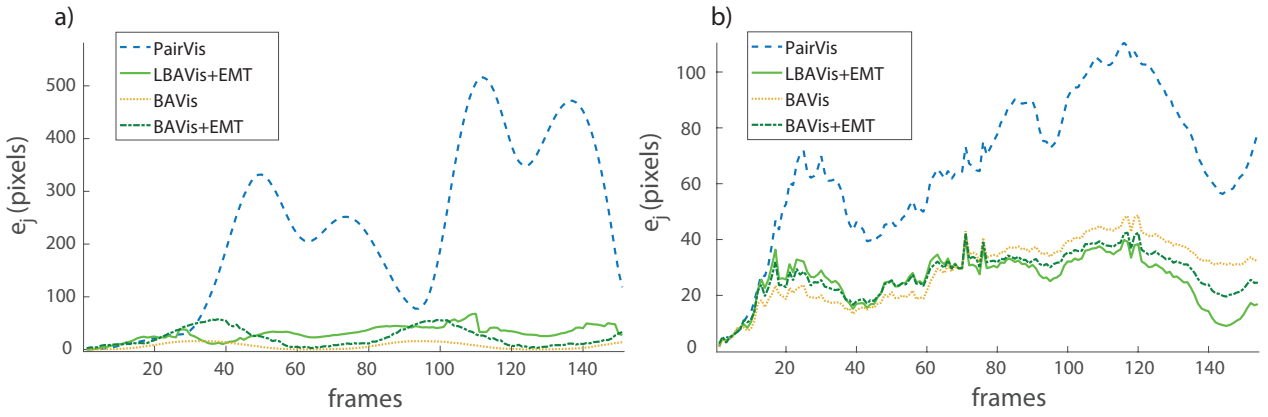


Fig. 3. Assessment of the accuracy of the PairVis (blue dashed), LBAVis+EMT (green), BAVis (yellow dotted) and BAVis+EMT (dark green dash-dotted) in a) SYN and b) PHB.

PairVis is expected due to small misregistrations in subsequent images, which leads to this long-term drift. In contrast, LBAVis+EMT maintains an approximately constant tendency over time, which demonstrates the

absence of long-term drift. Additionally, the accuracy of the proposed approach is very close to that of BAVis and BAVis+EMT. This shows the feasibility of sequential methods for mosaicking in a planar scenario in terms of accuracy, when the EMT system is guiding the estimation.

4.2.2 Sequential creation and blending of the mosaics

Figures 4a, b, and c show the results of running LBAVis+EMT in the SYN (3770 frames), PHB (902 frames), and EX (366 frames) datasets respectively. We provide videos (Video 1, MP4, 5.1MB, Video 2, MP4, 3.4MB, Video 3, MP4, 3.1MB) that illustrate the sequential creation of the mosaics by showing the subsequent blending of every new image into the mosaic image. We recommend the reader visualize the videos for a better understanding of the results. In this experiment, the goal is to demonstrate that our approach can create accurate long mosaics in a sequential fashion.

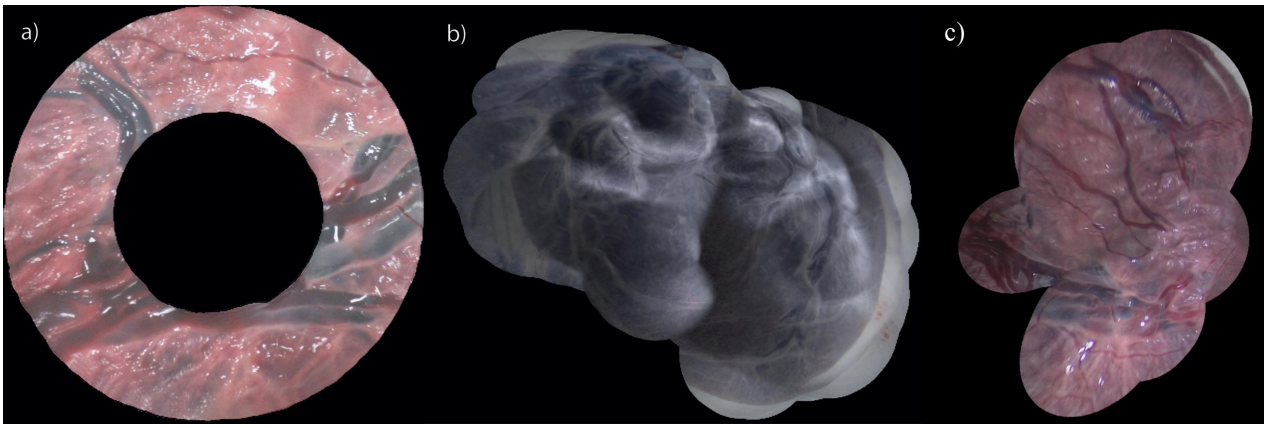


Fig. 4. Mosaic using LBAVis+EMT in the a) SYN (Video 1, MP4, 5.1MB) http://dx.doi.org/10.1117/12.2319839.1, b) PHB (Video 2, MP4, 3.4MB) http://dx.doi.org/10.1117/12.2319839.2, and c) EX (Video 3, MP4, 3.1MB) datasets. http://dx.doi.org/10.1117/12.2319839.3

5. DISCUSSION

Our results confirm that the fusion between the EMT and visual information using the proposed probabilistic model in a sequential fashion does not accumulate drift. However, there exists an error within the acceptable range of camera poses in which estimates can lay. We believe that the major cause of this error is the use of previous estimates as fixed camera poses, which encourages continuity on subsequent estimations. The effect of this error can be clearly appreciated in Fig. 4a. In contrast to the PairVis, which in Fig. 3 has drifted after approximately 20 frames, our approach is able to create a consistent mosaic after 3770 frames.

When the scene is revisited, images are not necessarily stitched in an exactly consistent location. This is a limitation of our approach, we do not constrain the revisited positions to match. However, an immediate extension that can tackle this problem would consist of the use of a spatiotemporal window to consider also regions of the space that are being revisited. While the probabilistic formulation would remain valid, the spatial window would optimize for loop closures, yet avoiding the computational cost associated with re-estimating the cameras within the loop, since we would rely on the EMT information to situate the loop roughly in a correct location from the beginning.

The placenta is not completely planar, the violation of this assumption will inevitably produce misregistration errors in those areas where the non-planarities are more prominent. Figure 4c demonstrates that these errors are small enough to consider the assumption of planarity valid in an ex vivo scenario. However, further research should be conducted on in vivo data.

6. CONCLUSIONS

We have presented a probabilistic model for robust drift-free sequential mosaicking that fuses imagery and data from an EMT system in the case where a planar or quasi-planar object is imaged in a hand-held motion. We have shown that our method does not accumulate error; a problem that affects all monocular pairwise mosaicking systems which use exclusively visual information. Therefore, we have been able to create long mosaics obtaining an accuracy comparable to the state-of-the-art Bundle Adjustment.

Given the low quality of fetoscopic images, we believe that the inclusion of the EMT system in the mosaicking process is fundamental to achieve a robust and accurate mosaic, independently of the number of frames.

Disclosures

We have nothing to disclose.

ACKNOWLEDGMENTS

This work was supported by an Innovative Engineering for Health award by Wellcome Trust [WT101957]; Engineering and Physical Sciences Research Council (EPSRC) [NS/A000027/1]. Jan Deprest is being funded by the Great Ormond Street Hospital Charity. Sebastien Ourselin receives funding from EPSRC (EP/H046410/1, EP/J020990/1, EP/K005278) and the MRC (MR/J01107X/1). Danail Stoyanov receives funding from the EPSRC (EP/N013220/1, EP/N022750/1, EP/N027078/1, NS/A000027/1), The Wellcome Trust (WT101957, 201080/Z/16/Z) and the EU-Horizon2020 project EndoVESPA (H2020-ICT-2015-688592). Juan Eugenio Iglesias receives funding from ERC Starting Grant agreement No 677697 (project BUNGEE-TOOLS). Marcel Tella-Amo is supported by the EPSRC-funded UCL Centre for Doctoral Training in Medical Imaging (EP/L016478/1). We would also like to thank Jyotirmoy Banerjee for the useful discussions as well as Efthymios Maneas for the contribution of the ex vivo placenta.

REFERENCES

[1] Lewi, L., Gucciardo, L., Van Mieghem, T., De Koninck, P., Beck, V., Medek, H., Van Schoubroeck, D., Devlieger, R., De Catte, L., and Deprest, J., "Monochorionic diamniotic twin pregnancies: natural history and risk stratification," *Fetal diagnosis and therapy* **27**(3), 121–133 (2010). [doi:10.1159/000313300].

[2] Baschat, A., Chmait, R. H., Deprest, J., Gratacós, E., Hecher, K., Kontopoulos, E., Quintero, R., Skupski, D. W., Valsky, D. V., and Ville, Y., "Twin-to-twin transfusion syndrome (ttts)," *Journal of perinatal medicine* **39**(2), 107–112 (2011).

[3] Slaghekke, F., Lopriore, E., Lewi, L., Middeldorp, J. M., van Zwet, E. W., Weingertner, A.-S., Klumper, F. J., DeKoninck, P., Devlieger, R., Kilby, M. D., et al., "Fetoscopic laser coagulation of the vascular equator versus selective coagulation for twin-to-twin transfusion syndrome: an open-label randomised controlled trial," *The Lancet* **383**(9935), 2144–2151 (2014). [doi:10.1016/S0140-6736(13)62419-8].

[4] Daga, P., Chadebecq, F., Shakir, D., Garcia Peraza Herrera, L., Tella Amo, M., Dwyer, G., David, A. L., Deprest, J., Stoyanov, D., Vercauteren, T., et al., "Real-time mosaicing of fetoscopic videos using sift," in [*Proceedings of SPIE*], **9786**, SPIE Medical Imaging (2016).

[5] Reeff, M., [*Mosaicing of endoscopic placenta images*], Hartung-Gorre (2011).

[6] Tella-Amo, M., Daga, P., Chadebecq, F., Thompson, S., Shakir, D. I., Dwyer, G., Wimalasundera, R., Deprest, J., Stoyanov, D., Vercauteren, T., et al., "A combined em and visual tracking probabilistic model for robust mosaicking: Application to fetoscopy," in [*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*], 84–92 (2016).

[7] Tran, P. T., Chang, P.-L., De Praetere, H., Maes, J., Reynaerts, D., Sloten, J. V., Stoyanov, D., and Poorten, E. V., "3d catheter shape reconstruction using electromagnetic and image sensors," *Journal of Medical Robotics Research* , 1740009 (2017).

[8] Franz, A. M., Haidegger, T., Birkfellner, W., Cleary, K., Peters, T. M., and Maier-Hein, L., "Electromagnetic tracking in medicinea review of technology, validation, and applications," *IEEE transactions on medical imaging* **33**(8), 1702–1725 (2014). [doi:10.1109/TMI.2014.2321777].

[9] Cleary, K., Zhang, H., Glossop, N., Levy, E., Wood, B., and Banovac, F., "Electromagnetic tracking for image-guided abdominal procedures: Overall system and technical issues," in [*2005 IEEE Engineering in Medicine and Biology 27th Annual Conference*], 6748–6753, IEEE (2005).

[10] Wood, B. J., Zhang, H., Durrani, A., Glossop, N., Ranjan, S., Lindisch, D., Levy, E., Banovac, F., Borgert, J., Krueger, S., et al., "Navigation with electromagnetic tracking for interventional radiology procedures: a feasibility study," *Journal of vascular and interventional radiology* **16**(4), 493–505 (2005). [doi:10.1097/01.RVI.0000148827.62296.B4].

[11] Vyas, K., Hughes, M., and Yang, G.-Z., "Electromagnetic tracking of handheld high-resolution endomicroscopy probes to assist with real-time video mosaicking," in [*SPIE BiOS*], 93040Y–93040Y, International Society for Optics and Photonics (2015).

[12] Dore, A., Smoljkic, G., Vander Poorten, E., Sette, M., Vander Sloten, J., and Yang, G.-Z., "Catheter navigation based on probabilistic fusion of electromagnetic tracking and physically-based simulation," in [*Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on*], 3806–3811, IEEE (2012).

[13] Brown, M., Lowe, D. G., et al., "Recognising panoramas.," in [*ICCV*], **3**, 1218 (2003).

[14] Szeliski, R., "Image alignment and stitching: A tutorial," *Foundations and Trends® in Computer Graphics and Vision* **2**(1), 1–104 (2006).

[15] Shum, H.-Y. and Szeliski, R., "Construction of panoramic image mosaics with global and local alignment," in [*Panoramic vision*], 227–268, Springer (2001).

[16] Atasoy, S., Noonan, D. P., Benhimane, S., Navab, N., and Yang, G.-Z., "A global approach for automatic fibroscopic video mosaicing in minimally invasive diagnosis," in [*International Conference on Medical Image Computing and Computer-Assisted Intervention*], 850–857, Springer (2008). [doi:10.1007/978-3-540-85988-8-101].

[17] Agrawal, M. and Konolige, K., "Real-time localization in outdoor environments using stereo vision and inexpensive gps," in [*Pattern Recognition, 2006. ICPR 2006. 18th International Conference on*], **3**, 1063–1068, IEEE (2006). [doi:10.1109/ICPR.2006.962].

[18] Davison, A. J., Reid, I. D., Molton, N. D., and Stasse, O., "Monoslam: Real-time single camera slam," *IEEE transactions on pattern analysis and machine intelligence* **29**(6), 1052–1067 (2007).