

SLIM (Slit Lamp Image Mosaicing)

K. Prokopetc^{1,2}

A. Bartoli²

¹ Quantel Medical, Cournon-d'Auvergne, France

² EnCoV / IP, UMR 6602 CNRS / Université Clermont Auvergne, Clermont-Ferrand, France

prokopec.kristina@mail.ru * adrien.bartoli@gmail.com

Résumé

La lampe à fente est un instrument essentiel pour les soins oculaires. Elle est utilisée dans les traitements par laser. La navigation utilise une mosaïque de la rétine, notamment pour aider au diagnostic. La construction de mosaïques précises est une tâche difficile. Les spécificités de l'ensemble d'imagerie introduisent des artefacts d'éclairage gênants. Une autre difficulté est la dérive accumulée. Ceux-ci non seulement dégradent la qualité de la mosaïque mais peuvent également affecter le diagnostic. Nous présentons une nouvelle méthodologie qui combine la réduction de la dérive et la manipulation des artefacts de réflexion dans SLIM, et améliore considérablement la qualité des mosaïques.

Mots Clef

dérive, reflet spéculaire, mouvement, rétine, mosaïque, fusion d'image, lampe à fente

Abstract

The slit lamp is an essential instrument for eye care. It is used in navigated laser treatment with retina mosaicing to assist diagnosis. The construction of seamless and accurate mosaics is an important and challenging task. Specifics of the imaging set-up introduce bothersome illumination artifacts. Another difficulty is accumulated registration drift. These not only degrade the quality of the mosaic but may also affect the diagnosis. We present a new methodology which combines drift reduction and reflection artifacts handling and significantly improves the quality of the mosaics in SLIM.

Keywords

drift, specular highlight, motion, retinal mosaicing, image blending, slit-lamp

1 Introduction

Retinal examination with the slit-lamp is the most important technique dating back to the 1980-ies which remains prevalent in clinical ophthalmology nowadays. It allows the eye to be examined with a light beam or 'slit' whose height and width can be adjusted. The slit of light, di-

rected at an appropriate angle, emphasizes the anatomic structures of the eye, allowing close inspection. It is also used for laser delivery in navigated panretinal photocoagulation (NPRP) - the standard treatment for numerous retinal diseases including diabetic retinopathy. Currently, NPRP can be performed by computer guided systems which combine real-time imaging, pre-operative planning and intra-operative navigation [21, 17, 18, 31]. Because the images captured with the slit-lamp have a narrow field of view (FOV) visualizing only thin portions of the retina, view expansion with image mosaicing became an important part of the slit-lamp based NPRP system developed at QuantelMedical, France (Fig. 1).

Obtaining a geometrically and photometrically accurate retinal mosaic is a difficult task due to the numerous challenging conditions. When performing retinal examination with a slit-lamp the imaging set-up is arranged such that the axis of the observation component is nearly coaxial with the axis of the illumination component. Both are fixed on the moving base, controlled by the ophthalmologist. The light beam is focused on the retina using a hand-held direct contact lens of strong convergence. This essential requirement unfortunately introduces bothersome illumination artifacts that populate the image and are difficult to separate from the retina. In addition, due to the 'chaining' nature of the mosaicing algorithm of the NPRP system, alignment errors tend to accumulate, causing images to drift in the mosaic. Our main objective in this work is to reduce the drift, handle illumination artifacts and obtain visually consistent mosaics.

The problem of accumulated drift was addressed in a wide variety of image registration applications [27, 14, 4, 7, 11]. The usual approach practiced by many is to perform global Bundle Adjustment (BA). In real-time systems it has been left as a post-processing step for a long time. However, in the past few years a number of real-time local BA-type refinement methods were proposed [15, 10, 6] which allow one to achieve a similar accuracy to conventional BA while reducing computational cost.

There also exist numerous methods for specular highlight removal and correction in medical and non-medical applications such as [12, 1, 5, 2, 21, 25, 22, 29, 30, 9, 28, 16, 24].

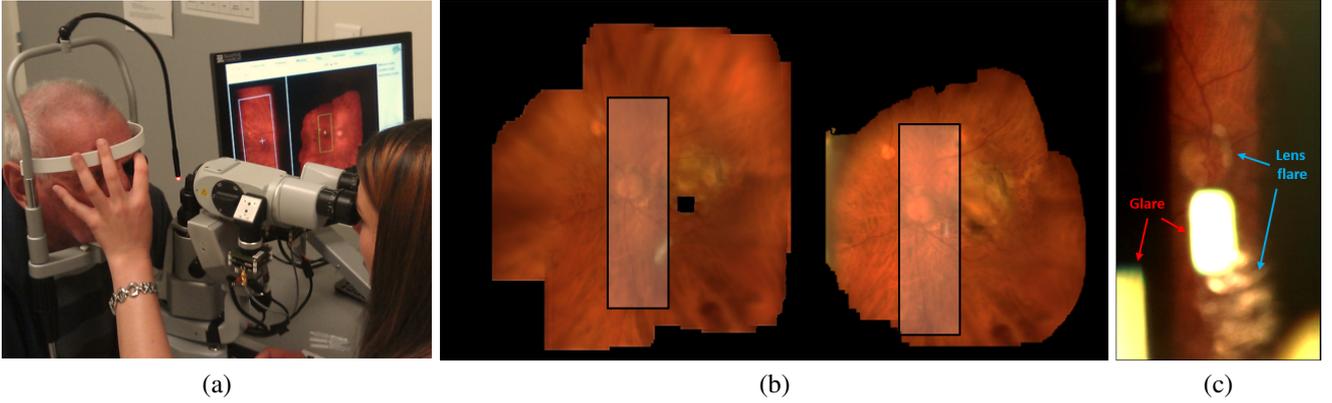


Figure 1: Retinal image mosaicing with a slit-lamp. (a) a slit-lamp NPRP system, developed at QuantelMedical, France and (b) sample mosaics constructed during an examination session with the slit-lamp, the FOV is shown as a rectangular region and (c) - a typical slit-lamp image demonstrating the appearance variation of light reflection of different origins. (Please refer to the electronic version for better visualization of all figures in this paper.)

Most of the single-image solutions are capable to correct strong glare but they share the same problem: they generally result in noticeable artifacts when applied directly in SLIM. Multi-image methods utilize the motion cues for highlight localization and correction. In SLIM, due to the specifics of the imaging set-up (see section 3.1), the apparent motion of specular highlights can be noticed but unlike in previous works, more than two consecutive observations are required for detection. Moreover, the limited FOV does not allow a frame to capture the highlight fully. Therefore, the motion cues are useful but shall be engaged as soft constraints. Learning appearance variation from multiple images has proved to outperform simpler methods [31]. However, the inability to model the complex color and intensity variation of the reflections associated with lens flare makes it unsuitable for our goals in SLIM.

In this paper we propose an improvement to [21]. We address the problem of accumulated registration drift through the creation of long-term high precision point correspondences. We associate a simple global model with local correction and key-frame based Bundle Adjustment. We also propose a two-stage solution addressing the problem of strong glares as a pre-processing step in the mosaicing pipeline, while correcting lens and haze at the image blending stage. Our main contribution is a SLIM dedicated method which handles registration drift and all types of illumination artifacts, as opposed to the methods from the state-of-the-art.

2 Methodology

2.1 Notation

- we use I to refer to the image and Latin bold to refer to image points $\mathbf{p}, \mathbf{q}, \mathbf{g}, \mathbf{c}$
- we use P to refer to the conditional probability and K is the number of components in Gaussian Mixture Model (GMM)

- w, μ, Σ are the GMM's parameters estimated with Expectation Maximization (EM)
- $\mathbf{p} \in \mathbb{R}^2$ is a vector of xy coordinates
- we use Greek character τ to refer to point tracks
- frame indexing is denoted as $f = 1, \dots, n_f$, where n_f is a total number of frames
- key-frame indexing is denoted as $k = 1, \dots, n_k$, where n_k is a total number of key-frames
- image point indexing is denoted as $i = 1, \dots, n_i$, where n_i is a total number of image points
- track indexing is denoted as $j = 1, \dots, n_j$, where n_j is a total number of tracks
- we use \mathbf{A} to note the affine transformation represented in a matrix form and w its corresponding transformation function
- the forward warping function is denoted as ω and the backward warping as ϕ respectively

2.2 Glare Removal and Retina Segmentation

Segmentation of the visible and informative retinal content from slit lamp images is a challenging task due to illumination artifacts. The work [31] was not found to be suitable for effective retina segmentation in our datasets. The concept of specular-free (SF) image widely used in the literature [25, 22, 29] is too coarse approximation. Nonetheless, it has been demonstrated to be effective for single-image glare removal. Incorporating contextual information is considered as one of the most effective approaches in many applications as was demonstrated in [19]. Retinal images obtained with a slit-lamp have a narrow FOV located in the center of the image resulting in a large part of the image containing dark pixels. This property can be used to obtain the region of interest (ROI) to reduce the processing load. Hence, our approach can be summarized in the following steps:

Step 1: pre-processing. First the image is converted to the LMS (Long, Medium, and Short light wavelengths) color space. This is commonly used color space to estimate the appearance of a pixel under a different illumination. Based on the observation that the maximum fraction of the unsaturated pixels in local patches changes smoothly we proceed with low-pass filtering similar to [29] and obtain I_{LP} . We then compute $C_{min} = \frac{\min(I_{LP})}{\max(I_{LP})}$ - the maximum chromaticity image as a pixelwise division of the minimum value over three components of I_{LP} and the mean value respectively. This computation results in a binary image, where the most glare pixels have intensity equal to 1.

Step 2: informative pixels selection. Given the priors on the location of the slit in the image we filter out highly improbable locations of the informative retinal content. For each pixel in the image we compute a conditional probability of the retinal content occurring at this pixel given the center of the image \mathbf{c}_I . We model this contextual constraint with a Gaussian Mixture Model (GMM):

$$P(\mathbf{p}|\mathbf{c}_I) = \sum_{i=1}^K w_i G(\mathbf{p} - \mathbf{c}_I; \mu_i, \Sigma_i) \quad (1)$$

where K is the number of GMM components and $\{w_i, \mu_i, \Sigma_i\}$, $i = 1, \dots, K$ are the GMM's parameters estimated with Expectation Maximization (EM). The model is learned offline on a set of annotated frames from different video sequences. Here, K was empirically tuned to represent two Gaussian components. We apply the learned GMM on a test frame and obtain a probability map.

Step 3: combination. Here we incorporate the positional prior learned in the previous step to filter out uninformative areas of the image and obtain the final segmentation of retinal content. Thus, we keep \mathbf{p} as a retinal content if $P(\mathbf{p}|\mathbf{c}_I) \geq t$, where t is a probability threshold which we empirically set to 0.6. We perform logical XOR operation with the GF image mask from *Step 1* within the estimated region. This allows us to keep only those pixels for the final result where the GF mask or the estimated region, but not both, contain a nonzero element at the same location.

2.3 Mosaicing with Drift Reduction

The solution to accumulated drift is to expand the span of tracks across the images. Our methodology relies on three main assumptions: (i) point correspondences presented in multiple views provide more constraints, (ii) a simple global motion model associated with local correction can be used to predict the track location, this may help to obtain tracks longer than short-inter-frames with improved precision, (iii) using a simple global model to initialize key-frame based local BA can be as accurate as performing global BA while being less computationally expensive. Our algorithm consists of the following parts:

Initialization. We start by obtaining a set of key-points $\{\mathbf{p}_i\}_{i=1}^{n_i}$ detected on the first frame $I_{f=1}$ and defining an

initial set of tracks $\{\tau_j\}_{j=1}^{n_j} = \{\mathbf{p}_i\}_{i=1}^{n_i}$. We also tag the first frame as a key-frame $I_{f=1} \rightarrow I_{k=1}$. Here and in the following steps all the computation and processing is performed on the segmented glare-free image obtained following the method described in section 2.2. In the experimental section we assess different types of key-point detectors, SIFT [13], the minimum eigen value algorithm (minEig) [26] and their impact on the performance of the proposed algorithm. We also use a uniform grid of points (UGrid) evenly placed on the area of the visible retina to complement the evaluation.

Motion estimation. Inter-frame motion estimation with a simple model as used in [21] seems to be robust but inaccurate, typically up to 5 pixels [17]. We can use this simple global model to create better inter-frame correspondences, and then tracks. The slit-lamp system's optics include several parts moving independently, namely the contact lens and the camera. This complicates the derivation of an accurate, simple and physically valid transformation. We use the affine transformation in our work as a best trade-off [17]. When the new frame I_f comes we estimate the motion to the last key-frame $\mathbf{A}_{f \rightarrow k-1}$ by solving a Linear Least Squares (LLS) problem where we minimize the sum of squared transfer discrepancies:

$$\min_{\theta} \sum_{i=1}^{n_i} \|\mathbf{q}_i - w(\mathbf{p}_i; \theta)\|_2^2 \quad (2)$$

The transformation function has the form $w(\mathbf{p}; \theta)$ where $\tilde{\theta}$ is an estimated (6×1) vector of affine motion parameters of the last key-frame and \mathbf{p}_i , \mathbf{q}_i are key-point correspondences from the current and the previous frames respectively.

Prediction. We propagate the existing *query* tracks τ_j using the Kanade-Lucas-Tomasi (KLT) algorithm [23] obtaining *candidate* tracks as:

$$\tau_j' = \mathcal{KLT}(\tau_j, I_{f-1}, I_f) \quad (3)$$

The key-point associated with the *candidate* track is then checked for zero-intensity (i.e. intensity values of all color channels equal to zero). If true it is then rejected as a faulty prediction because the track is considered valid only if it belongs to the visible retina. We have chosen KLT as it is an appearance based method which uses local search. It is fast and robust just enough to handle changes between consecutive frames. It copes with sudden motion better compared to statistical approaches such as the an Extended Kalman Filter (EKF), where the redundancy exists in time.

Track correction. We proceed with the refinement procedure to correct the position of the predicted *candidates* (figure 2). We first warp the new image using the previously estimated affine transformation as:

$$I_f^\omega = \omega(I_f, \mathbf{A}_{f \rightarrow k-1}) \quad (4)$$

We perform an exhaustive search in a 5×5 neighborhood w around the *query* tracks locations on the warped image I_f^ω

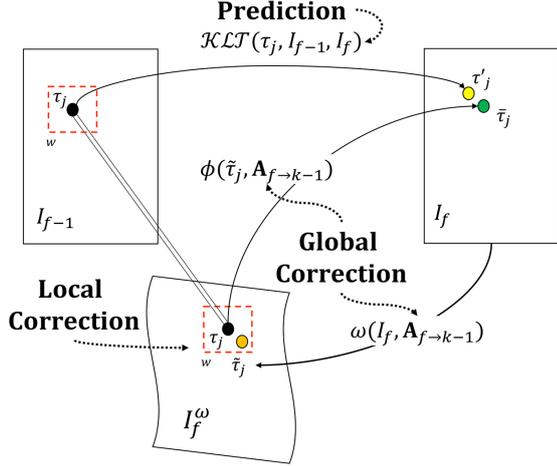


Figure 2: Schematic illustration of track prediction and correction on a sample track τ_j .

to find a possible update $\tilde{\tau}_j$ by minimizing a similarity metric. We search on the warped image because it allows us to find an estimate in a local area which is directly related to the perceived misalignment. We evaluate several metrics in this study, namely the Sum of Squared Distances (SSD), Normalized Cross Correlation (NCC) and Sum of Hamming Distances (SHD). The corrected position of the predicted track locations is computed using the previously estimated motion, where ϕ is the back-warping function:

$$\bar{\tau}_j = \phi(\tilde{\tau}_j, \mathbf{A}_{f \rightarrow k-1}) \quad (5)$$

Keyframe instantiation and local BA. We compute the tracking loss L in the current frame as the percentage of lost tracks from the last key-frame to provide the condition for inclusion of new key-points and then tracks $L = \frac{\text{sizeof}(\tau \in I_f)100}{\text{sizeof}(\tau \in I_k)}$. This does not indicate tracking re-initialization in case of full occlusion. It allows to assure that sufficiently many points are tracked at all times. Thus, if $L > 50\%$, we detect new key-points τ_f^{new} . We then filter out new tracks which fall in the predefined local neighborhood (7×7 pixels) and join the two sets of tracks. This is done to keep new tracks not too close to the existing ones and avoid populating them with redundant locations. The current frame is tagged as a new key-frame $I_f \rightarrow I_{k+1}$. We then invoke a local BA-type routine. An unknown 2D point \mathbf{g} is associated with each track $\tau_{k,j}$ and an affine transform $w(\mathbf{g}; \theta)$ with each key-frame. The presence/absence of a track in a key-frame is given by an indicator variable $v_{k,j} \in \{0, 1\}$. The reprojection error to minimize is:

$$\min_{\mathbf{g}_j, \theta} \sum_{k=1}^{n_k} \sum_{j=1}^{n_j} v_{k,j} \|\tau_{k,j} - w(\mathbf{g}; \theta)\|_2^2 \quad (6)$$

we solve this with matrix factorization in the LLS sense [8]. We repeat from the *Motion Estimation* step for the rest of the sequence.

2.4 Flare Correction and Blending

Localized flare patches in areas of uniform color and brightness in non-medical images can be easily corrected by copying parts of neighboring areas over the affected area. The situation is more complicated when the flare affects areas with lots of detail and tonal variations as retinal content. Correction is generally not possible without knowing beforehand what the affected areas should look like in the absence of flare. This requires a sophisticated per-pixel analysis in different views. Given a set of spatially aligned images we want to detect which pixels are likely to be pixels affected by lens flare. Once the lens flare regions are revealed, their visibility may be corrected by performing an adequate color mapping.

Step 1: pre-processing. Because reflection caused by lens flare has a complicated nature it is necessary to address the problem within an appropriate color space representation. Thus, for a given pixel on the mosaic $M(\mathbf{q})$, a set of overlapping frames are first transformed to the L^*a^*b color space. Following the same reasoning as described in the *step 1* of Section 2.2, we apply an image guided filter to the L component. Because the L component represents scene luminance and low-pass filtering adjusts the local intensity to its neighbors it is more likely to obtain well preserved boundaries of areas affected by lens flare.

Step 2: flare detection from color. Because regions affected by lens flare have specific colors, which are different from the rest of the retina, it motivates the use of color GMMs. We learn a simple GMM similarly to [5] offline on a set of manually annotated images where the pre-processing from the previous step was applied:

$$P(\mathbf{I}|\lambda) = \sum_{i=1}^K w_i G(\mathbf{I}; \mu_i, \Sigma_i) \quad (7)$$

with $K = 3$ Gaussian components. Here \mathbf{I} is the image pixel and $\lambda = \{w_i, \mu_i, \Sigma_i\}$, $i = 1, \dots, K$ are the GMM's parameters estimated with EM. We obtain a probability map for every L component in the observation set of frames on the mosaic using the trained GMMs. This indicates the probability that a given pixel in the observation belongs to the flared region. We use a Graph Cut algorithm [3] to mark the pixel as 'flare' or 'non flare'. As this is posed as a binary labeling problem, the Pott's Energy function is sufficient:

$$E(I) = \sum_{\mathbf{p} \in S} |I_{\mathbf{p}} - I'_{\mathbf{p}}| + \sum_{\mathbf{p}, \mathbf{q} \in N} K(\mathbf{p}, \mathbf{q}) T(I_{\mathbf{p}} \neq I_{\mathbf{q}}) \quad (8)$$

where $I = \{I_{\mathbf{p}} | \mathbf{p} \in S\}$ are the unknown 'true' labels over the set of pixels S and $I' = \{I'_{\mathbf{p}} | \mathbf{p} \in S\}$ are the observed labels. The Potts interaction is specified by $P(\mathbf{p}, \mathbf{q})$, which are the penalties for label discontinuities between adjacent pixels. The function T is an indicator function. This is optimally solved by a single execution of max-flow.

Step 3: blending. We count the number of pixels belonging to each label and identify the majority. We take the average luminance L of the majority as a L_t - *top luminance* and the average of the rest of the pixels as a L_b - *bottom luminance*. We then invoke an appropriate mapping function. This is inspired by [20]. Thus, if the majority is ‘flare’ pixels we apply ‘color burning’ - divide the inverted L_b by the L_t , and then invert the result as $C_{burn} = 1 - (1 - L_b)/L_t$. This darkens the L_t increasing the contrast. In the opposite case we apply ‘color dodging’ - divide the L_b by the inverted L_t such as $C_{dodge} = L_b/(1 - L_t)$. This lightens the L_b depending on the value of the L_t .

3 Experiments and Results

3.1 Dataset and Evaluation Strategy

The datasets used for evaluation were obtained from slit-lamp examination sessions performed on 11 different patients presented with healthy and unhealthy retinas at University Hospital of Saint-Étienne, France. The NPRP system developed at QuantelMedical was used. The videos were captured with a CCD camera at 60fps and were 2-3 minutes long. We took every 5th frame to produce image sequences to simplify the evaluation routine. In retinal imaging it is difficult to evaluate mosaicing methods objectively due to the lack of ground-truth for alignment. Here we provide an objective quantitative partial performance evaluation of the proposed drift reduction approach in two stages. First, the assessment of the steps of the method which potentially have strong influence on the result evaluated. This is followed by a comparison of the best performing combination to [21]. The proposed glare removal and retina segmentation were evaluated on a set of 270 manually annotated image frames sampled from the set of videos. This was to ensure the coverage of patient-specific and lens-specific specular highlight variation. The images were annotated with binary masks to separately assess the performance of glare removal and retina segmentation. The proposed blending technique for lens flare correction was rated on a sets of geometrically aligned video frames obtained by our mosaicing method described in section 2.3. Further details along with experimental results are provided in the following sections.

3.2 Are We Reducing Drift?

Long-term tracks is a fundamental part of BA-type refinement. The longer the track are, the better. We assessed the length of the tracks with and without the correction step of our method. To evaluate this we computed the average length of the tracks across different subsets of frames which were established each time a new key-frame was defined. Experiments have shown that tracking initialized with a UGrid and the SHD based track correction scheme provides long, consistent tracks. This gives us a solid base for BA initialization. We compare the method implemented with and without local BA to the baseline method

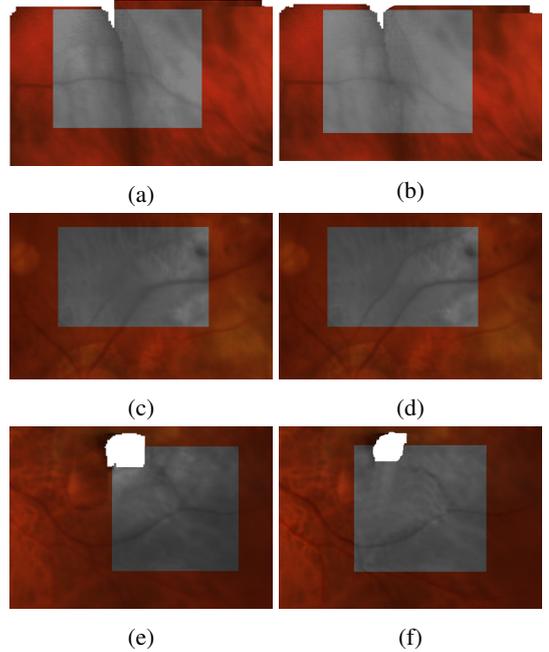


Figure 3: Examples of drift reduction results. First column - originals, second column - corrected versions.

[21]. We use a Loop Closure Error (LCE) metric [17] for this purpose. The metric shows how the composition of estimated transformations affects the global registration and accumulated drift. The idea is to initialize a uniform grid of points g_1, \dots, g_{n_l} at the first frame of the sequence and use the set of pairwise estimated transformations applied sequentially to transfer the grid through the sequence. The metric computes the discrepancy between the initial and resulting sets of points:

$$\xi_{LCE} = \sqrt{\frac{1}{l} \sum_{i=1}^{n_l} \|g_i - \zeta\|^2} \quad (9)$$

where $\zeta = w(\dots(w(g, \theta_{1,2}), \dots), \theta_{i,1})$. The comparison of our method and [21] are shown in table 1. The results show that the proposed method outperforms the baseline method. A significant improvement can be observed on the version of the proposed method where the local BA step was used.

	dataset#1	dataset#2	dataset#3	dataset#4
proposed (1)	30.43	21.75	48.02	49.12
proposed (2)	11.36	5.48	32.16	38.56
baseline [21]	34.18	28.64	48.15	50.72

Table 1: LCE across datasets. Proposed (1) - proposed method with UGrid based initialization and SHD based local correction. Proposed (2) is the proposed (1) + local BA.

A more consistent mosaicing result was achieved in figure 3. The vessel misalignment present in 3a was corrected and the vessel remains continuous in 3b. The blurred and duplicated vessels in 3c, 3e were also corrected and visual quality has been improved in 3d and 3f respectively.

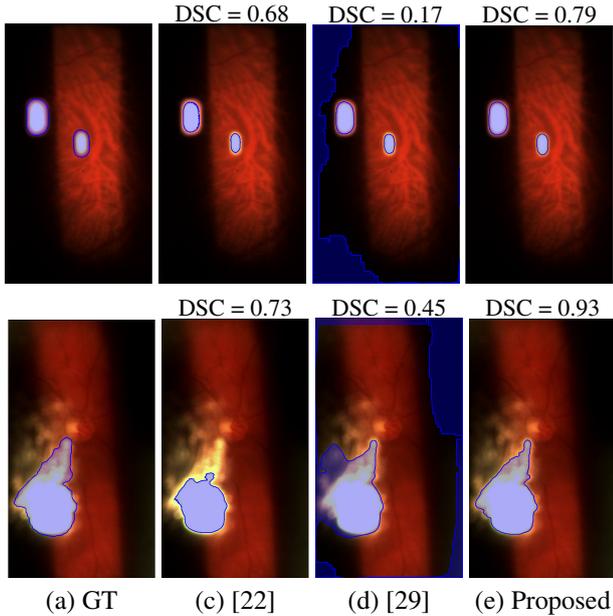


Figure 4: Glare removal results.

3.3 Is Illumination Handling Effective?

We start with the comparison of our glare removal technique with the existing methods [22, 29]. We manually annotated selected datasets by drawing the contour around regions obscured by highly saturated pixels. In simple cases, where the patient appeared to be less photosensitive and the image acquisition was not polluted by mixture of different degrees of reflections the glared region boundaries were easy to locate. Because most of the time it is difficult to observe a clear boundary between a glared region and the surrounding distorted areas, we opted for a middle-ground. The results for such two cases are shown in figure. 4. We computed the Dice Similarity Coefficient (DSC). The higher the value the more similar the algorithm’s output to the reference mask. For the simple case (figure 4(first row)) all the methods perform well while in the difficult case (figure 4(second row)) only the proposed method provides acceptable results.

We then combine the GF image with the spatial probability map to obtain the visible retinal content. The experimental results of our method compared to the simple thresholding technique used in [18] and the ML-based approach [31] as shown in figure 5. Here we also compute statistical measures for every output and average it over results on 270 annotated samples as shown in table. 2. One can see that our method provides higher values indicating better performance.

	Precision	Accuracy	Specificity	Sensitivity
Method [18]	0.30	0.70	0.58	0.86
Method [31]	0.90	0.92	0.94	0.89
Proposed	0.92	0.95	0.97	0.90

Table 2: Retinal content segmentation performance.

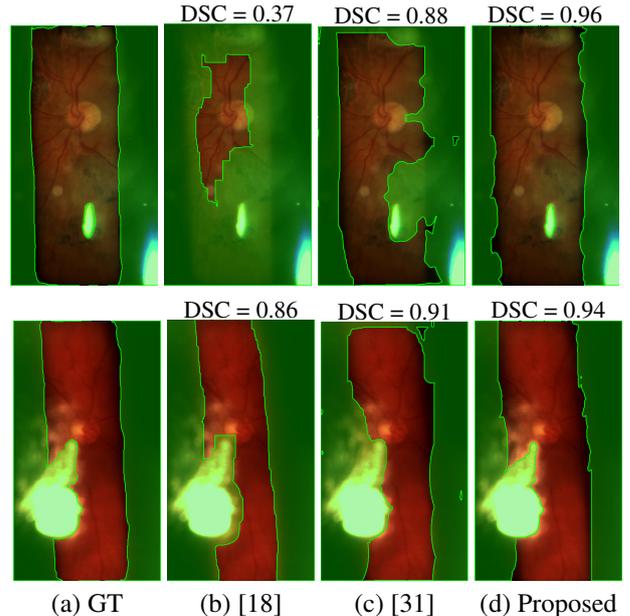


Figure 5: Comparative results of retina segmentation.

The most traditional way to evaluate the global photometric quality of slit lamp image mosaics is still based on the visual assessment of ophthalmologists. Even though the experts’ opinion is a good reference it is a subjective evaluation which may differ between experts and may prevent the mosaic from being used. Here we propose a new quantitative evaluation of the global photometric quality. We propose to use a Blending Consistency Measure (BCM). It assesses the quality of the blending by computing the standard deviation of a pixel’s intensity in the transformed image $I(\mathbf{q})$ from a set of corresponding locations in the mosaic $M_i, i = 1, 2, \dots, n$ as:

$$BCM = \sqrt{\frac{1}{n-1} \sum_{i=1}^n |I(\mathbf{q}) - \mu|^2} \quad (10)$$

$$\text{where } \mu = \frac{1}{n} \sum_{i=1}^n M_i$$

The results shown in figure. 6 (see next page) demonstrate one of the mosaics for visual assessment. We take the mosaicing result obtained by the modified version of the method in [18] where we removed the illumination correction (figure 6(a)). We then compute BCM for this uncorrected mosaic and the results obtained with the inclusion of the correction techniques from existing works in SLIM (figures 6(b), (c)) and the proposed method (figure 6(d)). The computed metric spans the range [0;255]. We show the computed results represented as a percentage value. The smaller the value, the better the blending consistency. As can be seen, the proposed method improves the global photometric quality of the mosaic and outperforms related works.

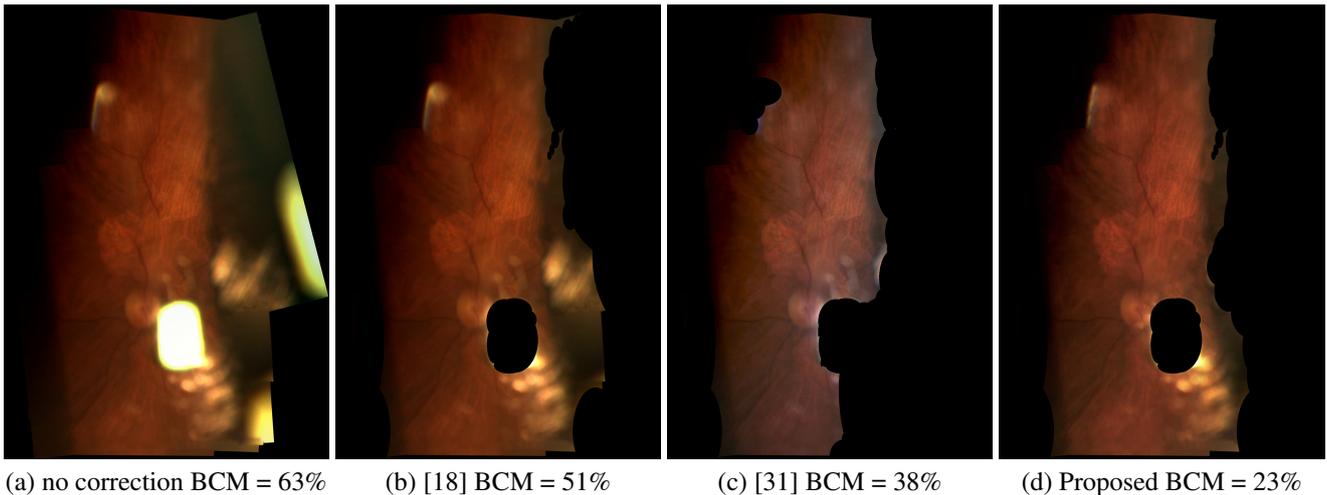


Figure 6: Comparison of mosaics constructed with the proposed method and the state-of-the-art.

4 Conclusion

In this paper we showed how to handle specular highlights of different degrees in slit lamp image mosaicing and reduce the mosaicing drift. To this end we studied several specular highlight removal and correction approaches proposed in the medical and non-medical domains and designed our own solution specifically adapted to our task. Finally, we improved previous work by proposing a fast single-image technique to remove glares and segment the visible retina using the concept of specular-free image and contextual information. Secondly, we incorporate the notion of the type of specular highlight and motion cue for intelligent image blending. Our experimental results showed that the proposed methodology exhibits a good efficiency, significantly outperforming the state-of-the-art in SLIM. We have also presented a method for drift reduction which we validated using a simple global motion model that can efficiently produce long-term tracks with better precision for long video sequences. We also demonstrated that using a grid of points distributed uniformly over the visible part of the retina generally provides a better initialization for tracking. We have proposed a new local refinement procedure which can be applied not only for mosaicing slit-lamp images but also within the scope of other applications such as object tracking in the non-medical domain. Our methodology is planned to be extended and integrated into the NPRP prototype developed at QuantelMedical.

References

- [1] Max Allan, Sébastien Ourselin, Steve Thompson, David J Hawkes, John Kelly, and Danail Stoyanov. Toward detection and localization of instruments in minimally invasive surgery. *IEEE Transactions on Biomedical Engineering*, 60(4):1050–1058, 2013.
- [2] Jane Asmuth, Bojidar Madjarov, Paul Sajda, and Jeffrey W Berger. Mosaicking and enhancement of slit lamp biomicroscopic fundus images. *British journal of ophthalmology*, 85(5):563–565, 2001.
- [3] Yuri Boykov and Marie-Pierre Jolly. Interactive organ segmentation using graph cuts. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 276–286. Springer, 2000.
- [4] Matthew Brown and David G Lowe. Automatic panoramic image stitching using invariant features. *International journal of computer vision*, 74(1):59–73, 2007.
- [5] Ajad Chhatkuli, Adrien Bartoli, Abed Malti, and Toby Collins. Live image parsing in uterine laparoscopy. In *2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI)*, pages 1263–1266. IEEE, 2014.
- [6] Javier Civera, Andrew J Davison, Juan A Magallón, and JMM Montiel. Drift-free real-time sequential mosaicing. *International Journal of Computer Vision*, 81(2):128–137, 2009.
- [7] Jari Hannuksela, Pekka Sangi, Janne Heikkilä, Xu Liu, and David Doermann. Document image mosaicing with mobile phones. In *Image Analysis and Processing, 2007. ICIAP 2007. 14th International Conference on*, pages 575–582, 2007.
- [8] R. Hartley and A. Zisserman. *Multiple view geometry in computer vision*. 2003.
- [9] Hyeonwoo Kim, Hailin Jin, Sunil Hadap, and Inso Kweon. Specular reflection separation using dark channel prior. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1460–1467, 2013.

- [10] Georg Klein and David Murray. Parallel tracking and mapping for small AR workspaces. In *Mixed and Augmented Reality, 2007. ISMAR 2007. 6th IEEE and ACM International Symposium on*, pages 225–234, 2007.
- [11] Kurt Konolige and Motilal Agrawal. FrameSLAM: From bundle adjustment to real-time visual mapping. *Robotics, IEEE Transactions on*, 24(5):1066–1077, 2008.
- [12] Holger Lange. Automatic glare removal in reflectance imagery of the uterine cervix. In *Medical Imaging*, pages 2183–2192. International Society for Optics and Photonics, 2005.
- [13] David G Lowe. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2):91–110, 2004.
- [14] Philip F McLauchlan and Allan Jaenicke. Image mosaicing using sequential bundle adjustment. *Image and Vision computing*, 20(9):751–759, 2002.
- [15] Etienne Mouragnon, Maxime Lhuillier, Michel Dhome, Fabien Dekeyser, and Patrick Sayd. Generic and real-time structure from motion using local bundle adjustment. *Image and Vision Computing*, 27(8):1178–1193, 2009.
- [16] Andreas Nussberger, Helmut Grabner, and Luc Van Gool. Robust aerial object tracking in images with lens flare. In *2015 IEEE International Conference on Robotics and Automation (ICRA)*, pages 6380–6387. IEEE, 2015.
- [17] Kristina Prokopetc and Adrien Bartoli. A comparative study of transformation models for the sequential mosaicing of long retinal sequences of slit-lamp images obtained in a closed-loop motion. *International Journal of Computer Assisted Radiology and Surgery*, pages 1–10, 2016.
- [18] Kristina Prokopetc and Adrien Bartoli. Reducing drift in mosaicing slit-lamp retinal images. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, pages 93–100. IEEE, June 2016.
- [19] Kristina Prokopetc, Toby Collins, and Adrien Bartoli. Automatic detection of the uterus and fallopian tube junctions in laparoscopic images. In *International Conference on Information Processing in Medical Imaging*, pages 552–563. Springer, 2015.
- [20] Erik Reinhard, Michael Stark, Peter Shirley, and James Ferwerda. Photographic tone reproduction for digital images. *ACM Transactions on Graphics (TOG)*, 21(3):267–276, 2002.
- [21] Rogério Richa, Rodrigo Linhares, Eros Comunello, Aldo von Wangenheim, Jean-Yves Schnitzler, Benjamin Wassmer, Claire Guillemot, Gilles Thuret, Philippe Gain, Gregory Hager, et al. Fundus image mosaicking for information augmentation in computer-assisted slit-lamp imaging. *IEEE transactions on medical imaging*, 33(6):1304–1312, 2014.
- [22] Hui-Liang Shen, Hong-Gang Zhang, Si-Jie Shao, and John H Xin. Chromaticity-based separation of reflection components in a single image. *Pattern Recognition*, 41(8):2461–2469, 2008.
- [23] Jianbo Shi and Carlo Tomasi. Good features to track. In *Computer Vision and Pattern Recognition, 1994. Proceedings CVPR'94., 1994 IEEE Computer Society Conference on*, pages 593–600, 1994.
- [24] Chao Sun, Shuaicheng Liu, Taotao Yang, Bing Zeng, Zhengning Wang, and Guanghui Liu. Automatic reflection removal using gradient intensity and motion cues. In *Proceedings of the 2016 ACM on Multimedia Conference*, pages 466–470. ACM, 2016.
- [25] Robby T Tan and Katsushi Ikeuchi. Separating reflection components of textured surfaces using a single image. *IEEE transactions on pattern analysis and machine intelligence*, 27(2):178–193, 2005.
- [26] Carlo Tomasi and Takeo Kanade. *Detection and tracking of point features*. School of Computer Science, Carnegie Mellon Univ. Pittsburgh, 1991.
- [27] Bill Triggs, Philip F McLauchlan, Richard I Hartley, and Andrew W Fitzgibbon. Bundle adjustment a modern synthesis. In *Vision algorithms: theory and practice*, pages 298–372. 1999.
- [28] Chenxue Xu, Xingzheng Wang, Haoqian Wang, and Yongbing Zhang. Accurate image specular highlight removal based on light field imaging. In *2015 Visual Communications and Image Processing (VCIP)*, pages 1–4. IEEE, 2015.
- [29] Qingxiong Yang, Shengnan Wang, and Narendra Ahuja. Real-time specular highlight removal using bilateral filtering. In *European Conference on Computer Vision*, pages 87–100. Springer, 2010.
- [30] Qingxiong Yang, Shengnan Wang, Narendra Ahuja, and Ruigang Yang. A uniform framework for estimating illumination chromaticity, correspondence, and specular reflection. *IEEE Transactions on Image Processing*, 20(1):53–63, 2011.
- [31] Sandro Zanet, Tobias Rudolph, Rogério Richa, Christoph Tappeiner, and Raphael Sznitman. Retinal slit lamp video mosaicking. *International journal of computer assisted radiology and surgery*, pages 1–7, 2016.