DirectFlow: a Robust Method for Ocular Torsion Measurement

Bruno Kozen Stahl[1], Leonardo Pavanatto Soares, Vicenzo Abichequer Sangalli,
Pedro Costa Klein, Rafael Neujahr Copstein and Márcio Sarroglia Pinho[2]
School of Technology, PUCRS
Porto Alegre, RS, Brazil
[1]bkstahl@gmail.com, [2]pinho@pucrs.br

***Abstract*—Measuring involuntary eye movement under specific stimuli is an important way to identify diseases such as balance disorders. Exams based on video-oculography (VOG) equipment are able to detect horizontal and vertical displacements of the pupil. However, detecting torsional movements is still a challenge. Although conventional methods have good accuracy, their results can be influenced by artifacts, such as a torsion center displacement, interference by illumination, reflections, and changes in the pupil dilation. We propose a novel method which improves the robustness of this measurement by applying the Lucas-Kanade Pyrm (LKP) optical flow technique to the captured image, directly over the iris, rather than making a polar transformation. Retaining this additional information allows multiple features over the iris to be analyzed individually and as a group, providing correction of the torsion center displacements, filtering features with reflections and adapting to different pupil dilations before the torsion angle is calculated. The accuracy and performance of this method were evaluated by comparing it against a conventional method when detecting torsional movements on videos with a known ground truth. Moreover, a simplified version of the proposed method is also evaluated, in order to analyze the impacts of a torsion center displacement. Results show that the proposed method has higher accuracy and equivalent performance to the conventional method.**

***Index Terms*—image processing, ocular torsion, video-oculography**

I. Introduction

The measurement of ocular movement is an important way to detect certain types of movements and obtain information for the diagnosis of many vestibular, neurological, and ophthalmological disorders [1]. These movements, called nystagmus, are rhythmic, repetitive and involuntary oscillations of a single or both eyes, in a horizontal, vertical or torsional way. Torsional eye movement is defined as a rotation performed by the eye around the line of sight [2]. Since visual observation of these movements might become impractical, video-oculography (VOG) equipment are often used to perform these measurements. VOG equipment is usually composed of one or two cameras attached to some sort of glasses, being able to record the movements in the eyes of the person wearing it.

Although the analysis of vertical and horizontal eye movement is already commercially available through video-oculography (VOG) equipment, the measurement of torsional movements still presents some challenges that can diminish the measured accuracy of torsional movements: displacements in the torsion's center, which is also the location of the pupil's center; reflections on the iris, created by illumination on the eyes; and pupil dilation, since exams that use VOG equipment are frequently performed in dark or dimmed-light environments.

We propose Direct Flow, a method that measures ocular torsional movements while being resistant to reflections, and both dilation and displacements of the pupil. Current methods face limitations under scenarios with those characteristics, as reviewed in Section II. A group of features is selected and directly tracked over the iris image using the Lucas-Kanade Pyrm (LKP) optical flow technique [3], enabling the analysis of more information than if a polar conversion was performed. A torsion angle can then be calculated for each pair of consecutive frames, and any displacement, reflection, and dilation that might have happened between them can be corrected before the torsion angle is calculated, by analyzing each of the tracked features both individually and compared to the average of features.

This article is organized as follows: Section II explores related work on ocular torsion measurement, along with their solutions and setbacks considering the existing challenges; Section III describes in details the proposed method, dividing each of the steps taken to achieve the result; Section IV defines a comparison between the proposed method and other approaches; Section V presents the results found during the experiment, along with a discussion; finally, Section VI describes the final remarks of this work.

II. Related Work

Infrared lighting is commonly used by VOG equipment to improve the quality of acquired ocular images. When the eye is subjected to this frequency of light, the pupil becomes more prominent, facilitating the location of borders and the pupil's center [4].

Lee et al. [5] proposed a method for measuring ocular torsion by tracking iris features. Initially, the method uses known image processing techniques to find both the center and size of the pupil and the iris [6], [7]. After that, a circular image of the iris is cropped from the original image. Since measuring eye torsion from this image can be a complex and time-consuming

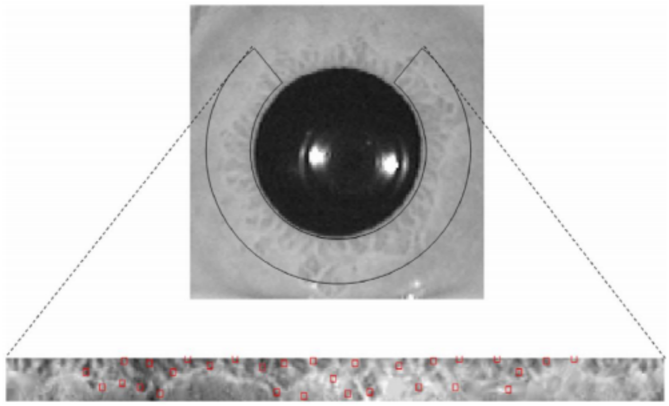

Fig. 1. Polar transformation of iris image. *The red dots represent the image features selected for tracking*

process, it is transformed into a new rectangular image. This is done through a conversion algorithm, from polar coordinates to a Cartesian plane [8], as can be seen in Figure 1.

The transformation of the iris image into a rectangle maintains the orientation of the torsional movements, converting them into horizontal displacements, which can be more easily tracked [5]. In this method, the red dots (Figure 1) are selected using Bouguest's [3] edge detection technique. These dots are the features to be tracked on the transformed image by the optical flow technique from Lucas-Kanade [9]. In addition to having its accuracy in the order of 0.15°, the method is able to process videos at up to 30 frames per second (FPS).

In order to correctly measure the torsion angle between two frames of a video, the two eye images must be rotating on the same axis. Since the eyes can move throughout the video, all of the existing methods suffer from the same problem: displacements as small as 1 pixel between the torsion axes of the images can cause errors of up to 1° [5]. The techniques used to perform image alignment have a success rate of only 80% [10].

Ong and Haslwanter [11] developed a method which also uses features of the transformed iris to measure the torsional movement of the eye, which is resistant to uneven illumination and significant changes of eye or camera position. Regarding tolerance to reflections on the iris, Ong [11] uses the technique of Maximally Stable Volumes (MSVs) [12], which performs poorly, reaching a maximum of 0.5 FPS. Regarding pupil dilation, the works of Hoshimo and Nakamomi [13] and Lee et al. [5], insert an artificial illumination with visible light to force iris contraction, which makes certain VOG exams unfeasible since they require dark or dimmed-light environments.

Otero-Millan et al. [2] proposed using template matching on the entire iris, taking into account iris occlusions and pupil by the eyelids. It also applies a polar transformation with a geometric correction, from which the eyelids are masked, and the features extracted from the iris. Then, the template matching technique is used to compare the iris pattern with a reference image. Real-time recordings at 100 Hz were measured for both eyes, with a noise of less than 0.1°.

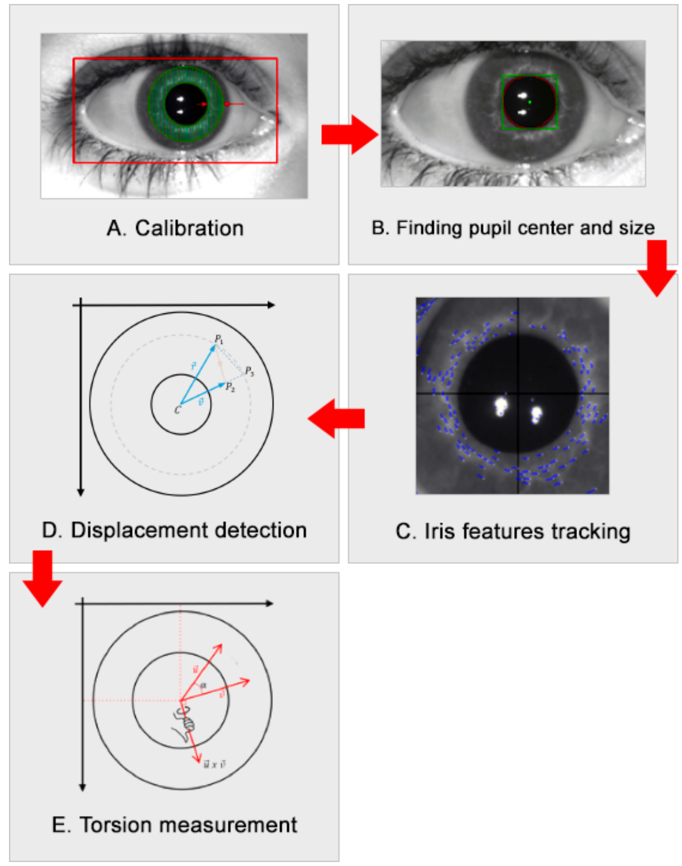

Fig. 2. Overview of each step of the proposed technique.

III. Proposed Method

This work proposes DirectFlow, a novel method that is resistant to torsion's center displacement, light's reflection and pupil's dilation, therefore allowing the measurement of ocular torsion angles with higher precision than conventional methods. This is possible because feature tracking is performed directly on the image of the iris instead of using a polar transformation, as in the work of Lee et al. [5] and Ong and Haslwanter [11]. Retaining this additional information allows multiple features over the iris to be analyzed individually and as a group, providing correction of the torsion's center displacements, filtering features with reflections and adapting to different pupil dilation before the torsion angle is calculated. Therefore, it is easier to determine if there was any problem between two sequential frames used in the tracking process. For each pair of frames (current and previous), the coordinates of the features in the current frame are corrected in order to compensate for the detected displacement. Moreover, features that are not behaving as expected, due to reflections or changes in the pupil's dilation can be disregarded. The calculation of the ocular torsion angle is only performed after the features have been filtered and corrected. The steps of the proposed method are presented in Figure 2, and further explained in the next subsections.

In the first step, a calibration process is performed to minimize problems in image acquisition and also to obtain the size of the iris. In the second step, two consecutive frames of a video are selected and the calculations of both the center position and size of the pupil are performed on both images. In the third step, the iris region is cut into a new image and iris features are found. In the fourth step, based on the movement profile of these features, possible center displacements between the two images used in the previous stage are detected. If there is any displacement, the coordinates of the second frame are corrected to compensate for translation. In the fifth, and last step, the ocular torsion angle is calculated based on the center of the pupil in the first image and the displacements measured and updated in the previous steps.

A. Manual calibration

The calibration step defines a region of interest (ROI) that will be used for tracking in the next steps. It is also where the size of the iris is obtained.

In the works of Khan et al. [14] and Daugman [15], the iris size is obtained by iris segmentation methods. However, these methods have errors that can impair the detection of angular torsion. For this reason, in this work, the iris size is defined manually by an operator and since it is performed only once, in the first frame of the video, this does not actually impact processing time. Figure 2-A shows an example picture with the ROI painted in red and the size of the iris in green.

B. Finding pupil center and size

The calculations of both the size and position of the pupil make up the second stage of the method. Finding the pupil in an image is straightforward because it is a well-defined region with a darker tone than the rest of the image. Having the pupil position and size makes it easier to find the iris and track its features.

From the ROI, a threshold based segmentation is performed to obtain an image with only the darkest areas of the eye, including the pupil (Figure 3-B). With the segmented image, the Canny [16] filter is applied to identify the edges of these areas previously found (Figure 3-C). To ensure the integrity of the pupil's contour, a morphological closing technique is applied [17](Figure 3-D). Over this image, a contour detection algorithm is applied to obtain information about the areas of the image. Then, the contour with the biggest area is selected as the pupil.

The center of the pupil is calculated from the midpoint of the contour points. The size of the pupil is obtained through the dimensions of the pupil's bounding box. Figure 2-B shows an example picture with the pupil's center and bounding box painted in green, and the contour of the pupil in red.

C. Tracking of iris features

This step consists in the application of an optical flow tracking technique in two consecutive frames of a video to measure the displacement of features in the iris region. The analysis of these displacements will allow the angular torsion angle calculation in the next steps.

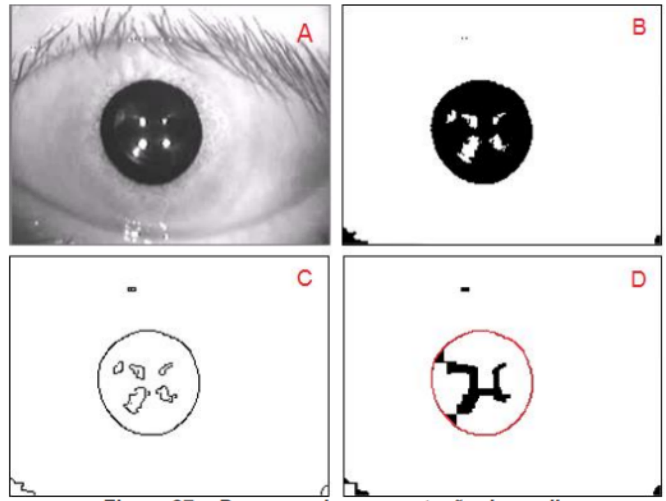

Fig. 3. Frame segmentation. *(A) Original image. (B) Threshold based segmentation. (C) Canny filter result. (D) shows the morphological closing technique*

To align the center of these images and reduce computational costs, they are cropped to fit the iris region. This region is defined as a square of side equal to the iris diameter. This cropping procedure facilitates the calculation of the angle and also enables the application of optical flow tracking techniques since both images rotate around the same axis. In case there is a pupil dilation between two frames, the bigger frame is resized to match it's neighbor, keeping pupil on the same position on both frames.

After generating the cropped images of the iris, the Lucas-Kanade Pyrm (LKP) optical flow technique [3] is applied directly on the iris image, to track the displacement of its features. Differently from other techniques [5], [11], a polar transformation is not performed in our technique. The relevant features were obtained using Shi and Tomasi method [18]. According to studies carried out by Zhang, Gao and Bakos [19], this image cropping procedure helps to minimize computational costs, since the LKP technique execution time varies according to the image resolution. Figure 2-C shows an example picture with the pupil's center positioned at the center of the image, and the tracked features' movements painted in blue.

D. Detection of pupil displacement

This step consists of detecting horizontal and vertical displacements of the pupil based on the movement of the features being tracked in the iris. This displacement is obtained by averaging the displacements of all the features that are within the iris region.

In Figure 2-D, for example, it is possible to see a feature that is moving from point P_1 to P_2. However, if the feature was only rotating around the center of the pupil, it would be expected for it to move from P_1 to P_3. This third point is obtained from a vector that has its origin in C, the same magnitude as the vector $\overrightarrow{r}$ and the same direction of the vector $\overrightarrow{v}$.

After calculating the displacement of each feature, the average of these displacements is calculated. This average is used to compensate the center displacement.

Another possible cause for these displacements is the occurrence of artifacts while tracking the features, caused by the existence of reflections. These failures, however, affects only a small number of tracked features, as they occur when there is a center displacement.

E. Measurement of torsion angle

In this last step, the method selects valid displacements of features, calculating the torsion angle and torsion direction for each displacement. The degree of eye torsion is obtained through the median of these angles.

Before measuring the torsion angle, invalid displacements are disregarded. There are three cases where this can happen: displacements that have one or more features outside the iris region are disregarded. Features that had a radius smaller than the radius of the pupil or larger than the radius of the iris; when a feature moves in the opposite direction in relation to the other features in the same pair of images. This occurs in cases of reflections caused by the illumination used in the VOG equipment; when the displacement is not rotational around the center of the pupil. These artifacts, in general, are generated by errors in the mechanisms of similarity detection among the features.

The torsion angle of each displacement is represented by α, as can be seen in Figure 2-E. It is calculated through the scalar product between two vectors that go from the center of the pupil to the position of the feature in the current frame ($\overrightarrow{u}$) and in the previous frame ($\overrightarrow{v}$).

The ocular torsion angle is obtained through the median of all displacement angles of the iris features. The average was not used, as it is very sensitive to outliers.

The torsion direction is equivalent to the direction in which most iris features are moving. To calculate the torsional direction of each feature, the cross product is calculated between the vectors $\overrightarrow{u}$ and $\overrightarrow{v}$. Depending on the direction of the resulting vector, the torsion angle will be positive (clockwise torsion) or negative (counterclockwise torsion).

IV. Experiment

An experiment was carried out to determine the accuracy and performance of the proposed method. It was compared to two other methods: PolarFlow and Direct. This section presents the process of acquiring the ocular torsion images, describing the equipment used to obtain the videos and both processes of obtaining and generating the ground truth used to evaluate the accuracy of the proposed method.

A. Dataset

The ocular torsion movements of the videos used in this work were obtained in two ways: the first one was through the process of intorsion and extorsion on the user's head; and the second one was through the rotation of the image of an eye, as it was done by Lee et al. [5].

Fig. 4. Experimental Equipment. *During tests, a camera is fixed to one of the eyes, while the other opening is kept closed.*

The first category of videos was obtained using an experimental equipment with a DMM 22BUC03-ML infrared camera, from The Imaging Source, which captures videos of 640x480 at a rate of 30 FPS. The device was mounted on a diving mask and the lenses were painted black to prevent the influence of external light, while holes were made to attach the camera and also a plastic cover. Figure 4 shows two images of this equipment.

The process of intorsion and extorsion allows generating torsional nystagmus in people without any type of pathology. This process was performed by tilting the user's head to the right and left. Three videos were created in this category, captured from three persons with different iris colors.

The second category of videos was obtained through the artificial rotation and/or translation of the image of an eye, which was published in Forensic Informatics Biometric Repository (FIB-R) [20]. This image was selected to produce this video because it has a higher resolution than the other videos used in this work. The capture equipment was not informed. In addition, the pupil will not dilate during the course of the video and its center is fixed, reducing the number of errors caused by this type of problem.

Two videos were generated for this category by performing a clockwise rotation of the image, where each frame was gradually rotated until it reached 90°. After that, it was rotated counterclockwise using the same step size. Although both videos have rotations, one of them also has displacements inserted in some of its frames to test their effect in the ocular torsion measurement. Displacements of at most 6 pixels in each axis were distributed on some of its frames.

Table I shows the details of the videos used in this study.

B. Ground Truth

An auxiliary software was created in order to define a ground truth for the first category of videos. In this program, a set of contours is found on the previous frame and projected onto the current one. If the structures found on the previous frame have moved, the contours will not match the current frame correctly. In this case, the contours are rotated by the user until their positions match the current frame. These

TABLE I

DETAILS OF THE TORSIONAL NYSTAGMUS VIDEOS. *Last column shows the video source. CAP means the video was captured, and ROT, means the video was artificially generated from the ocular image rotation.*

Video	Resolution	Frames	Iris color	Source
1	640x480	120	Light brown	CAP
2	640x480	10	Black	CAP
3	640x480	10	Green	CAP
4	752x480	80	Dark brown	ROT
5	752x480	80	Dark brown	ROT

rotations are stored and determine the torsion angle between each frame.

Although a manual definition of the torsion rotation angle can lead to errors in the ground truth, due to the difficulty to inform the angles precisely, all methods were compared to the same ground truth, indicating that the same possible error should equally affect all methods.

In the videos of the second category, where the image of an eye was rotated and translated, it was not necessary to manually measure the torsion angles in each frame, since they were known from when the video was generated.

C. Procedure

We compared our method to two other approaches: PolarFlow and Direct. The first method, PolarFlow, is based on the work of Lee, Choi and Park [5], and represents the traditional way of tracking ocular torsion by applying a polar transformation on iris region and then track features on this transformed image.

The second method, Direct, is a simplification of our method, where the displacement detection step is not performed. A comparison of the followed steps in each method can be seen in Figure 5.

All methods evaluated in this section used the same technique to calculate both the center and size of the pupil. They also used the same technique to track features through optical flow, which is the LKP algorithm [3].

V. RESULTS

This section shows the results obtained by each of the methods. The comparison parameters between the methods are an average absolute error, maximum absolute error, and speed.

The average absolute error measured in degrees, defines the average of the absolute errors made during the angular torsion measurement of all captured video frames. In this metric, Direct recorded an error of 0.39 ($\sigma = 0.03292$). DirectFlow recorded an error of 0.17 ($\sigma = 0.025$). PolarFlow recorded an error of 0.29 ($\sigma = 0.0092$).

The maximum error, also measured in degrees, define the highest error obtained during the measurement of ocular angular torsion of all captured video frames. In this metric, Direct recorded an error of 1.23 ($\sigma = 0.1792$). DirectFlow recorded an error of 0.57 ($\sigma = 0.14$). PolarFlow recorded an error of 0.78 ($\sigma = 0.094$).

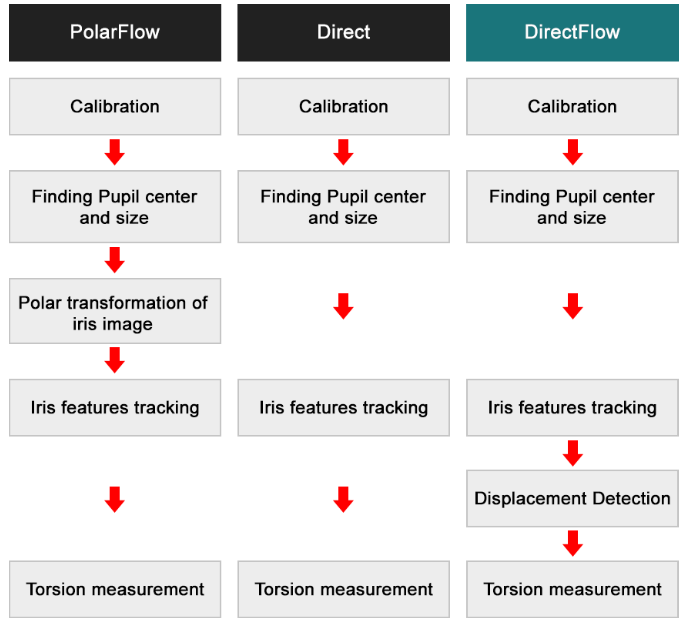

Fig. 5. Steps of compared techniques. *DirectFlow does not perform a polar transformation of the iris, which allows for a displacement detection component*

TABLE II

AVERAGE ERROR IN DIFFERENT TYPES OF VIDEOS. *Video 4 didn't have displacements. Video 5 had displacements.*

Video	Direct	DirectFlow	PolarFlow
Video 4	0.0178	0.0176	0.0503
Video 5	1.604	0.0137	0.2613

The speed, measured in Frames Per Second (FPS), accounts for the performance of all captured videos. The Direct method had a performance of 17.49 FPS ($\sigma = 1.08$). The DirectFlow method measured 16.26 FPS ($\sigma = 0.92$). The PolarFlow method measured 14.41 FPS ($\sigma = 0.79$).

During the analysis of these results, it was observed that the center displacement between frames is one of the main causes of errors in the detection of the eye torsion. As can be seen in Table II, the average error from methods Direct and PolarFlow were higher in video 5, which had artificial displacements, than video 4, which had the same artificial rotations but didn't have displacements.

To visualize the existent amount of error in a video with displacement, a histogram of errors was created from video 5, which had known artificial displacements introduced in it. The distribution of errors can be seen in Figure 6.

This analysis has shown that most errors on the proposed method are smaller than 0.05°, while the PolarFlow method has most errors between 0.05° and 1.00°. The Direct method had the worst result, with most errors higher than 1.5°.

VI. CONCLUSION

This work presented a new method to measure ocular torsion via direct optical flow. In this method, feature tracking was performed directly on the iris image without applying any

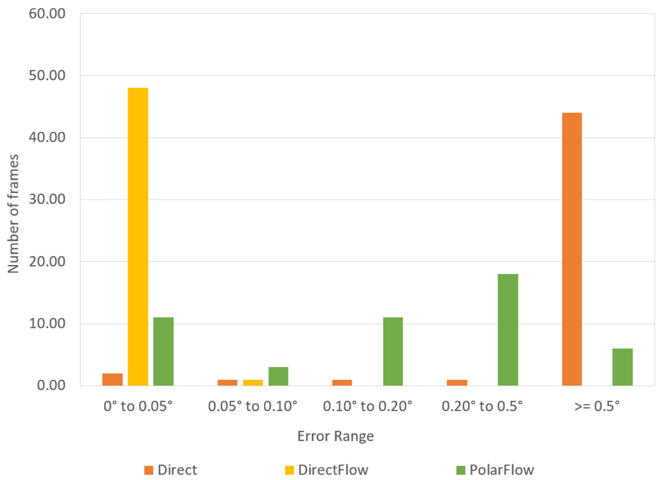

Fig. 6. Histogram of errors. *Ocular angular torsion measurement error distribution among the studied methods for video 5 (with artificial displacement)*

polar transformation to the images. This approach allowed it to determine if there was any displacement in the center of torsion along with the torsion angle calculation. Moreover, selecting multiples features over the iris image allowed for the detection of light's reflection and pupil's dilation. This enabled the measurement of ocular torsion angles with greater precision than conventional methods.

The proposed method was tested in five videos with ocular torsion movements. Two categories of videos were obtained. The first, through intorsion and extorsion processes, by using an experimental VOG equipment. In the second, two videos were generated based on a high-resolution image, with artificial rotation and/or translation being applied on each frame. One of the latest, which had rotation and translations applied to it was used to identify how much the error spreads for each method, and thus, discover how much susceptible each of them was to torsion center displacements.

In the results of the first category of videos, the proposed method was able to measure ocular torsion movements with an average error of 0.17° and a maximum error of 0.57° at a processing speed close to 17 FPS, even in the presence of reflections and variations in the diameter of the pupil. These results indicate that the proposed method has higher precision and equivalent performance to the previous methods, which presented an average error of 0.29 degrees and a maximum error of 0.78° at a processing speed close to 15 FPS. In the second category of videos, the proposed method had most of its error distribution between 0° and 0.05°, while the conventional approach had a wider error range, between 0° and 0.5°.

Therefore, the proposed method is a robust way to measure torsional movements, being resistant to torsion center displacements, interference by illumination reflections and changes in the pupil dilation.

VII. Acknowledgments

Our research is partially funded by (a) the National Institute of Science and Technology in Medicine Assisted by Scientific Computing (Grant CNPq 465586/2014-7) and (b) the Coordenação de Aperfeiçoamento de Pessoal de Nivel Superior – Brasil (CAPES) – Finance Code 001.

References

[1] R. J. Leigh and D. S. Zee, *The neurology of eye movements*. Oxford University Press, USA, 2015, vol. 90.

[2] J. Otero-Millan, D. C. Roberts, A. Lasker, D. S. Zee, and A. Kheradmand, "Knowing what the brain is seeing in three dimensions: A novel, noninvasive, sensitive, accurate, and low-noise technique for measuring ocular torsion," *Journal of vision*, vol. 15, no. 14, pp. 11–11, 2015.

[3] J.-Y. Bouguet, "Pyramidal implementation of the affine lucas kanade feature tracker description of the algorithm," *Intel Corporation*, vol. 5, no. 1-10, p. 4, 2001.

[4] B. Kunka and B. Kostek, "Non-intrusive infrared-free eye tracking method," in *Signal Processing Algorithms, Architectures, Arrangements, and Applications Conference Proceedings (SPA), 2009*. IEEE, 2009, pp. 105–109.

[5] I. Lee, B. Choi, and K. S. Park, "Robust measurement of ocular torsion using iterative lucas–kanade," *Computer methods and programs in biomedicine*, vol. 85, no. 3, pp. 238–246, 2007.

[6] J. Ivins, J. Porrill, and J. Frisby, "Deformable model of the human iris for measuring ocular torsion from video images," *IEE Proceedings-Vision, Image and Signal Processing*, vol. 145, no. 3, pp. 213–220, 1998.

[7] Y. Ebisawa, "Improved video-based eye-gaze detection method," *IEEE Transactions on instrumentation and measurement*, vol. 47, no. 4, pp. 948–955, 1998.

[8] E. Maeland, "On the comparison of interpolation methods," *IEEE transactions on medical imaging*, vol. 7, no. 3, pp. 213–217, 1988.

[9] B. D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," in *Proceedings of the 7th International Joint Conference on Artificial Intelligence - Volume 2*, ser. IJCAI'81. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1981, pp. 674–679. [Online]. Available: http://dl.acm.org/citation.cfm?id=1623264.1623280

[10] F. Timm and E. Barth, "Accurate eye centre localisation by means of gradients." *Visapp*, vol. 11, pp. 125–130, 2011.

[11] J. K. Ong and T. Haslwanter, "Measuring torsional eye movements by tracking stable iris features," *Journal of neuroscience methods*, vol. 192, no. 2, pp. 261–267, 2010.

[12] M. Donoser and H. Bischof, "3d segmentation by maximally stable volumes (msvs)," in *Pattern Recognition, 2006. ICPR 2006. 18th International Conference on*, vol. 1. IEEE, 2006, pp. 63–66.

[13] K. Hoshino and H. Nakagomi, "Measurement of rotational eye movement under blue light irradiation by tracking conjunctival blood vessel ends," in *System Integration (SII), 2013 IEEE/SICE International Symposium on*. IEEE, 2013, pp. 204–209.

[14] T. M. Khan, M. A. Khan, S. A. Malik, S. A. Khan, T. Bashir, and A. H. Dar, "Automatic localization of pupil using eccentricity and iris using gradient based method," *Optics and Lasers in Engineering*, vol. 49, no. 2, pp. 177–187, 2011.

[15] J. Daugman, "New methods in iris recognition," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 37, no. 5, pp. 1167–1175, 2007.

[16] J. Canny, "A computational approach to edge detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-8, no. 6, pp. 679–698, Nov 1986.

[17] R. Klette and A. Rosenfeld, *Digital geometry: Geometric methods for digital picture analysis*. Elsevier, 2004.

[18] J. Shi and C. Tomasi, "Good features to track," in *1994 Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, Jun 1994, pp. 593–600.

[19] F. Zhang, Y. Gao, and J. D. Bakos, "Lucas-kanade optical flow estimation on the ti c66x digital signal processor," in *High Performance Extreme Computing Conference (HPEC), 2014 IEEE*. IEEE, 2014, pp. 1–6.

[20] FIB-R. (1999) Forensic informatics biometric repository. [Online]. Available: http://fib-r.com/irises/