Cataract-101 – Video Dataset of 101 Cataract Surgeries

Klaus Schoeffmann
Alpen-Adria-Universität Klagenfurt,
Institute of Information Technology
Klagenfurt, Austria
klaus.schoeffmann@itec.aau.at

Mario Taschwer
Alpen-Adria-Universität Klagenfurt,
Institute of Information Technology
Klagenfurt, Austria
mario.taschwer@itec.aau.at

Stephanie Sarny
Klinikum Klagenfurt,
Dept. of Ophthalmology & Optometry
Medical University of Graz
Klagenfurt, Graz, Austria
stephanie.sarny@kabeg.at

Bernd Münzer
Alpen-Adria-Universität Klagenfurt,
Institute of Information Technology
Klagenfurt, Austria
bernd.muenzer@itec.aau.at

Manfred Jürgen Primus
Alpen-Adria-Universität Klagenfurt,
Institute of Information Technology
Klagenfurt, Austria
juergen.primus@itec.aau.at

Doris Putzgruber
Klinikum Klagenfurt,
Dept. of Ophthalmology & Optometry
Klagenfurt, Austria
Doris.Putzgruber-Adamitsch@kabeg.at

ABSTRACT

Cataract surgery is one of the most frequently performed microscopic surgeries in the field of ophthalmology. The goal behind this kind of surgery is to replace the human eye lense with an artificial one, an intervention that is often required due to aging. The entire surgery is performed under microscopy, but co-mounted cameras allow to record and archive the procedure. Currently, the recorded videos are used in a postoperative manner for documentation and training. An additional benefit of recording cataract videos is that they enable video analytics (i.e., manual and/or automatic video content analysis) to investigate medically relevant research questions (e.g., the cause of complications). This, however, necessitates a medical multimedia information system trained and evaluated on existing data, which is currently not publicly available. In this work we provide a public video dataset of 101 cataract surgeries that were performed by four different surgeons over a period of 9 months. These surgeons are grouped into moderately experienced and highly experienced surgeons (assistant vs. senior physicians), providing the basis for experience-based video analytics. All videos have been annotated with quasi-standardized operation phases by a senior ophthalmic surgeon.

CCS CONCEPTS

• **Information systems** → **Test collections**; • **Applied computing** → *Health informatics*;

KEYWORDS

medical data sets, surgical actions, ophthalmology, medical multimedia, multimedia information systems

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
MMSys'18, June 12–15, 2018, Amsterdam, Netherlands
© 2018 Association for Computing Machinery.
ACM ISBN 978-1-4503-5192-8/18/06…$15.00
https://doi.org/10.1145/3204949.3208137

ACM Reference Format:
Klaus Schoeffmann, Mario Taschwer, Stephanie Sarny, Bernd Münzer, Manfred Jürgen Primus, and Doris Putzgruber. 2018. Cataract-101 – Video Dataset of 101 Cataract Surgeries. In *MMSys'18: 9th ACM Multimedia Systems Conference, June 12–15, 2018, Amsterdam, Netherlands.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3204949.3208137

1 INTRODUCTION

In the field of ophthalmic surgery, interventions are performed on the human eye with tiny instruments under a microscope. Examples are retina laser treatments, refractive surgery, or cataract surgery (lens replacement). Although somehow similar to endoscopic surgeries [3], ophthalmic surgeries are radically different in terms of setup. For endoscopic surgeries, an endoscope projects a video onto a large display in the operation room, which is used by the performing surgeon to control the intervention. The endoscope, however, can also serve as a source of information to a group of trainees who watch the operation for educational purposes. This is impossible for ophthalmic surgeries by design, since the operating surgeons work with a microscope that only allows one additional trainee to follow the operation (via an additional ocular). This fact significantly hinders the teaching and training of young surgeons, which is especially unsatisfying as ophthalmic surgery is a particularly challenging surgical discipline and requires special operation techniques and psychomotor skills that need to be trained intensively.

However, ophthalmic surgeries can be recorded and revisited to be watched in full detail again with the help of additionally mounted cameras. Although a mounted camera shows a different perspective than the direct view of the ocular, it can be used for educational, scientific, and documentary purposes. More importantly, this allows the investigation of medical research questions in a postoperative fashion with the help of video content analysis and a medical multimedia information system [7]. To give an example for a medical research objective, the correlation between a specific complication and certain movements of operation instruments, or the time some instruments remain in the human eye, could be investigated. In the case of cataract surgery, more specific research questions could be addressed, such as: how much time different artificial lenses need to unfold; or, in conjunction with appropriate patient monitoring, whether specific surgical actions correlate with certain postoperative diseases, such as macular edema.

In this work, we provide a public dataset of videos recorded from 101 regular cataract surgeries (i.e., without complications) with a total duration of about 14 hours, performed by four different surgeons with different levels of experience: 56 of the surgeries are performed by two highly experienced ophthalmic senior surgeons, and 45 are performed by two ophthalmic assistant surgeons. All surgery videos have been manually annotated by a senior ophthalmic surgeon with ten quasi-standardized operation phases in cataract surgery. The dataset provides a solid basis for postoperative video content analysis of cataract surgeries in general (e.g., training video content classifiers for surgical actions in cataract surgery; evaluating operation phase segmentation of cataracts, etc.). Moreover, it gives insight into different phases of cataract surgery and allows to compare different techniques of different surgeons, as well as to evaluate the impact of the surgeons' experience level (e.g., differences in terms of movements or smoothness of actions; different phase durations; etc.). In this paper, we provide some statistics and first evaluations of such comparisons, demonstrating the usefulness of the provided dataset.

The rest of the paper is organized as follows. Section 2 gives more details on ophthalmic surgery and describes the ten different operation phases of cataract surgery. This section also gives an overview of related work on operation phase classification for cataracts. Section 3 introduces the dataset, describes the annotation format, and presents results of a first statistical analysis involving the experience level of surgeons. Section 4 concludes this dataset paper.

2 OPHTHALMIC SURGERY

2.1 Disciplines of Ophthalmic Surgery

The field of ophthalmic surgery comprises six major disciplines: first, oculoplastic, adnexal, and lacrimal surgery containing operations of the eye lid, the lacrimal duct and the orbit; second, corneal surgeries like corneal transplantation; third, cataract and refractive surgery covering phacoemulsification; fourth, surgery due to glaucoma containing also laser; fifth, retinal surgery called vitrectomy; sixth, strabismus surgery due to malposition of the eyes. In the following, we focus on modern cataract surgery, called phacoemulsification, whose purpose is to replace a dysfunctional lens with an artificial one.

2.2 Operation Phases in Cataract Surgeries

Cataract surgery can be considered a quasi-standardized procedure because it follows a consistent pattern of surgical actions described in literature [4–6]. However, since the definition of operation phases may vary in clinical practice, we adopt the following sequence of operation phases used at Klinikum Klagenfurt:

2.2.1 Incision. A sharp blade is used to create an incision through the cornea, which provides intraocular access for instruments. The paracentesis is followed by a 'clear cornea incision' which is less than 3 mm wide and is large enough to insert the phaco handpiece (see Figure 1).

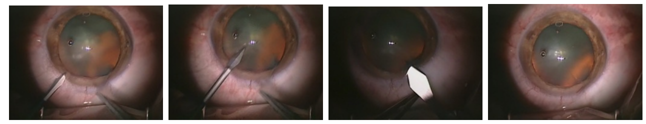

Figure 1: Different moments of *Incision.*

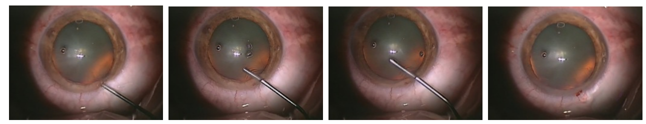

Figure 2: Different moments of *Viscous Agent Injection.*

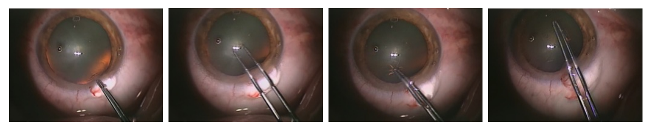

Figure 3: Different moments of *Rhexis.*

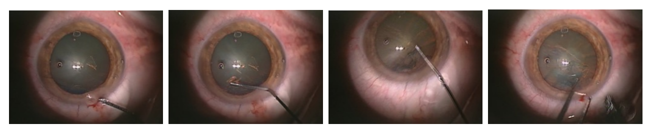

Figure 4: Different moments of *Hydrodissection.*

2.2.2 Viscous Agent Injection. Viscous agent is injected to widen the anterior chamber and to protect the corneal endothelium and the intraocular structures (see Figure 2). *Important note:* this operation phase is usually repeated before Phase 8 (lens implant setting-up), but we use the same phase identifier for both occurrences, because they are visually not distinguishable.

2.2.3 Rhexis. The anterior capsule of the lens is opened. The surgeon begins with a central radial cut. At the end of the cut, a tear is built and allows the anterior capsule to fold over itself. This tear is grasped and a flap is carried around in a circular way (see Figure 3).

2.2.4 Hydrodissection. The surgeon injects electrolyte solution and epinephrin under the rhexis to separate the peripheral cortex of the lens from the capsule. This facilitates the rotation of the nucleus and hydrates the peripheral cortex (see Figure 4).

2.2.5 Phacoemulsificiation. With ultrasound power, the phaco tip emulsifies the anterior central cortex. A deep central linear groove through the nucleus is made and the lens is cracked into two parts. The lens is rotated and chopped into pieces, which can be emulsified. During this procedure, it is essential to keep the posterior capsule intact (see Figure 5).

2.2.6 Irrigation and Aspiration. Remaining parts of the cortex are extracted (see Figure 6).

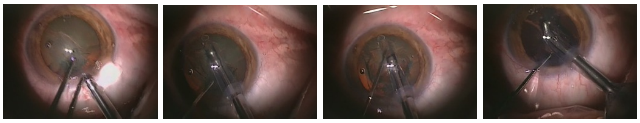

Figure 5: Different moments of *Phacoemulsification*.

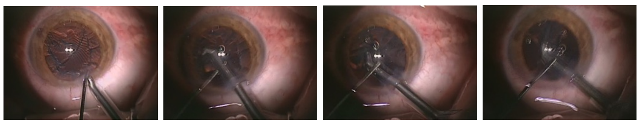

Figure 6: Different moments of *Irrigation and Aspiration*.

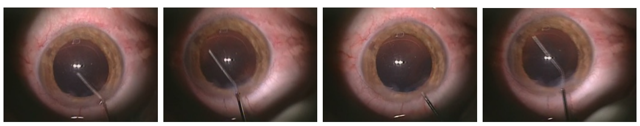

Figure 7: Different moments of *Capsule Polishing*.

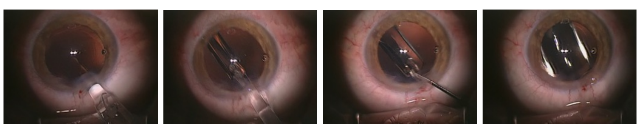

Figure 8: Different moments of *Lens Implant Setting-Up*.

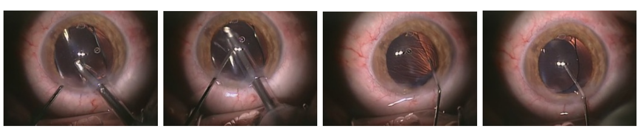

Figure 9: Different moments of *Viscous Agent Removal*.

2.2.7 Capsule Polishing. The posterior capsule is polished in order to avoid opacification of the capsule (see Figure 7).

2.2.8 Lens Implant Setting-Up. The folded artificial lens is inserted. The lens is slowly unfolding and is pushed into the capsular bag (see Figure 8).

2.2.9 Viscous Agent Removal. Viscous elastic agent is removed from anterior chamber and capsule bag (see Figure 9).

2.2.10 Tonifying and Antibiotics. The corneal incision is hydrated with electrolyte solution and antibiotics are injected. This induces temporary stromal swelling and closure of incision. Only if it leaks, a suture is required (see Figure 10).

2.3 Postoperative Video Content Analysis of Cataract Surgeries

The literature on postoperative video content analysis of cataract surgeries is currently rather sparse. Only a few methods have been proposed so far—mainly for operation phase segmentation and surgical action detection, which we summarize in the following. Lalys et al. [2] used visual information such as color, texture, and shape for the classification of surgical tasks in cataract surgery, by employing support vector machines (SVM) as machine learning approach. The classified images were aligned to already annotated recordings using a hidden markov model (HMM) and dynamic time warping (DTW). Charriere et al. [1] used a Bayesian network and two conditional random fields for classification of operation phases in cataract surgery videos. Quellec et al. [5] introduced a method that divides cataract surgeries into ten phases. Each phase is divided into an action phase, where the surgical task is performed, and an idle phase, in which basically nothing happens in the operation area, because the next step of the procedure is prepared outside of the microscope's view (or instruments are exchanged). Recorded videos were used to learn the differences between action phases and idle phases. Furthermore, a conditional random field was employed to align phases of new videos to existing ones. While these works build on the sequential order of the surgical workflow to detect phase transitions, we recently used classification of single frames in cataract surgery videos using convolutional neural networks (CNNs) [4]. The major advantage of this approach is that it is much more flexible, because it can be easily extended to additional classes pertaining to optional operation phases, out-of-order phases, or complications.

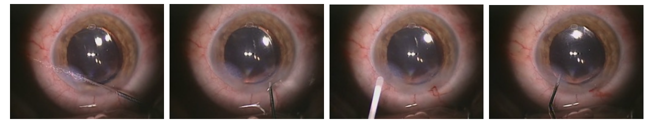

Figure 10: Different moments of *Tonifying and Antibiotics*.

3 DATASET DETAILS

The dataset consists of videos recordings from 101 cataract operations that were performed by four different surgeons at the Department for Ophthalmology and Optometry at the largest public hospital in Klagenfurt (Klinikum Klagenfurt), Austria. These videos were collected over a period of nine month and annotated by a senior ophthalmic surgeon with the ten surgical phases of cataract operations described in Section 2.2. The four operating surgeons are grouped into two different levels of experience, according to the total number of performed surgeries and their position. Two of them are moderately experienced assistant surgeons (level 1), the other two are highly experienced senior surgeons (level 2).

In total, video recording of the provided dataset amount to a duration of 14 hours, 2 minutes, and 5 seconds (1,263,116 frames). All videos have PAL resolution (720x540 pixels) and are encoded as MP4 files, using H.264/AVC with profile High as video codec (25 fps, about 1.25 MBit/s birate), resulting in a total storage requirement of 8.4 GiB. The dataset is available at the following URL: https://doi.org/10.5281/zenodo.1220951

3.1 Annotation Format

The annotation of the 101 videos consists of three CSV files (semicolon-separated values): `annotations.csv`, `phases.csv`, and `videos.csv`.

	Surgeon 1	2	3	4
Total Operations	25	24	32	20
Total Frames	465193	212250	252278	333395
Avg. Length/Op (s)	310	147	131	277
Total Phases	304	270	360	239
Avg. Phases/Op	12.16	11.25	11.25	11.95
Experience Level	1	2	2	1

Table 1: Dataset Statistics Grouped by Surgeon

The first file consists of 1266 phase annotations, for each of which the *video ID*, the *starting frame number* of this phase, and the corresponding the *phase ID* are given (MP4 files use the same video ID as part of its name). The second file maps the phase IDs to *phase names*, according to Section 2.2. Finally, the third file gives more information to each video: the number of frames in the video (operation duration), frame rate of the video, the performing surgeon (ID) of the intervention and the surgeon's level of experience (1 is lower than 2).

3.2 The Impact of the Surgeon's Experience

As shown in Table 1, the expert surgeons performed 56 (24+32) cataract surgeries and the assistant surgeons 45 (25+20). All these operations have a varying average duration (from 131s to 310s), which obviously correlates with the level of experience. We can also see that the average number of phases per operation is higher than 10 (i.e., the number of phases described in Section 2.2) and differs among the surgeons, which is grounded in the fact that some phases need to be repeated. In particular, as noted in Section 2.2, Phase 2 (viscous agent injection) occurs twice in every surgery. Moreover, the average numbers of phases per surgery indicate that less experienced surgeons tend to repeat operation phases more frequently than experienced ones.

Figures 11 and 12 show boxplots for the duration of operation phases (measured in number of video frames) for highly experienced surgeons (level 2) and moderately experienced surgeons (level 1), respectively. These figures confirm conclusions that can be drawn already from Table 1: less experienced surgeons consistently needed roughly twice as long as experienced surgeons to carry out each operation phase, and the variance in duration for a given operation phase is also consistently larger.

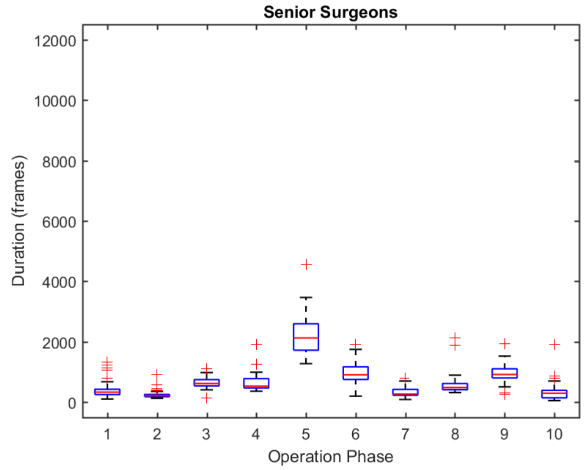

Figure 11: Boxplot of duration of operation phases when performed by highly experienced surgeons.

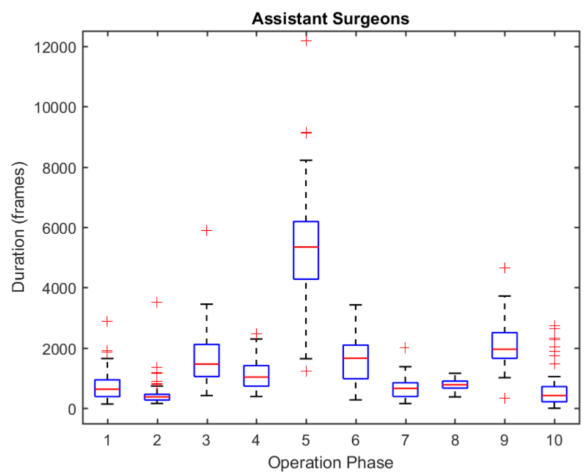

Figure 12: Boxplot of duration of operation phases when performed by moderately experienced surgeons.

4 CONCLUSION

We present a dataset of videos recorded during cataract surgeries. The dataset consists of 101 videos performed by four different surgeons (two highly experienced senior surgeons and two less experienced assistant surgeons). It is annotated with ground truth of ten quasi-standardized operation phases typically performed for such operations (without serious complications). The dataset can be used as a source for automatic video content analysis, such as investigation of specific phases, instruments, or movements—or training content classifiers (e.g., to automatically detect operation phases and segment videos accordingly). Furthermore, it can serve as a public dataset to be used by international researchers working on medical multimedia systems for evaluating and comparing their methods.

ACKNOWLEDGMENTS

This work is supported by Universität Klagenfurt and Lakeside Labs GmbH, Klagenfurt, Austria and funding from the European Regional Development Fund and the Carinthian Economic Promotion Fund (KWF) under grant KWF 20214 u. 3520/26336/38165.

REFERENCES

[1] Katia Charrière, Gwénolé Quellec, Mathieu Lamard, David Martiano, Guy Cazuguel, Gouenou Coatrieux, and Béatrice Cochener. 2017. Real-time analysis of cataract surgery videos using statistical models. *Multimedia Tools and*

Applications (23 May 2017). https://doi.org/10.1007/s11042-017-4793-8

[2] Florent Lalys and Pierre Jannin. 2014. Surgical process modelling: a review. *International Journal of Computer Assisted Radiology and Surgery* 9, 3 (2014), 495–511. https://doi.org/10.1007/s11548-013-0940-5

[3] Bernd Münzer, Klaus Schoeffmann, and Laszlo Böszörmenyi. 2017. Content-based processing and analysis of endoscopic images and videos: A survey. *Multimedia Tools and Applications* (jan 2017). https://doi.org/10.1007/s11042-016-4219-z

[4] Manfred Jüergen Primus, Doris Putzgruber-Adamitsch, Mario Taschwer, Bernd Münzer, Yosuf El-Shabrawi, Laszlo Böszörmenyi, and Klaus Schoeffmann. 2018. Frame-Based Classification of Operation Phases in Cataract Surgery Videos. In *MultiMedia Modeling*, Klaus Schoeffmann, Thanarat H. Chalidabhongse, Chong Wah Ngo, Supavadee Aramvith, Noel E. O'Connor, Yo-Sung Ho, Moncef Gabbouj, and Ahmed Elgammal (Eds.). Springer International Publishing, Cham, 241–253.

[5] Gwénolé Quellec, Katia Charrière, Mathieu Lamard, Zakarya Droueche, Christian Roux, Béatrice Cochener, and Guy Cazuguel. 2014. Real-time recognition of surgical tasks in eye surgery videos. *Medical Image Analysis* 18, 3 (2014), 579–590. https://doi.org/10.1016/j.media.2014.02.007

[6] G. Quellec, M. Lamard, B. Cochener, and G. Cazuguel. 2014. Real-Time Segmentation and Recognition of Surgical Tasks in Cataract Surgery Videos. *IEEE Transactions on Medical Imaging* 33, 12 (Dec. 2014), 2352–2360. https://doi.org/10.1109/TMI.2014.2340473

[7] Klaus Schoeffmann, Bernd Münzer, Michael Riegler, and Pål Halvorsen. 2017. Medical Multimedia Information Systems (MMIS). In *Proceedings of the 2017 ACM on Multimedia Conference (MM '17)*. ACM, New York, NY, USA, 1957–1958. https://doi.org/10.1145/3123266.3130142