Published in final edited form as:

Ophthalmic Surg Lasers Imaging. 2010 ; 41(1): 60. doi:10.3928/15428877-20091230-11.

Stability of 35 mm Scanners as Used in Ophthalmologic Research

Karl Jensen, MBA, Kristine E. Lee, MS, Michael D. Knudtson, MS, Ronald Klein, MD, MPH, and Barbara E.K. Klein, MD, MPH

Department of Ophthalmology and Visual Sciences, University of Wisconsin School of Medicine and Public Health, Madison, WI

Abstract

Background and Objective—To assess the consistency of digitization of 35mm slides, as practiced in ophthalmologic research and estimate the impact of variation on semi-automated retinal vessel width measurements.

Study Design and Methods—A single retina slide was repeatedly digitized under various conditions on three scanner models. Average color levels were extracted from the resulting images, from which vessel widths were graded. The color channel level variations and possible correlation with width were analyzed.

Results—The Nikon 5000 scanner had average coefficients of variation (CV) of 0.4, 2.3, and 0.5 for the red, green, and blue channel levels across all runs. The p-values of the correlation between the red, green, and blue color channel levels, and the width of the large retinal arteriole, were 0.89, 0.27, and 0.58.

Conclusion—Our results suggest that the tested scanners digitize the 35mm slides in a reliable manner without biasing the retinal vessel measurements.

Keywords

general ophthalmology

The 35 mm color slide has been used to document eye disease in ophthalmic practice and research for many years.^{1–3} Research centers such as those at the University of Wisconsin have accumulated large libraries of color slides, representing longitudinal views of the retinas and lenses of study participants in epidemiological studies and clinical trials as well as patients receiving clinical care for retinal disease. More recently, digital fundus photography has become a popular option,⁴ paralleling developments in the mainstream photography world.

In many cases, an image originally acquired on film is used as input to an automated or semi-automated process, necessitating digitization. At the University of Wisconsin, images originally acquired on film, in some cases from longitudinal studies spanning decades, are digitized for use in semi-automated assessments of retinal vessel diameters. Thirty-five millimeter color slides are digitized using general-purpose flatbed scanners with 35 mm adapters or dedicated 35 mm scanners. The latter are generally recommended for journalistic and scientific 35 mm conversion due to their greater spatial resolution and more consistent quality.⁵ The dedicated 35 mm film scanners have been widely used in projects where retinal or lens images were originally acquired on film and then digitized.^{6–9} At the University of

Correspondence to: Karl Jensen, MBA, UW-Madison Department of Ophthalmology & Visual Sciences, 610 North Walnut Street, 442 WARF, Madison, WI 53726-2336, 608-263-9369 Fax 608-263-0279, jensen@epi.ophth.wisc.edu.

No conflicting relationship exists for all authors for this manuscript.

Wisconsin, evaluation of the retinal vessel diameters is commonly performed on these images. An underlying assumption is that the digital scanning process is stable, that is, variations in light sensitivity of the 35 mm scanner do not affect the scientific results. The question of potential variations in the sensitivity of the 35 mm scanner has not been widely addressed in the vision research literature, although related issues have been discussed in the radiology field.^{10–12}

Anecdotal evidence of instability of digital scanners led us to design an experiment where a color slide was repeatedly scanned and the resulting images compared. We used popular models of 35 mm scanners and assessed the impact of this variation on retinal vascular caliber gradings.

Methods

A given Early Treatment of Diabetic Retinopathy (ETDRS) Field 1 slide, shot on Kodak Ektachrome 100 Plus film and centered on the optic disc, was scanned in four runs of 20 instances each on three different scanner models. The slide was of a healthy eye, of one of the authors, and of good photographic quality.

The scanners used were the Nikon LS-2000 (Nikon Corp, Tokyo, Japan), the Nikon 5000, and the Polaroid 4000 (Polaroid Corp, Waltham, MA). For the Nikon LS-2000, the Nikon Scan 2.5 software was used and the images were scanned at 2700 dots per inch (dpi). For the Nikon 5000, the Nikon Scan 4.0 software was used and the images were scanned at 2540 dpi. For the Polaroid 4000, the Polaroid Polacolor Insight 5.5 software was used and the images scanned at 3000 dpi. Images were acquired at 8 bits per channel and saved in the Tagged Image File Format (TIFF) format.

The four runs were organized into sessions of two cold runs and two warm runs. Cold runs were defined as sessions where equipment was turned on for more than 30 minutes and less than four hours, and warm runs were defined as sessions where equipment was on for more than four hours. A Matlab (The Mathworks, Natick, MA) software program averaged the pixel color values across the entire image separately for the red, green, and blue channels. The result was three numbers for each image. The three numbers characterized the image and were used roughly to quantify the differentness of the images. For our purposes, we looked for a tool that would identify cases of scanner drift and inconsistency. We had received reports of dramatic increase in scanner sensitivity in one of the channels, particularly the red channel, so we focused on color channel levels. The numerical results showed whether or not the results of two scans are identical, and modeled the level of difference. The resulting 240 images were processed using custom-built Matlab routines to determine the average pixel brightness of the red, green, and blue channels. The analysis looked at differences between the warm and cold runs, as well as differences within the runs, to determine if the variations were due to temperature. Vessels within each image were measured three times using the Interactive Vessel ANalysis (IVAN) software¹³ and three variations of the IVAN grading protocol (The IVAN Grading Protocol, University of Wisconsin Ocular Epidemiology Reading Center internal document, 2006).

The IVAN grading protocol specifies the procedure for choosing the vessels to grade for width, where to measure them, and how to handle issues such as branching of the vessels within the evaluation zone. It delineates methods to adjust mis-readings by the automated process. For this experiment, we chose to focus on four specific retinal vessels; two arterioles and two venules. Three variations of the standard protocol were developed.

In the automated approach (Phase 1), the only grader intervention was to guide the IVAN software to evaluate the pre-selected vessels. Then, using a semi-automated approach (Phase 2), grader intervention included correction, if needed, of the automated placement of the grid over the optic nerve. The third phase used the standard IVAN grading protocol to a greater

degree, in that the grader was permitted to truncate and otherwise modify the automated grading. The IVAN software attempts to find the optic nerve head and center the grid on it; occasionally this automatic centration fails and is manually corrected. Under our modified versions of the grading protocol, in Phase 1 we noted, but did not correct, the mis-centration. In Phase 2 and 3, we noted and corrected these errors.

Statistical analysis was performed with the SAS System v 9.1.3 (SAS Institute, Cary, NC). The mean, standard deviation (SD), and coefficients of variation (CV) of the color channel levels were found using procedure MEANS, which performs simple statistics. The analysis of the effect of the warm-up period and the search for correlation between color channels and vessel caliber grading were performed using ANOVA (ANalysis of VAriance) in procedure Generalized Linear Model (GLM). ANOVA is robust to unbalanced data sets. For purposes of analysis, the luminance levels for each channel were stratified into low, medium, and high tertiles of equal size by scanner and these tertiles were regressed against the vessel widths.

Results

The Nikon 5000 scanner had the lowest CV for the blue and red channels, averaging each color channel across all runs (Table 1.). The Nikon LS-2000 scanner had the lowest CV for the green runs (Table 2). The SD for the Nikon scanners ranged from 0.43 to 1.21 in the various channels, with higher SD's for the Polaroid scanner, ranging from 6.73 in the green channel to 9.01 in the red channel. All channel intensity levels are based on eight-bit pixels and thus potentially can range from 0 to 255. Due to scanning error, the data for run 1 (cold), of the Nikon LS-2000 was discarded.

Analysis of variance was used to compare the warm and cold runs by color channel. The data were missing for one Nikon LS-2000 cold run. Significant differences were found in the color channel levels, in most cases, between the two cold runs and between the two warm runs, as well as between the warm and cold runs, suggesting that the overall variability, while of interest, did not relate to temperature.

The vessel caliber measurements were regressed against the categorized luminance values of the three channels among the three scanners, with all runs combined. Figures 1–4 show scatter plots of the vessel widths, for the four measured vessels, against the color channel levels for the Nikon 5000 scanner. Here for simplicity we show only the results from Phase 3 of the grading protocol, and for the Nikon 5000 scanner. In these figures the color channel levels are normalized to the average value for that scanner and color. Regression lines are displayed. The measured vessel widths seem to vary by around three microns, in the bulk of the cases, and the color channel levels vary by about five brightness levels. The data show no significant relation between the color channel level and the vessel width, except for some relation between the small arteriole and the green channel which rises to the level of significance.

In 3 of 80 scans from the Nikon 5000 scanner, the software incorrectly centered the grid to a large degree, necessitating re-centration of the grid by the grader. Notwithstanding the occasional incorrect centration, we found that the average CV of the measured vessel widths was quite small; for the Nikon 5000 scans, the CV was 0.3 for the large arteriole and 0.4 for the large venule.

Discussion

This study grew out of reported inconsistencies with the 35 mm scanners extensively used at the Ocular Epidemiology Reading Center. We decided to examine the reliability of the scanning process, investigating the physical consistency of the resulting images as well as the impact of any variation on our computer-assisted retinal vessel caliber grading. Our overall test was to

scan the same slide repeatedly. Ideally, the result of these repeated scans would be a series of identical image files. Because the scanner and associated personal computers are electronic-mechanical devices, we expected some variation and found some in color channel levels. Interesting patterns were shown, and may bear further investigation. However, they did not appear to correlate with the warm-up period, and they did not significantly affect the vessel width measurements.

A review of the literature did not find extensive coverage of this topic in vision research publications, although many projects have digitized 35 mm slides in the course of their activity (Ferrier NJ. Automated identification of the anatomical features in slit lamp photographs of the lens. IOVS 2002;43:E-Abstract 435).^{14–17} In 1991, Cideciyan, Nagel and Jacobson developed a mathematical analysis and model of the overall noise produced by the process of retinal film photography and digital scanning in the course of research into retinitis pigmentosa.¹⁸ They described the mechanical operation of the Nikon LS3500 scanner and developed an image processing technique to restore detail lost in the process. In 2000, Wilson, Nemeth, Edwards and Soliz attempted to find the best set of scanning parameters to support human recognition and grading of retinal features such as vessels, cup and disc, and drusen.¹⁹

In the field of radiology oncology, a number of investigators have reported on the use of and issues associated with consumer-level flatbed and slide scanners for digitizing dosimetry film. Lynch, Kozelka, Ranade et al. analyzed various performance characteristics of the Macbeth densitometer as well as the Epson and Microtek flatbed scanners used to read radiochromic film.²⁰ They found that after repeated scans the flatbed scanners warmed the film itself, which in turn affected the values they were measuring, and that the scanner did not respond uniformly over the entire area of the film. Hupe and Brunzendorf examined the use of the Nikon 8000 ED slide scanner for radiochromic film dosimetry.²¹ This scanner used light-emitting diode (LED) illumination which did not heat up the film as much as the lamps used in the flatbed scanners noted above. They also observed the non-uniform characteristics of the charge-coupled diode receptor array but found that the chief contributor to uncertainty in the process was the non-uniformity of the film itself. Paelinck, De Neve, De Wagter et al. evaluated the Epson Pro 1680 Expression scanner and found a warm-up effect where the reading was slightly higher on the first scan. They did not find other forms of drift over time. Along with the other investigators, they noted the “response over the scanner field is not uniform.”²²

Semiconductor theory indicates that warmer temperatures result in more noise in the electronic response of the optical sensor.²³ The above researchers did not find conclusive evidence of a warm-up effect, and neither did we as a result of our research.²³ Variations over time are illustrated in figures 1–4. Note that the luminance of the color channels is measured on a 0–255 scale,²⁴ so a change of 2.0 would be less than 1% of the entire range.

The Nikon scanners had lower CV in the color channel levels, suggesting a better suitability for this type of application.

The core of our analysis was to repeatedly measure the width of certain vessels in the selected eye using variations of our IVAN vessel caliber reading protocol. We made certain modifications to focus precisely on potential vessel width variations. Our hypothesis was that the variations in color channel level, particularly where color levels varied with respect to each other, might influence either the automated or human finding of the vessel walls. As noted above, no clear-cut correlation was found.

We were also interested in any other insights the experiment offered about the vessel caliber grading process using the IVAN software. The IVAN software attempts to automatically detect the optic nerve head and center a grading grid on it. We would not expect any incorrect

centration, as all images were of the same slide and the original photo quality was good. Nonetheless, we found a rate of 2.4% incorrectly centered grids. Despite this, we found that the average CV of the measured vessel widths was quite small; for the Nikon 5000 scans, the CV was 0.3 for the large artery and 0.4 for the large vein. In comparison, Newsom, Sullivan, Rassam et al. found CV of 1.5–7.5% for automated vessel diameter measurement and CV of 6–34% for an observer-driven measurement method.²⁴

In conclusion, the film digitization process did not produce systematic bias that would influence the measurement of vessel caliber or the consequent calculation of summary variables based on individual vessel width measurement. The variations in the vessel caliber were low and lent credence to the reproducibility of the IVAN-based vessel width grading process.

Acknowledgments

The authors wish to thank Mr. Scott Brandenburg, Ms. Holly Cohn, Ms. Lisa Grady, and Ms. Mary Kay Aprison for their assistance.

Supported by the National Institutes of Health grant EY06594 (R Klein, BEK Klein).

Reference List

1. Diabetic Retinopathy Preferred Practice Pattern®. San Francisco, CA: American Academy of Ophthalmology; 2003. American Academy of Ophthalmology Retina Panel.
2. Klein R, Klein BE, Neider MW, Hubbard LD, Meuer SM, Brothers RJ. Diabetic retinopathy as detected using ophthalmoscopy, a nonmydriatic camera and a standard fundus camera. *Ophthalmology* 1985;92:485–491. [PubMed: 4000642]
3. Gilchrist J. Computer processing of ocular photographs--a review. *Ophthalmic Physiol Opt* 1987;7:379–386. [PubMed: 3331732]
4. Fulton, W. A few scanning tips. Plano, TX: 1998.
5. Duncan DD, Shukla OB, West SK, Schein OD. New objective classification system for nuclear opacification. *J Opt Soc Am A Opt Image Sci Vis* 1997;14:1197–1204. [PubMed: 9168593]
6. Hubbard LD, Brothers RJ, King WN, et al. Methods for evaluation of retinal microvascular abnormalities associated with hypertension/sclerosis in the Atherosclerosis Risk in Communities Study. *Ophthalmology* 1999;106:2269–2280. [PubMed: 10599656]
7. Klein R, Klein BE, Marino EK, et al. Early age-related maculopathy in the cardiovascular health study. *Ophthalmology* 2003;110:25–33. [PubMed: 12511342]
8. Chapman N, Witt N, Gao X, et al. Computer algorithms for the automated measurement of retinal arteriolar diameters. *Br J Ophthalmol* 2001;85:74–79. [PubMed: 11133716]
9. Lynch BD, Kozelka J, Ranade MK, Li JG, Simon WE, Dempsey JF. Important considerations for radiochromic film dosimetry with flatbed CCD scanners and EBT GAFCHROMIC film. *Med Phys* 2006;33:4551–4556. [PubMed: 17278806]
10. Hupe O, Brunzendorf J. A novel method of radiochromic film dosimetry using a color scanner. *Med Phys* 2006;33:4085–4094. [PubMed: 17153388]
11. Paelinck L, De Neve W, De Wagter C. Precautions and strategies in using a commercial flatbed scanner for radiochromic film dosimetry. *Phys Med Biol* 2007;52:231–242. [PubMed: 17183138]
12. Mosher A, Klein BE, Klein R, Knudtson MD, Ferrier NJ. Comparison of retinal vessel measurements in digital vs film images. *Am J Ophthalmol* 2006;142:875–878. [PubMed: 17056378]
13. Cideciyan, AV.; Nagel, JH.; Jacobson, SG. Modeling of high resolution digital retinal imaging engineering in Medicine and Biology Society; Proceedings of the Annual International Engineering in Medicine and Biology Society; 1991. p. 264-266.
14. Wilson, M.; Nemeth, SC.; Edwards, A.; Soliz, P. Optimizing retinal image digitization for improved digital processing and visualization; Proceedings of the 14th IEEE Symposium on Computer-Based Medical Systems (CMBS'01); 2001 July 26–27; Bethesda, MD; Los Alamitos, CA: IEEE Computer Society; 2001.

15. Russ, JC. The Image Processing Handbook. 2nd edition. Boca Raton, FL: CRC Press; 1995. p. 28
16. Newsom RS, Sullivan PM, Rassam SM, Jagoe R, Kohner EM. Retinal vessel measurement: comparison between observer and computer driven methods. Graefes Arch Clin Exp Ophthalmol 1992;230:221–225. [PubMed: 1597285]

Figures 1-4.

Nikon 5000 color channel levels vs vessel width. Results from all runs are plotted. Four vessels were measured and plotted here, respectively the larger arteriole, smaller arteriole, larger venule, and smaller venule. All runs adjusted for color channel values.

Table 1

Channel Levels in Nikon 5000 by Run, Comparison Between Runs

Channel		P-value Cold 1 vs Cold 2		Run 3 (Warm)	Run 4 (Warm)	P-value Warm 3 vs Warm 4		Average Across All Runs
		Run 1 (Cold)	Run 2 (Cold)			Warm 3	Warm 4	
Red	Mean	142.4	141.8	142.1	142.4			142.2
	SD	0.19	0.12	0.61	0.81	.03	.22	0.56
	CV	0.14	0.08	0.43	0.57			0.40
Green	Mean	50.3	51.4	52.3	53.4			51.9
	SD	0.34	0.31	0.67	0.47	<.001	<.001	1.21
	CV	0.67	0.61	1.29	0.89			2.33
Blue	Mean	83.0	82.9	83.1	83.5			83.1
	SD	0.03	0.01	0.52	0.52	<.001	<.001	0.43
	CV	0.04	0.01	0.63	0.63			0.52

Cold, equipment turned on for 30 minutes; Warm, equipment on for more than four hours; NA, not applicable; SD, standard deviation; CV, coefficients of variation.

Table 2

Channel Levels in Nikon LS-2000 by Run, Comparison Between Runs

Channel		P-value Cold 1 vs Cold 2		P-value Warm 3 vs Warm 4		P-value Average Across All Runs	
		Run 1 (Cold)	Run 2 (Cold)	Run 3 (Warm)	Run 4 (Warm)	Cold	Warm
Red	Mean	*	132.4	130.1	129.8		130.8
	SD	*	0.38	NA	0.17	<.001	<.001
	CV	*	0.29	0.08	0.13		0.92
Green	Mean	*	50.13	48.8	48.9		49.3
	SD	*	0.11	NA	0.09	0.43	<.001
	CV	*	0.22	0.19	0.22		1.26
Blue	Mean	*	69.4	67.3	67.3		68.0
	SD	*	0.12	NA	0.08	<.001	<.001
	CV	*	0.18	0.12	0.14		1.01
							1.48

* Data missing.

Cold, equipment turned on for 30 minutes; Warm, equipment on for more than four hours; NA, not applicable; SD, standard deviation; CV, coefficients of variation.

Table 3

Channel Levels in Polaroid 4000 by Run, Comparison Between Runs

Channel		P-value Cold 1 vs Cold 2		Run 3 (Warm)	Run 4 (Warm)	P-value Warm 3 vs Warm 4		Average Across All Runs
		Run 1 (Cold)	Run 2 (Cold)			Warm 3	Warm 4	
Red	Mean	150.9	155.5	155.2	133.6			148.8
	SD	0.15	0.18	0.13	0.21	<.001	<.001	9.01
	CV	0.10	0.12	0.08	0.15			6.05
Green	Mean	69.7	72.6	73.3	56.7			68.11
	SD	0.23	0.42	0.25	0.43	.43	<.001	6.74
	CV	0.33	0.59	0.34	0.75			9.89
Blue	Mean	92.7	96.5	96.5	76.9			90.65
	SD	0.23	0.23	0.23	0.21	<.001	<.001	8.12
	CV	0.25	0.24	0.24	0.27			8.96

Cold, equipment turned on for 30 minutes; Warm, equipment on for more than four hours; NA, not applicable; SD, standard deviation; CV, coefficients of variation.