Journal of Imaging Science and Technology® 66(3): 030508-1–030508-6, 2022.

The Objective Measurement and Subjective Perception of Flexible ENT Endoscopes' Image Quality

G. Geleijnse▲†, L. L. Veder†, M. M. Hakkesteegt, and R. M. Metselaar
Erasmus Academic Medical Center, Department Ear Nose & Throat, Dr Molewaterplein 40, Rotterdam, The Netherlands, 3015CA
E-mail: g.geleijnse@erasmusmc.nl

Abstract. *ENT-flexible endoscopes are an important tool for ear, nose and throat (ENT) professionals to examine the upper airway. Although image quality has improved significantly in the past decade, there is no generally accepted approach to measure this objectively. Sharpness, visual noise and color fidelity are aspects of image quality that can objectively be measured. The purpose of this study was to explore the relationship between these quality metrics and the subjective perception of image quality by ENT-professionals. The image quality of six different flexible endoscopes was assessed objectively and subjectively. Objective measurements were obtained using the Rez Checker Target Nano Matte and comprised sharpness (MTF50), visual noise and color fidelity (CIE ΔE 2000). Subjective image quality ranking was obtained by presenting images of a single larynx to 30 ENT-professionals in a forced pairwise comparison and asking them to select the image with the best image quality. Differences in image quality between endoscopes are reliably detected by objective measurement and subjective assessment. A strong positive correlation was found between sharpness and subjective ranking ($p < 0.005$). Visual noise and color fidelity may be relevant, but did not correlate with the subjective assessment and were probably overshadowed by the strong correlation between sharpness and subjective ranking in the data. The authors found that number of pixels on screen to display the registered image differs per type of endoscope, however more pixels do not necessarily imply a sharper image. The authors will continue their investigation of image quality metrics and their relation to diagnostic accuracy which can provide feedback on design and manufacturing optimization to the industry.* © *2022 Society for Imaging Science and Technology.*
[DOI: 10.2352/J.ImagingSci.Technol.2022.66.3.030508]

1. INTRODUCTION

Flexible endoscopes are essential to examine nose, throat and upper airway [1]. Earlier, these fiber optic endoscopes captured images that were observed directly with the eyepiece or using a small camera that was connected to it. Fiber optic endoscopes have been gradually replaced by digital endoscopes due to higher image quality [2–6].

Although ear, nose and throat (ENT) endoscopes are commonly used, literature concerning the diagnostic accuracy is limited. Five studies compared fiber optic to digital endoscopes [2–6]. In two studies of Eller et al. [2, 3], the use of digital endoscopes showed to be equivalent or slightly better compared to the use of fiber optic endoscopes, but signs of reflux appeared to be better seen using fiber optic endoscopes. Plaat et al. showed that ENT professionals preferred the image quality of digital endoscopes, but found no significant difference in diagnostic accuracy [4]. Scholman et al. showed that high definition digital endoscopes provide better image quality and significantly improved diagnostic accuracy [5, 6]. These clinical studies do not mention a consistent approach to measure image quality objectively. Image quality is a collective term which includes a variety of characteristics, such as sharpness, visual noise and color fidelity [7, 8]. Objectively quantifying those characteristics requires an *in vitro* setting, with a setup that positions the endoscope to capture the image of the target in a uniform and reproducible way. This allows objective comparison without burdening patients, and has a better controlled environment so measurements can be reproduced and compared over time. It is important for hospitals, equipment manufacturers and researchers to define, measure and communicate image quality and determine the characteristics that are important for improving the accuracy of diagnosis. For example, the field of Radiology has taken great strides to objectively measure image quality and produce consistently high-quality images [9, 10].

Singular studies have objectively measured aspects of image quality of flexible endoscopes [11–14], but a consistent approach is missing. Camera manufacturers have developed and continuously improved standards that can be adopted and adapted to objectively measure aspects of image quality of endoscopes [15, 16]. The *Rez Checker Target Nano Matte* (Figure 1) is a test chart that was designed by *Image Science associates* to be more suited to narrow illumination geometries, notably those used in endoscopic imaging. It supports the objective measurement of the image quality metrics, including sharpness, noise and color fidelity. In our earlier study, we showed that these methods can be applied to different digital ENT endoscopes and yield measurement results that are objective, reproducible and specific per type of endoscope [15].

The purpose of this study was to explore the relationships between the objectively measured sharpness, visual noise and color fidelity using the *Rez Checker Target Nano Matte* and the subjective perception of image quality by ENT-professionals of flexible endoscopes to see if the objective metrics are relevant in the field of ENT endoscopy.

▲IS&T Member.
†These authors contribute equally to this work.

Received Sept. 11, 2021; accepted for publication Jan. 13, 2022; published online March 11, 2022. Associate Editor: Marius Pedersen.
1062-3701/2022/66(3)/030508/6/$25.00

Figure 1. Rez checker target nano matte designed by Image Science associates. The circles are used as reference for automatically determining the regions of interest within the test image. The colored patches are used for measuring color fidelity. The gray patches are used for measuring the noise and the optoelectronic conversion function which is necessary for calculating the modulation transfer function. The gray slanted edges are used for measuring the step response and calculating the modulation transfer function to assess sharpness. Note that images cannot be reproduced with original resolution and color because of degradation associated with various steps for preparation and printing of the manuscript.

2. METHODS

Six different types of flexible ENT endoscopes that were available at our clinic were used for this study. Based on our clinical experience we included one low image quality endoscope and one high image quality endoscope. The other four endoscopes were offered by three manufacturers as state-of-the-art endoscopes for the pediatric care with a small diameter and good image quality. Since the purpose was to study image quality metrics and not to compare the types of endoscopes, we refer to scope A to F throughout the manuscript and disclose the specific type of endoscope and video processor here once for sake of reproducibility of the study: (A) Pentax fiber optic endoscope *FNL-10RP3* connected to a *Xion HD camera EndoSTROBE PL Spectar*, (B) *Storz 11102 CMAC* connected to a *8403 ZXK*, (C) *Olympus ENF-V4* connected to a *CV-170*, (D) *Pentax VNL8-J10* connected to a *Defina HDK 3000*, (E) *Pentax VNL9-CP* connected to a *VIVIDEO CP-1000* and (F) *Xion HD* connected to a *EndoSTROBE PL Spectar*. Images were acquired using the settings as indicated by the vendor representatives. We would like to stress that these types of endoscopes are not necessarily the best performing endoscope per manufacturer.

The included endoscopes have different diameters. Manufacturers use different methods to measure endoscope diameter, therefore we measured all endoscopes uniformly by placing the endoscope tip parallel to the jaws of a *Mitutoyo* caliper.

2.1 *Objective Image Quality Assessment*

Sharpness was measured using the slanted edge method, the default standard and is described in detail in ISO 12233 [15, 17]. This method measures the modulation transfer function (MTF), which is a measure of how well a system can transfer contrast at a certain spatial frequency, i.e., resolution. The MTF of each endoscope was measured objectively by imaging the *Rez Checker Target Nano Matte* at 3.0 cm distance. The *Image Science associates* developed this test chart specifically to test the image quality of endoscopes and contains a horizontal and vertical slanted edge. A custom written Matlab script (MATLAB 2019b, Natick, MA, USA) detected the three black circles automatically. These circles served as reference for determining the regions of interest within the image on the test chart. The sharpness was quantified by the frequency (line pairs per mm) at which the MTF first has a value of 50% (MTF50) [15].

Each system projects its image on screen using a certain number of pixels. The number of effectively used pixels (rows and columns) were counted.

Visual noise was computed as a weighted sum of variances from the luminance and two chrominance channels of pixel values on the twelve gray patches [18, 19]. The visual noise values were combined into one average. Gray patches that showed oversaturation were excluded from the average, because these patches typically have no noise and yield an underestimated average noise.

Color fidelity was quantified by calculating the perceptual color difference (CIE ΔE 2000) between the eighteen target color patches and the imaged color patches [20]. All color differences were combined into one average.

2.2 *Subjective Image Quality Assessment*

This study was reviewed and approved by the Medical Ethical Review Committee Erasmus MC (MEC-2021-0387). All six endoscopes were used to image the larynx of a single volunteer. The first series of images was acquired using endoscope E and stored using an *Epiphan DVI2USB3* frame grabber as uncompressed 24-bit BMP-files. The best out of several images was selected based on visible anatomy and lack of motion blur. The selected image was shown as an example to the clinical examiners using the other endoscopes and they were asked to reproduce this image as accurately as possible. The best out of several images were selected for each type of endoscope. The selected images were used for a pairwise comparison of all six endoscopes ($n = 6$), resulting in $(n^2 - n)/2 = (6^2 - 6)/2 = 15$ test pairs of images to be compared [21]. To test the reproducibility of the observers, the images were swapped and mirrored resulting in another 15 retest pairs to be compared. The total of 30 pairs were imported in *MS PowerPoint*, scrambled and presented to ENT-professionals on a color calibrated diagnostic display (EIZO RadiForce MX315W 4096x2160). The images were presented with their native resolution by carefully formatting the uncompressed images in *MS PowerPoint*. The resulting presentation was checked by counting the pixels using the "print-screen" function. The characteristics sharpness, noise

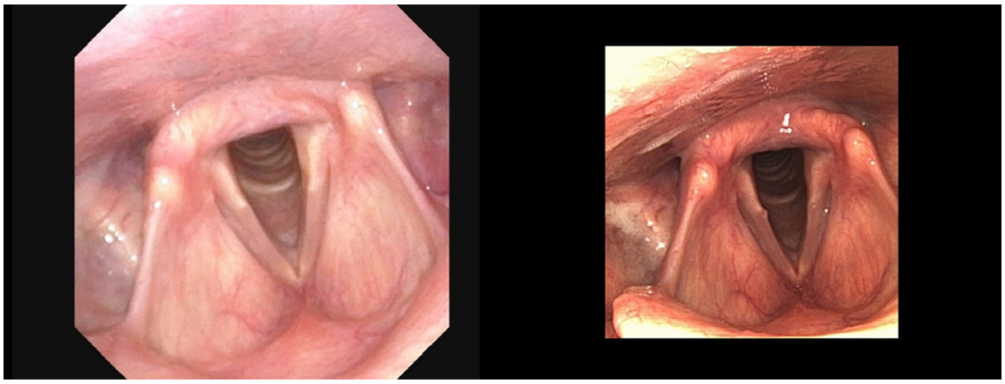

Figure 2. Example images for the pairwise comparison of the endoscope C (left) to endoscope E (right). The actual screen had more pixels (4096x2160) and the images were displayed with their native resolution, resulting in a smaller relative area. See note in Fig. 1.

and color fidelity were verified by comparison of print screens of *MS PowerPoint* in presentation mode displaying the test images for objective measurements to the original test images. The monitor was calibrated to sRGB color space with a color temperature of 6500 K, gamma of 2.2 and the luminance ranged from 0.5 to 300 cd/m^2. An example of a pairwise comparison screen is presented in Figure 2. The observers were asked to select the image with the best image quality characteristics for diagnostic purposes and neglect the influence of position and viewing angle. The ranking score was calculated as the sum of votes per type of endoscope.

2.3 *Correlation of Objective and Subjective Quality Assessment*

The consistency of the subjective and objective quality assessment was determined by calculating the correlation between the sum of votes and the MTF50 for sharpness, the average visual noise, and the average color differences. The correlation was calculated using the Spearman correlation coefficient [22]. The Bonferoni correction [23] was applied to correct for multiple testing resulting in p-values smaller then 0.017 to be significant.

3. RESULTS

The diameters of the endoscopes are measured and listed in Table I.

3.1 *Objective Quality Assessment*

The MTF50 values are shown in Table II and indicate that the horizontal sharpness differs from the vertical sharpness. These values were combined into an average as a measure for the overall sharpness, and included in Table I. The noise and color fidelity were measured and are presented in Table I as well as the number of effectively used horizontal and vertical pixels on the display.

3.2 *Subjective Image Quality Assessment*

Thirty ENT professionals participated, including 16 ENT medical specialists (oncologists, pediatricians and generalists), 11 ENT-residents, 2 physician assistants and 1 speech therapist specialized in laryngoscopy. Each observer required about 2-3 minutes to complete the comparison session. The number of votes for each pairwise comparison, test and retest, is presented in Table III. The first six columns provide detailed insights in the voting data we collected, which are summarized and easy to interpret in the final column. We provide an example to aid the interpretation of Table III: on the third row and fourth column, endoscope C is compared to endoscope D. The first number indicates that 7 observers preferred endoscope C over endoscope D. The second number indicates that 4 observers preferred endoscope C upon retest. These test and retest results are linked to the fourth row and third column where 23 observers preferred endoscope D over endoscope C. On retest 26 observers preferred endoscope D. Detailed analysis of this pair showed that observers can vote different upon retest: three observers changed from C to D and two observers changed from D to C. The final column lists the sum of votes per endoscope during test and retest. The sum of votes for the test pairs are also included in Table I.

Sharpness was consistently mentioned as the criterion to differentiate; color and image size were not mentioned as differentiators. Noise was only mentioned when comparing endoscopes, A to B.

3.3 *Correlation Objective and Subjective Quality Assessment*

The Spearman correlation coefficient of the subjective quality score and the objective MTF50 was calculated as 0.94 ($p < 0.005$). The correlation coefficients for visual noise -0.09 ($p = 0.87$) and color fidelity -0.31 ($p = 0.54$) were not significant.

4. DISCUSSION

We objectively assessed image quality by measuring sharpness, visual noise and color fidelity of different types of flexible ENT endoscopes and explored the relationships between these metrics and the subjective perception of image quality.

We found not only objective and subjective differences between fiber optic endoscope A and digital endoscopes

Table I. Summary of objective & subjective image quality data. MTF50 (larger MTF values indicate sharper images), visual noise (lower values indicate less noise) and CIE ΔE 2000 are an average of ten repeated measurements (lower values indicate better color fidelity).

	Diameter [mm]	Number of pixels [horizontal × vertical]	Sum of votes	Sharpness MTF50 average [lp/mm]	Visual noise average	Color fidelity CIE ΔE 2000 average
Endoscope A	3.9	830 × 830	1	1,71	33,37	10,26
Endoscope B	2.9	1024 × 768	29	1,66	2,03	11,51
Endoscope C	2.9	1080 × 1080	69	2,15	2,75	9,30
Endoscope D	3.2	1120 × 845	88	2,60	4,47	10,74
Endoscope E	3.2	800 × 800	118	3,58	4,07	11,27
Endoscope F	4.2	1475 × 1080	145	4,54	3,48	7,97

Table II. Objectively measured horizontal and vertical sharpness expressed as MTF50 in line pairs per millimeter. Each MTF50 measurement was repeated ten times and averaged. Larger MTF50 values indicate sharper images capable of depicting finer details. It can be observed that the horizontal sharpness is different from the vertical sharpness.

	MTF50 horizontal [lp/mm]	MTF50 vertical [lp/mm]	MTF50 average [lp/mm]
Endoscope A	2,13	1,72	1,71
Endoscope B	1,57	1,75	1,66
Endoscope C	2,37	1,93	2,15
Endoscope D	2,87	2,33	2,60
Endoscope E	3,83	3,33	3,58
Endoscope F	4,84	4,24	4,54

B-E, but also between digital endoscopes B-E. This offers a new perspective on the previous studies performed by Eller, Plaat and Scholman who compared fiber optic endoscopes to digital endoscopes [2–6]. The results of these studies are limited to the specific types of endoscopes that are used in the study and cannot be applied to all types of fiber optic and digital endoscopes in general. Relating to these clinically relevant results is difficult since there are no quality metrics reported by the investigators or manufacturers.

The objective measurements of sharpness showed that the minimum amount of discernable detail varied per type of endoscope (Table I). MTF50, used for sharpness, correlated very well with the subjectively perceived diagnostic image quality by ENT-professionals. Only endoscope A was ranked different by the observers than was expected on the MTF50. This can be attributed to the large visual noise when compared to the other endoscopes, due to the discrete fiber pattern (Table I). The strong correlation between ranking and sharpness makes sense, because the main goal of diagnostic imaging is the detection of abnormalities. Sharper images provide more contrast and more detail, which are mandatory conditions for the detection of differences between types of tissues. Although it seems plausible that sharper images will enable ENT-professionals to better detect abnormalities, we have not proven this yet. Sharpness of endoscopes can be different for the vertical or horizontal direction (Table II). Apparently, some small details that can just be perceived may be missed if the endoscope is rotated by 90°. This was also concluded by Komatsu et al. [24].

While the correlation coefficients for visual noise and color fidelity were not significant, we do not think that these metrics are irrelevant, but the role of sharpness in this study was merely dominant. The large correlation coefficient indicates that almost all variation in the data is explained by the sharpness and there is too little variation left to be explained by any other characteristic.

Visual noise is easily perceived on uniform areas, while biological tissues are typically non-uniform due to the presence of blood vessels and anatomic structures. Non-uniformity of the scene typically lowers the sensitivity to noise. The variation in visual noise levels between endoscopes can be perceived on the *in vitro* images of the test chart containing uniform patches, but this variation may not be sufficient to be perceived *in vivo* except for endoscope A, which contained excessive noise due to the fiber optic structure.

Color fidelity relates to the absolute color accuracy which is different from contrast and could for example be useful for detecting inflammation, indicated by increased redness. However, inflamed tissues are identified by comparison of color to surrounding tissues, limiting the importance of color fidelity and increasing the importance of contrast. Koningsberger et al., showed that color vision deficiency does not affect the endoscopic competence of gastroenterologists [25].

The subjective pairwise comparison showed clear differences in perception of image quality of endoscopes that were used in this study. Observers voted unanimously or a convincing majority preferred one endoscope. Votes of ENT-residents and more experienced ENT-specialists were similar, as residents were required to have more than 3 months of clinical experience in using endoscopes.

Table III. Votes per pairwise comparison (test | retest) and sum of votes (test | retest).

	Endoscope A	Endoscope B	Endoscope C	Endoscope D	Endoscope E	Endoscope F	Sum of votes
Endoscope A	–	1 | 1	0 | 0	0 | 0	0 | 0	0 | 1	1 | 2
Endoscope B	29 | 29	–	0 | 1	0 | 1	0 | 0	0 | 0	29 | 31
Endoscope C	30 | 30	30 | 29	–	7 | 4	2 | 2	0 | 0	69 | 65
Endoscope D	30 | 30	30 | 29	23 | 26	–	5 | 5	0 | 0	88 | 90
Endoscope E	30 | 30	30 | 30	28 | 28	25 | 25	–	5 | 2	118 | 115
Endoscope F	30 | 29	30 | 30	30 | 30	30 | 30	25 | 28	–	145 | 147

Some observers did change preference, which is natural according to how the just-noticeable-difference is defined: it is the difference in image quality, where 50% of the observers perceive a difference, when the other 50% cannot perceive a difference in image quality and will choose randomly. Observers changed preferences more frequently when endoscopes were compared to a closely ranked endoscope, indicating that the image quality difference is harder to perceive. Sharpness was consistently mentioned as the criterion to make a difference, while color and image size were never mentioned as criteria. Noise was only mentioned when comparing endoscope, A to B, probably because the sharpness was similar and visual noise differed.

Pairwise comparison is a powerful tool for subjective ranking [21]. It is an easy task for observers compared to sequentially scoring image quality, which is based on memory. Scores based on memory may shift e.g., a poor-quality image may be scored as "normal" if it is preceded by images with poorer quality. Pairwise comparison does have a risk: circular ranking may occur, e.g., the first image is scored better than the second, and the second better to the third, but the third better to the first. We saw no circular ranking in this study.

Another finding was that the number of pixels on screen differs per type of endoscope and that more pixels do not necessarily imply a sharper image. For example, endoscope B displayed far more pixels compared to endoscope E, but the sharpness was lower and the subjectively perceived image quality was obviously poorer as well. Terms like SD, HD and 4K refer to the number of pixels on screen [26] and can be mistaken for better image quality. More pixels on screen are indeed a requirement for displaying sharper images, but do not imply that the image itself contains more details.

The findings of this study have to be seen in light of the following limitations. Firstly, image quality is an umbrella term, and there are more metrics than we measured that cover different aspects that might be relevant for diagnostic purposes [7]. Other aspects could be confounding with sharpness and contribute to the strong correlation we found. Secondly, sharpness was measured using the modulation transfer function, using a method that is designed to discard noise. An endoscope with a good MTF, but high levels of noise, can have a poor signal-to-noise ratio. We did not incorporate an image quality metric that combines sharpness and noise that could possibly have an even better correlation than we found for sharpness alone. Lastly, this study relied on ENT-professionals being able to judge the image quality for diagnostic purposes and not for personal appeal or high fidelity like art preservation. Further investigation of objective image quality metrics and their relation to diagnostic accuracy is required to understand which metrics are important for better diagnostics. This is likely to vary on the basis of medical specialization. Subsequently it can provide feedback to the industry on the desired metrics to optimize. We recommend measuring and reporting image quality metrics in future studies to the diagnostic accuracy of endoscopes in order to compare and combine study results and perhaps set requirements for standard of care.

5. CONCLUSION

The variation of image quality between endoscopes can be reliably detected by objective and subjective assessments. Our data show that sharpness has a strong positive correlation with the subjective quality ranking and probably overshadowed the correlation between visual noise or color fidelity in the data. Sufficient pixels on screen are necessary for displaying sharp images, but do not imply a sharp image. Further investigation of image quality metrics and their relation to diagnostic accuracy can provide feedback on design and manufacturing optimization to the industry.

5.1 *Disclosures*

This study was driven by curiosity. The authors have no relevant financial interests in the manuscript and no other potential conflicts of interest to disclose.

ACKNOWLEDGMENT

We would like to thank Jet de Gier and Sahar Mokhles for recording the laryngoscopic images and the ENT-professionals for comparing the image pairs. We also thank EIZO Europe GmbH for loaning the diagnostic monitor and remote support for calibration. We acknowledge Olympus Nederland B.V, PENTAX Nederland B.V. and KARL STORZ Endoscopie Nederland B.V. for their support and loaning endoscopes and video processors. Further support was solely from institutional and/or departmental sources. There are no sources of funding to be acknowledged.

REFERENCES

[1] B. C. Paul, S. Chen, S. Sridharan, Y. Fang, M. R. Amin, and R. C. Branski, "Diagnostic accuracy of history, laryngoscopy, and stroboscopy," The Laryngoscope **123**, 215–219 (2012).

[2] R. Eller, M. Ginsburg, D. H.-A, Y. Lurie, K. Lyons, and R. Sataloff, "Flexible laryngoscopy: A comparison of fiber optic and distal chip technologies. Part 1: vocal fold masses," J. Voice **22**, 746–750 (2008).

[3] R. Eller, M. Ginsburg, D. H.-A, Y. Lurie, K. Lyons, and R. Sataloff, "Flexible laryngoscopy: A comparison of fiber optic and distal chip technologies-part 2: Laryngopharyngeal reflux," J. Voice **23**, 389–395 (2009).

[4] B. D. Plaat, B. F. van der Laan, J. Wedman, G. B. Halmos, and F. G. Dikkers, "Distal chip versus fiberoptic laryngoscopy using endoscopic sheats: diagnostic accuracy and image quality," Laryngology **8**, 2227–2232 (2014).

[5] C. Scholman, J. M. Westra, M. A. Zwakenberg, F. G. Dikkers, G. B. Halmos, J. Wedman, J. E. Wachters, B. F. van der Laan, and B. E. Plaat, "Differences in the diagnostic value between fiberoptic and high definition laryngoscopy for the characterisation of pharyngeal an laryngeal lesions: A multi-observer paired analysis of videos," Clin. Otolaryngology **45**, 119–125 (2020).

[6] C. Scholman, J. M. Westra, M. A. Zwakenberg, F. G. Dikkers, G. B. Halmos, J. Wedman, J. E. Wachters, B. F. vanderLaan, and B. E. Plaat, "High-definition videolaryngoscopy is superior to fiberoptic laryngoscopy: a 111 multi-observer study," Eur. Arch. Otorhinolaryngol **278**, 1927–1932 (2021).

[7] D. Williams, P. D. Burns, and L. Sraff, "Imaging performance taxonomy," Proc. SPIE **7242**, 724208 (2009).

[8] D. Williams, "Debunking of specsmanship: Progress on ISO/TC42 standards for digital capture imaging performance," *Proc. IS&T 2003 PICS Conf.* (IS&T, Springfield, VA, 2003).

[9] E. Samei and M. J. Flynn, "A method for measuring the presampled MTF of digital radiographic systems using an edge test device," Med. Phys. **25**, 102–113 (1998).

[10] E. Samei, N. T. Ranger, J. T. Dobbins Iii, and Y. Chen, "Intercomparison of methods for image quality chracterization. I. Modulation transfer function," Med. Phys. **33**, 1454–1465 (2006).

[11] P. F. Niederer, J. Haefliger, P. Blessing, Y. Lehareinger, D. Doswald, and N. Felber, "Image quality of endoscopes," Proc. SPIE **4158** (2001).

[12] F. Vaz, L. Ripley, D. Lim, R. Kanegaonkar, and M. Harries, "Optical quality of the nasendoscope with and without the endosheath," J. Laryngol. Otol. **120**, 385–388 (2006).

[13] A. Tsunoda, K. Tsunoda, T. Sumi, S. Kishimoto, and K. Kitamura, "High-resolution ENT video endoscope with superior image quality equivalent to that of gastric video endoscopes," Otolaryngology **4** (2014).

[14] J. J. Bae, A. Vavilin, J. J. You, H. Kim, S. S. Ryu, J. J. Jang, and J. Woonggyu, "Smartphone-based endoscope system for advanced point-of-care diagnostics: Feasibility study," JMIR Mhealth Uhealth **5**, 1–16 (2017).

[15] G. Geleijnse, M. M. Hakkesteegt, G. J. de Groot, and M. R. Metselaar, "Measuring Image Quality of ENT Chip-on-tip Endoscopes," J. Imaging Sci. Technol. **65** (2021).

[16] R. Nishitha, S. Amalan, S. Sharma, A. K. Gurrala, S. Preejith, J. Joseph, and M. Sivaprakasam, "Image quality assessment for endoscopy applications," *IEEE Int'l. Symposium on Medical Measurements and Applications (MeMeA), Neuchâtel* (IEEE, Piscataway, NJ, 2021).

[17] ISO 12233 :2017 Photography – Electronic still picture imaging – Resolution and spatial frequency responses.

[18] ISO 15739: 2017 Photography – Electronic still-picture imaging – Noise measurements.

[19] S. C. Kelly and B. W. Keelan, "ISO 12232 revision: determination of chrominance noise weights for noise-based ISO calculation," Proc. SPIE **5668** (2005).

[20] M. Luo, G. Cui, and B. Rigg, "The Development of the CIE 2000," Color Res. Appl. **26**, 340–350 (2001).

[21] R. K. Mantiuk, A. Tomaszewska, and R. Mantiuk, "Comparison of four subjective methods for image quality assessment," Comput. Graph. Forum **31**, 1–13 (2012).

[22] C. Spearman, "The proof and measurement of association between two things," The Am. J. Psychol. **15**, 72–101 (1904).

[23] C. E. Bonferroni, Teoria statistica delle classi e calcolo delle probabilità, Florence: R Istituto Superiore di Scienze Economiche e Commerciali di Firenze, (1936).

[24] K. Komatsu, Y. Kitsuta, T. Ishii, D. Yamaguchi, T. Gunshin, S. Nakajima, T. Matsubara, S. Matsuse, K. Nakamura, T. Hiruma, R. Ito, H. Fujita, R. Inokuchi, and N. Yahagi, "Comparison of image quality of Video-laryngoscopes," Jpn J. Trauma Emerg. Med. **2**, 63–70 (2011).

[25] J. Koningsberger, D. Norren, J. van Niel, and W. Dekker, "Does color vision deficiency in the endoscopist influence the accuracy of endoscopic diagnosis? An anonymous study with Dutch gastrointestinal endoscopists," Endoscopy **26**, 549–553 (1994).

[26] ASGE Technology Committee "High-definition and high-magnification endoscopes," Gastrointest Endosc. **80**, 919–927 (2014).