Published in final edited form as:

J Voice. 2014 January ; 28(1): 98–105. doi:10.1016/j.jvoice.2013.08.015.

Detection of chronic laryngitis due to laryngopharyngeal reflux using color and texture analysis of laryngoscopic images

Daniel R. Witt¹, Huijun Chen, M.D.², Jason D. Mielens, B.S.¹, Kieran E. McAvoy, B.S.¹, Fan Zhang, M.D.², Matthew R. Hoffman, Ph.D.¹, and Jack J. Jiang, M.D., Ph.D.^{1,2}

¹Department of Surgery, Division of Otolaryngology – Head and Neck Surgery, University of Wisconsin School of Medicine and Public Health, Madison, WI, 53792

²Department of Otolaryngology – Head and Neck Surgery, Shanghai EENT Hospital, Fudan University

Abstract

Objective—To determine if pattern recognition of hue and textural parameters can be used to identify laryngopharyngeal reflux (LPR).

Methods—Laryngoscopic images from 20 subjects with LPR and 42 control subjects without LPR were obtained. LPR status was determined using the Reflux Finding Score. Color and texture features were quantified using hue calculation and 2D Gabor filtering. Five regions were analyzed: true vocal folds, false vocal folds, epiglottis, interarytenoid space, and arytenoid mucosae. A multilayer perceptron artificial neural network with varying numbers of hidden nodes was used to classify images according to pattern recognition. Receiver operating characteristic (ROC) analysis was used to evaluate diagnostic utility and intraclass correlation coefficient analysis was performed to determine to gauge interrater reliability.

Results—Classification accuracy when including all parameters was $80.5 \pm 1.2\%$ with an area under the ROC curve of 0.887. Classification accuracy decreased when including only hue ($73.1 \pm 3.5\%$; area under the curve = 0.834) or texture ($74.9 \pm 3.6\%$; area under the curve = 0.852) parameters. Interrater reliability was 0.97 ± 0.03 for hue parameters and 0.85 ± 0.11 for texture parameters.

Conclusions—This preliminary study suggests that a combination of hue and texture features can be used to detect chronic laryngitis due to LPR. A simple, minimally invasive assessment would be a valuable addition to the currently invasive and somewhat unreliable methods currently used for diagnosis. Including more data will likely improve classification accuracy. Additional investigations will be performed to determine if results are in accordance with those provided by pH probe monitoring.

Keywords

laryngopharyngeal reflux; hue; Gabor filter; artificial neural network; laryngitis

© 2013 The Voice Foundation. Published by Mosby, Inc. All rights reserved.

Corresponding author: Jack J. Jiang, M.D., Ph.D., 1300 University Ave, 2725 Medical Sciences Center, Madison, WI 53792, 608-265-9854, jjjiang@wisc.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Conflicts of interest: None

INTRODUCTION

Approximately 15% of all patients presenting to the otolaryngology office have chronic laryngopharyngeal reflux (LPR).¹⁻³ 24-hour pH probe measurements indicative of LPR have been observed in 50% of patients with voice complaints.⁴ Despite the high incidence of this pathology in voice patients, current methods of diagnosing LPR can be unreliable.

LPR is the regurgitation of gastric contents onto the mucosal linings of the pharynx, larynx, and upper aerodigestive tract, resulting in a spectrum of nonspecific symptoms.⁵ Both Goldberg⁶ and Koufman⁷ identified a causal relationship between the presence of acidic gastric juice and mucosal tissue damage in animals, suggesting a similar damaging effect could occur on the laryngeal mucosa during LPR. The presence of acid and pepsin in this sensitive region causes a variety of physiological responses, such as laryngeal edema and erythema, mucosal hypertrophy,⁴ granuloma, carcinoma, and subglottic stenosis.^{8,9} These physical signs are often considered during LPR diagnosis along with common symptoms such as throat clearing, persistent cough, globus sensation, and changes in voice quality.¹⁰ Thus, there is an array of nonspecific signs and symptoms that point to LPR as an underlying etiology, making diagnosis controversial.

Despite the widespread prevalence of this disorder, current diagnosis is fairly subjective and can be inaccurate. Branski¹¹ demonstrated low inter-rater reliability among otolaryngologists assessing the same physical laryngeal signs, such as erythema, edema, and granulation. Some clinicians accept 24-hour ambulatory pH probe testing to be the gold standard in LPR diagnosis^{9, 12, 13}. 24-hour pH probes have shown 96% sensitivity and specificity for identifying the presence of acid reflux near the lower esophageal sphincter;¹⁴ however, other studies have reported much lower sensitivity and reproducibility (55%) for identifying abnormal amounts of proximal esophageal acid reflux.¹⁵ The accuracy of 24-hour pH probe testing can be limited due to incorrect probe placement, irregularity of reflux during testing, and low pH values used to identify periods of reflux.^{13, 16, 17} Moreover, cost factors, inconvenient procedure length, and the invasive nature of this method can deter patients from undergoing pH probe impedance testing.¹¹⁻¹³ Given some of the technical limitations and low availability of pH probe equipment in some clinics, the Reflux Finding Score (RFS) can serve as a reasonable, non-invasive surrogate for the pH probe. Park et al. reported high sensitivity (87.8%) and lower specificity (37.5%) for the RFS in diagnosing pH probe-validated LPR in patients with globus (n=57).¹⁸

New methods to diagnose LPR are warranted. Computer-based analysis of laryngeal image color and texture features offers an alternative which is objective, quantitative, and convenient. Few studies have been conducted on color and texture analyses of the larynx, with little investigation into the specific physical signs of LPR. Texture, as it pertains to digital images, constitutes descriptive metrics that describe human-perceived 'textures' of an image. That is, how segmented, patterned, 'bumpy', or 'smooth' an image appears. Verikas¹⁹ performed a study that examined the diagnostic capability of color and texture analysis in differentiating between healthy and abnormal vocal folds, and between various subgroups of vocal fold mass lesions. Hanson et al. performed a study quantifying erythema in chronic posterior laryngitis resulting from LPR.²⁰ Different areas of the larynx were analyzed for relative erythema and treatment efficacy was monitored using Red-Green-Blue (RGB) color analysis. This method focused solely on obtaining an erythema index for a limited number of regions, with no consideration of texture patterns.

We propose using pattern recognition of color and texture features obtained from laryngoscopic images to identify LPR. We hypothesized that hue and a combination of two-dimensional (2D) Gabor textural parameters would distinguish images obtained from

subjects with LPR from those obtained from controls. To test this hypothesis, we classified images based on hue and texture parameters using an artificial neural network (ANN).

MATERIALS AND METHODS

Selection Criteria

This study was conducted under the approval of the ethics committee of the Shanghai EENT Hospital. Laryngeal images were obtained from 20 subjects with LPR (18 males; age range: 37-67 years; mean age: 54.3 years) and 42 control subjects (25 males; age range: 24-78 years; mean age: 48.7 years) who required laryngoscopy for other laryngeal abnormalities. Subjects for both groups had laryngoscopy performed at the Eye, Ear, Nose, & Throat Hospital of Fudan University in Shanghai, China. Subjects in the LPR group presented to the clinic with signs and symptoms of LPR and were diagnosed with LPR according to the Reflux Finding Score assessment.⁵ A score exceeding 7 indicated the presence of LPR.⁵ Subjects comprising the control group were individuals presenting with unknown laryngeal complications who had an RFS ≤ 7 (since LPR was diagnosed when RFS ≥ 8). Patients comprising the control group presented to the clinic with varying laryngeal complications, including nasopharyngeal discomfort, foreign body sensation in pharynx, cysts near tongue base, chronic neck discomfort, and chronic sore throat.

ANN analysis requires knowledge of the class to which each dataset belongs before the pattern recognition algorithm can be evaluated. There is currently no perfect assessment of LPR. One currently available option, the Reflux Finding Score, demonstrates excellent intra- and interrater reliability⁵ and the cutoff employed in this study was demonstrated previously to be effective in distinguishing persons with and without LPR.⁵ While 24-hour ambulatory pH probe is considered a gold standard in LPR diagnosis, it has numerous well-noted limitations^{13,16,17} and is not performed regularly at our clinic. Accordingly, the Reflux Finding Score was considered an adequate indicator of LPR presence or absence for this preliminary study.

Image Collection

Sixty-two videos were obtained as part of standard clinical assessment. A 70° laryngoscope (model 8706 CA; KARL STORZ; Tuttlingen, Germany) with a mounted GPKS822 micro camera (Panasonic; Secaucus, NJ, USA) was used. This model delivers images of 752×582 pixels at a sensitivity of 6 Lux/F1.4. Importantly, the white balance function was used to prevent any potential effects of variable lighting on image color.

Protocol

Videos were assessed on VLC media software version 1.1.11 (VideoLAN Organization; Paris, France) which enabled the extraction of individual still image frames (figure 1). Images were saved as JPEG image files and then individually analyzed for hue and texture.

The following regions were evaluated: true vocal folds, false vocal folds, epiglottis, interarytenoid space, and arytenoids (figure 2). Each region was manually selected using a sensor sketch pad (Wacom Intuos 4 PTK-640; Wacom Co., Ltd.; Toyonodai Otonemachi, Kazo-shi, Saitama, Japan). These site selections were performed by two image analysts. Both analysts maintained similar technique to achieve a high level of standardization in region selection. The mean hue and texture parameters were derived from three independent boundary selection trials. Following site selection, mean hue and texture outputs were assessed for each selected region.

Data description

Part I: Hue detection—Images were transformed to hue-saturation-value color space. Once hue was identified, regions of the larynx were evaluated for redness. To achieve this hue, our images were transformed from the RGB color space into the hue-saturation-value color space using novel software developed in MATLAB (MathWorks; Natick, MA). The hue-saturation-value color space is derived directly from the RGB color space in a geometric sense.²¹ Each RGB value is represented as a tuple, or vector sum of red, green, and blue components. The equations for this transformation are as follows:

$$M=\max(R, G, B) \quad m=\min(R, G, B) \quad C=M - m \quad H' = \begin{cases} \frac{G-B}{C} \bmod 6, & \text{if } M=R \\ \frac{B-R}{C} + 2, & \text{if } M=G \\ \frac{R-G}{C} + 4, & \text{if } M=B \\ \text{undefined}, & \text{if } C=0 \end{cases} \quad (1)$$

$$H = 60^\circ \times H'$$

where R is the red component of the color tuple, G is the green component, and B is the blue component. These three values correspond to a unique color expressed by each pixel in a digital image. M is the value of the color component with the largest magnitude (R, G, or B depending on the tuple) and m is the minimum color component. Depending on which color value is M, H' (the raw hue) is converted for each image pixel using the R, G, and B values. As shown, C=Mm, where C is chroma and measures the colorfulness relative to the brightness of the perceived white, given a specific lighting. If C=0, such that M=m, H' renders that pixel as undefined and separate from the global hue calculation. Finally, a scaled hue value, H, is calculated and used in data analysis.

Conceptually, the RGB values naturally form a cube in three-dimensional space (figure 3). By turning this cube on a corner, with black (0,0,0) pointing down and white (1,1,1) directly above it, a measurement of hue can be obtained by calculating the angle between red (1,0,0) and the hue in question.

Through this conversion, hue angles were derived for all pixels in the laryngeal image. These angles lie on a hue wheel (figure 4A). Rotation around the circumference of this hue space corresponds to changing hue (i.e., changes in the color we observe optically). Thus, the software provides an average hue value for a manually selected region of an image, averaging the hue angles from each pixel in the region.

For the purpose of this investigation, the hue angles could lie within the domain of $-180^\circ < H < 180^\circ$. This was done so that the 0° hue measure was situated as a reference point in the classical red area of the hue wheel (figure 4B). For this reason, the dependent color-related variable was hue (i.e., the absolute distance away from the 0° reference point, regardless of direction).

Part II: Multi-channel 2D Gabor Filtering (Texture detection)—Image texture is a quantitative set of variables which mathematically describe perceived texture: the proximal contrasts in light intensity that the human eye captures and renders as a coherent image. This set of metrics—specifically, image texture—provides information about local spatial arrangement of color or color intensities in a specified region of an image.²² Hence, the goal of texture detection is distinct from that of color detection. Color detection aims to quantify the colors (e.g., hue) present in an image. Texture detection aims to quantify contiguous differences in color intensity of an image region; therefore, the specific hue of adjacent regions is not important in this case, as long as there are proximate disparities in color intensity.

Since our study involves natural images of the larynx with varying visual texture, our study adopted a statistical approach for texture analysis. Multi-resolution 2D Gabor filters were used for texture analysis.^{23,24} This method was selected as it allows a comprehensive evaluation of variable texture features for subsequent classification^{24–26} and mimics the way in which the human visual system recognizes textures.^{27–29}

Gabor filtering is an edge detection system, modeled as a complex sinusoid. The form of the Gabor function is as follows:

$$g(x, y; \lambda, \theta, \psi, \sigma, \gamma) = \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right) \cos\left(2\pi \frac{x'}{\lambda} \psi\right) \quad (2)$$

where λ = wavelength (scale) of the filter; θ = orientation; ψ = phase; σ = size; and γ = roundness.

In the filtered images, white regions correspond to where the edges are detected, with brightness reflecting the intensity of an edge position (figure 5). Edges, or color discontinuities in an image, correspond to larger variations in the amplitude response returned for each Gabor filter.²⁴ As seen in figure 5, the lightness (amount of white versus black) of these images reflects the total number and strength of edges detected within the image for that given orientation and scale. The identification of edge intensities (i.e., the degree of color intensity contrast in contiguous local regions) allows for effective quantification of a perceived texture. Texture, in general, can be fine, coarse, smooth, rippled, mottled, irregular, or lineated.²² The topographical profile of supraglottal mucosa is often described with these subjective textural descriptions in the clinic. For example, polyps, nodules or general tissue granulation will all demonstrate altered edge boundaries or localized protrusions, which can be identified by Gabor filters.

As shown, Gabor filters operate as a function of orientation and scale, among other variables. A filter with a given scale can present variability when its orientation is altered. Figure 5 represents a filter bank consisting of three different orientations, each with the same scale. Within the figure, A is a synthetic texture, and B-D represent Gabor filter outputs for different theta values ($\theta = 0^\circ, 45^\circ$, and 90° , respectively). Each filter responds to different regions of the synthetic texture based on the orientation of the lines in that texture. Thus, a single Gabor filter can only detect edges of a given wavelength oriented in a given direction. In virtually all natural textures, there are complex edge patterns that would be missed if using only one filter.²³ Therefore, a series of Gabor filters at varying wavelengths and orientations was used so that a more accurate description of the region could be obtained; this technique is known as a filter bank.^{24,25} By having a set of uniquely tuned filters, one effectively segments an image into sub-images that detail specific regional texture. Each sub-image can then contribute to the profile of the image under analysis, and be used as a parameter for image classification.

We used 24 discrete filters (six orientations and four scales). With eight discrete anatomical regions per source image, our total number of textural features is 192 per source image. Conceptually, this is the equivalent to having 192 sub-images whose features add up to the original image being evaluated. Thus, instead of analyzing the aggregate sum of the textural features, each image was decomposed into sub-images for a more comprehensive view of texture components. The filter bank was applied to each of the eight anatomic locations.

Data Analysis

Classification of images using an artificial neural network (ANN)—The hue index and textural features formed the input for classification using the ANN. A multi-layer

perceptron (MLP) ANN was used to provide non-linear, discriminant analysis of the image features.³⁰ A standard MLP function updated weight and bias values according to Levenberg-Marquardt optimization.

The MLP consisted of an input layer for data entry, a layer of hidden nodes (nodes=5, 10, 15, or 20), and an output layer which provided the classification outcome (i.e., non-LPR or LPR) (figure 6). Hue and texture data were presented to the input layer, computations were performed in the layer of hidden nodes, and an output value was obtained at each node of the output layer determining the class into which the particular data set was classified.

Image data were partitioned into training, validation, and testing sets in a 60-20-20 ratio, respectively. This division ensures classification is both accurate (adequate number of samples in training set) and generalizable (adequate number of samples in testing set). As there were more images available from controls (n=42) than from subjects with LPR (n=20), all classifications were performed with a subset of the control data so that classifications were balanced. This ensures the chance probability is 0.5. The samples which were included in the control group were varied across classifications to ensure all were included at some point. Five classifications were performed for each number of hidden nodes and average classification rates were obtained. Classification accuracy was also evaluated using receiver operating characteristic (ROC) analysis. Classifications were performed with the entire feature set (all hue and texture parameters) as well portions of it (all hue parameters; all texture parameters; hue and texture parameters from each region individually).

To evaluate interrater reliability, two raters performed all hue and texture parameter extraction. Intraclass correlation coefficient (ICC) analysis was performed for each parameter for each region. Due to the large number of texture parameters (24 parameters for each of 8 regions), average ICC for each region is reported along with standard deviation.

RESULTS

Summary hue values are provided in table 1. Optimal total classification accuracy was 80.5 $\pm$ 1.2% observed at 10 hidden nodes (table 2), with an area under the ROC curve of 0.887 (figure 7a). When including only hue parameters, classification accuracy was 73.1 $\pm$ 3.5% with an area under the ROC curve of 0.834 (figure 7b). When including only texture parameters, classification accuracy was 74.9 $\pm$ 3.6% with an area under the ROC curve of 0.852 (figure 7c). Single region classifications produced comparable results across locations.

Both hue and texture computations displayed high interrater reliability. Across regions, average ICC for hue was 0.97 $\pm$ 0.03 and average ICC for texture was 0.85 $\pm$ 0.11 (table 3).

DISCUSSION

We hypothesized that an ANN-based pattern recognition of hue and texture features would be able to distinguish between non-LPR and LPR laryngoscopic images. Our approach yielded an optimal classification accuracy of 80.5 $\pm$ 1.2%. The corresponding ROC curve provided an area under the curve equal to 0.887, indicating good sensitivity and specificity. While this is a preliminary study with a limited number of subjects, early results are encouraging and demonstrate the potential for this method to identify LPR based on hue and texture features of the laryngeal mucosa.

Few studies have quantified the degree of erythema present in patients with reflux. Hanson et al. noted significantly increased redness in the posterior arytenoid mucosa, vocal processes, and true vocal folds during reflux episodes.¹⁹ That study represents a valuable contribution in advancing objective, quantitative, endoscopic-based reflux assessment;

however, it did not include texture and was based on isolated measures of erythema rather than pattern recognition of multiple measurements. Considering the complexity of this disorder and its variation in presentation, a more detailed, comprehensive assessment is desirable.

To assess for LPR, our method first quantified prominent physical signs of the laryngeal mucosa. Common abnormalities associated with LPR include erythema, laryngeal edema, pseudosulcus vocalis, ventricular obliteration, posterior commissure hypertrophy, laryngeal granulomas, and lymphoid hypertrophy.^{4,12,31,32} Color and texture analyses can potentially identify the underlying quantitative properties of such abnormalities and are not dependent on isolating discrete physical abnormalities. Though isolated measurements of hue or texture in a single region would be nonspecific, pattern recognition nonlinear signal processing can be used if these parameters are applied to the entire larynx. Interestingly, the texture parameters appeared to have a bigger impact on classification accuracy than the hue parameters. This could reflect relevant differences in presentation, but could also be an artifact of artificial neural network processing, which typically performs better when more parameters are included. As 24 times the amount of texture parameters compared to hue parameters were included, the relatively small differences in classification accuracy and area under the ROC curve are likely negligible.

This study has several limitations which will be addressed in future investigations. Most notably, a challenge when proposing a new test to diagnose LPR is validating it against an existing diagnostic modality. We used the Reflux Finding Score assessment to confirm subjects had LPR. In a study by Belafsky et al., the 95% confidence interval for the Reflux Finding Score in control subjects had an upper boundary of 6.8. This study provides support for the cutoff value of 7 used to indicate the presence of LPR.⁵ Additionally, new support for the Reflux Finding Score was provided in a large series by Habermann et al. evaluating sensitivity to treatment with proton pump inhibitors.³³ Additional studies comparing our results using pattern recognition of laryngeal hue and texture to those obtained using pH probe monitoring are warranted.

Pertinent limitations of this study include a modest sample size and the potential for measurement error due to technical methodology inconsistency, as may occur due to lighting glare and different amounts of saliva on the arytenoid mucosa. We controlled for differences in lighting by using hue instead of color and also by using the white balance function when recording images. We also eliminated images which had noticeable glare. Methodological modifications to completely eliminate even seemingly small effects of glare will be the subject of further development. Furthermore, a relatively small sample size was included and the two groups had unequal numbers of subjects. Future studies with more subjects and a wider representation of LPR severity would provide a more generalizable color and texture profile.

Additionally, the control subjects were undergoing laryngoscopy for a voice concern unrelated to reflux and thus did not represent a classical healthy control group. While the subjects had low RFS scores, they may have had some color or textural features different from those of completely healthy subjects. It is true that there are other laryngeal disorders and underlying conditions that can result in some of the nonspecific signs and symptoms of LPR. However, the utility of the RFS is that we can have a standard for determining if a collection of laryngeal signs indicates the presence of LPR. For example, an individual with only erythema or only edema will not be included in the LPR group, since their mucosal presentation will not warrant a LPR diagnosis according to the RFS criteria.

Given the control group, our method aims to diagnose LPR among patients with other laryngeal signs and symptoms, and thus serves a practical, valuable service in the clinical setting. The goal of our hue and texture analysis is to provide a nuanced, robust distinction solely between laryngeal mucosa affected by LPR and that which is not. The clinical utility is based on our method's capacity to be trained in an *a posteriori* fashion using an objective LPR diagnostic standard (i.e., according to class determined by the RFS) and provide subsequent discrimination between larynges with or without the most common signs of LPR. Those which do not exhibit LPR could still have an array of mucosal abnormalities; however, these patients would not have physical signs suggestive of LPR presence in our binary classification model. Thus, image analysis would benefit the clinician who would otherwise make a diagnosis of LPR based solely on subjective interpretation of nonspecific laryngeal signs.

Hue and texture quantification allows for objective visualization of the larynx by creating a quantified color and texture profile, independent of subjective clinical observations. Pace et al. used a pattern recognition approach to identify gastro-esophageal reflux disease.³⁴ The method displayed high accuracy; however, their study relied on 101 clinical variables, many of which were based on patient self-reports. Our method requires only a single laryngoscopic image and includes only objective, quantitative data. Key to the high classification accuracy achieved is the number of parameters included in the analysis (8 hue features and 192 textural features per image). Manual interpretation of this large amount of data would be challenging and timeconsuming at the least; however, pattern recognition with an artificial neural network (ANN) is straightforward and efficient. The ability to synthesize a large amount of information and provide a simple output is a key benefit of machine learning techniques and is relevant to medical decision-making, including the diagnosis of LPR.

CONCLUSION

This preliminary study suggests that a combination of laryngeal hue and texture features could potentially be used to identify laryngopharyngeal reflux. More investigation would be valuable to further assess the classification accuracy associated with the tested physical parameters and other variations of Gabor filtered textural features. Additional research should also focus on the LPR classification accuracy observed by our method when it classifies images based on diagnoses from other objective standards (e.g., pH probe monitoring). The high classification accuracy achieved in this study is encouraging and provides preliminary support that such an approach could be clinically valuable.

Acknowledgments

This study was funded by NIH grant numbers R01 DC05522 and F31 DC012495 from the National Institute on Deafness and other Communicative Disorders and grant number 81028004 from the National Natural Science Foundation of China.

REFERENCES

1. DeVault KR. Overview of therapy for the extraesophageal manifestations of gastroesophageal reflux disease. *Am J Gastroentero.* 2000; 95:S39–S44.
2. Postma GN, Johnson LF, Koufman JA. Treatment of laryngopharyngeal reflux. *Ear Nose Throat J.* 2002; 81:24–26. [PubMed: 12353429]
3. Vaezi MF, Hicks DM, Abelson TI, Richter JE. Laryngeal signs and symptoms and gastroesophageal reflux disease (GERD): a critical assessment of cause and effect association. *Clin Gastroenterol Hepatol.* 2003; 1:333–344. [PubMed: 15017651]

4. Koufman JA, Amin MR, Panetti M. Prevalence of reflux in 113 consecutive patients with laryngeal and voice disorders. *Otolaryngol Head Neck Surg.* 2000; 123:385–388. [PubMed: 11020172]
5. Belafsky PC, Postma GN, Koufman JA. The validity and reliability of the reflux finding score (RFS). *Laryngoscope.* 2001; 111:1313–1317. [PubMed: 11568561]
6. Goldberg HI, Dodds WJ, Gee S, Montgomery C, Zboralske FF. Role of acid and pepsin in acute experimental esophagitis. *Gastroenterol.* 1969; 56:223–230.
7. Koufman JA. The otolaryngologic manifestations of gastroesophageal reflux disease (GERD): a clinical investigation of 225 patients using ambulatory 24-hour pH monitoring and an experimental investigation of the role of acid and pepsin in the development of laryngeal injury. *Laryngoscope.* 1991; 101:1–78. [PubMed: 1895864]
8. Qadeer MA, Swoger J, Milstein C, et al. Correlation between symptoms and laryngeal signs in laryngopharyngeal reflux. *Laryngoscope.* 2005; 115:1947–1952. [PubMed: 16319603]
9. Rees CJ, Belafsky PC. Laryngopharyngeal reflux: Current concepts in pathophysiology, diagnosis, and treatment. *Int J Speech Lang Pathol.* 2008; 10:245–253. [PubMed: 20840040]
10. Ford CN. Evaluation and management of laryngopharyngeal reflux. *JAMA.* 2005; 294:1534–1540. [PubMed: 16189367]
11. Branski RC, Bhattacharyya N, Shapiro J. The reliability of the assessment of endoscopic laryngeal findings associated with laryngopharyngeal reflux disease. *Laryngoscope.* 2002; 112:1019–1024. [PubMed: 12160267]
12. Ulualp SO, Toohill RJ. Laryngopharyngeal reflux: state of the art diagnosis and treatment. *Otolaryngol Clin North Am.* 2000; 33:785–802. [PubMed: 10918661]
13. Sataloff RT, Hawkshaw MJ, Gupta R. Laryngopharyngeal reflux and voice disorders: an overview on disease mechanisms, treatments, and research advances. *Discov Med.* 2010; 10:213–224. [PubMed: 20875343]
14. Fuchs KH, DeMeester TR, Albertucci M. Specificity and sensitivity of objective diagnosis of gastro-esophageal reflux. *Surgery.* 1987; 102:575–580. [PubMed: 3660234]
15. Vaezi MF, Schroeder PL, Richter JE. Reproducibility of proximal probe pH parameters in 24-hour ambulatory esophageal pH monitoring. *Am J Gastroenterol.* 1997; 92:825–829. [PubMed: 9149194]
16. Wiener GJ, Koufman JA, Wu WC, Cooper JB, Richter JE, Castell DO. Chronic hoarseness secondary to gastroesophageal reflux disease: documentation with 24-h ambulatory pH monitoring. *Am J Gastroenterol.* 1989; 84:1503–1508. [PubMed: 2596451]
17. Ahmad I, Batch AJ. Acid reflux management: ENT perspective. *J Laryngol Otol.* 2004; 118:25–30. [PubMed: 14979968]
18. Park KH, Choi SM, Kwon SU, Yoon SW, Kim SU. Diagnosis of laryngopharyngeal reflux among globus patients. *Otolaryngol Head Neck Surg.* 2006; 134:81–85. [PubMed: 16399185]
19. Verikas A, Gelzinis A, Bacauskiene M, Uloza V. Towards a computer-aided diagnosis system for vocal cord diseases. *Artif Intell Med.* 2006; 36:71–84. [PubMed: 16412950]
20. Hanson DG, Jiang J, Chi W. Quantitative color analysis of laryngeal erythema in chronic posterior laryngitis. *J Voice.* 1998; 12:78–83. [PubMed: 9619981]
21. Smith AR. Color gamut transform pairs. *ACM Computer Graphics (SIGGRAPH).* 1978; 12:12–19.
22. Haralick RM, Shanmugam K, Dinstein I. Textural Features for Image Classification. *IEEE Transactions on Systems, Man, and Cybernetics.* 1973; 6:610–621.
23. Daugman JG. Uncertainty Relation for Resolution in Space, Spatial-Frequency, and Orientation Optimized by Two-Dimensional Visual Cortical Filters. *J Optical Soc Am a- Optics Image Sci Vision.* 1985; 2:1160–1169.
24. Bovik AC, Clark M, Geisler WS. Multichannel Texture Analysis Using Localized Spatial Filters. *IEEE Trans Pattern Analysis Mach Intell.* 1990; 12:55–73.
25. Dunn D, Higgins WE. Optimal Gabor filters for texture segmentation. *IEEE Trans Image Process.* 1995; 4:947–964. [PubMed: 18290045]
26. Chan W, Coghil G. Text analysis using local energy. *Pattern Recognition.* 2001; 34:2523–2532.
27. Hubel DH, Wiesel TN. Receptive fields and functional architecture in two nonstriate visual areas 18 and 19 of the cat. *J Neurophysiol.* 1965; 28

28. Campbell FW, Kulikowski JJ. Orientational selectivity of the human visual system. *J Physiol.* 1966; 187:437–445. [PubMed: 5972183]
29. Daugman JG. Complete Discrete 2-D Gabor Transforms by Neural Networks for Image- Analysis and Compression. *IEEE Trans Acoustics Speech Signal Proc.* 1988; 36:1169–1179.
30. Baxt WG. Application of artificial neural networks to clinical medicine. *Lancet.* 1995; 346:1135–1138. [PubMed: 7475607]
31. Gaynor EB. Laryngeal complications of GERD. *J Clin Gastroenterol.* 2000; 30:S31–S34. [PubMed: 10777169]
32. Grontved AM, West F. pH monitoring in patients with benign voice disorders. *Acta Otolaryngol Suppl.* 2000; 543:229–231. [PubMed: 10909027]
33. Habermann W, Schmid C, Neumann K, DeVaney T, Hammer HF. Reflux symptom index and reflux finding score in otolaryngologic practice. *J Voice.* 2012 May; 26(3):123–127.
34. Pace F, Buscema M, Dominici P, et al. Artificial neural networks are able to recognize gastro-oesophageal reflux disease patients solely on the basis of clinical data. *Eur J Gastroenterol Hepatol.* 2005; 17:605–610. [PubMed: 15879721]

Figure 1.
Images of larynges without (A) and with laryngopharyngeal reflux (B).

Figure 2.

Regions of interest: (A) left true vocal fold, (B) right true vocal fold, (C) left false vocal fold, (D) right false vocal fold, (E) left arytenoid, (F) right arytenoid, (G) interarytenoid space, and (H) epiglottis.

Figure 3.
Three-dimensional model of red-green-blue (RGB) color space.

Figure 4.
 (A) Two-dimensional hue wheel indicating perceived color and the corresponding hue angle; (B) three-dimensional model of hue-saturation-value adapted color space.

Figure 5. Effect of Gabor filter angle. (a) Synthetic texture; (b) Gabor filter output at $\theta=0^\circ$; (c) Gabor filter output at $\theta=45^\circ$; (d) Gabor filter output at $\theta=90^\circ$. Line of reference for θ is the horizontal axis of the image.

Figure 6.
Schematic of multi-layer perceptron artificial neural network used in this study.

Figure 7.

A) Receiver operating characteristic (ROC) curve for all parameters, with area under the curve (AUC) of 0.887. B) ROC curve for only hue parameters, with AUC of 0.834. C) ROC curve for only texture parameters, with AUC of 0.852.

Table 1

Comparison of absolute hue values across regions in non-LPR (n=42) and LPR (n=20) subjects.

Region	Non-LPR (°)	LPR (°)
Left true vocal fold	15.88 ± 3.97	17.05 ± 6.35
Right true vocal fold	15.91 ± 3.83	17.16 ± 5.54
Left false vocal fold	11.33 ± 3.15	11.41 ± 2.95
Right false vocal fold	11.10 ± 3.21	11.90 ± 3.16
Epiglottis	10.08 ± 3.33	12.09 ± 3.79
Left arytenoid	10.49 ± 2.85	10.80 ± 3.83
Right arytenoid	10.22 ± 3.12	11.60 ± 4.68
Interarytenoid	10.04 ± 3.22	11.33 ± 4.28

Values are presented as mean ± standard deviation.

Table 2

Summary classification accuracies produced by multi-layer perceptron neural network analysis with varying number of hidden nodes.

Parameters included	Classification accuracy
All parameters	80.52 $\pm$ 1.15
Hue only	73.08 $\pm$ 3.51
Texture only	74.88 $\pm$ 3.62
Left vocal fold	74.04 $\pm$ 2.36
Right vocal fold	73.32 $\pm$ 3.05
Left false vocal fold	73.20 $\pm$ 7.67
Right false vocal fold	71.42 $\pm$ 4.55
Epiglottis	70.36 $\pm$ 3.42
Left arytenoid	75.66 $\pm$ 3.34
Right arytenoid	76.44 $\pm$ 5.39
Interarytenoid space	65.36 $\pm$ 2.45

Values are presented as mean $\pm$ standard deviation. LPR = laryngopharyngeal reflux.

Table 3

Intraclass correlation coefficients (ICC) representing interrater reliability.

Parameters included	Hue	Texture
Total	0.97 ± 0.03	0.85 ± 0.11
Left vocal fold	0.99	0.64 ± 0.12
Right vocal fold	0.98	0.84 ± 0.10
Left false vocal fold	0.98	0.77 ± 0.09
Right false vocal fold	0.98	0.97 ± 0.01
Epiglottis	0.98	0.99 ± 0.00
Left arytenoid	0.96	0.84 ± 0.03
Right arytenoid	0.91	0.86 ± 0.02
Interarytenoid space	0.97	0.86 ± 0.03

Texture values represent the average of the 24 measurements for that region.