Decision Graph Embedding for High-Resolution Manometry Diagnosis

Julian Kreiser, Alexander Hann, Eugen Zizer, Timo Ropinski

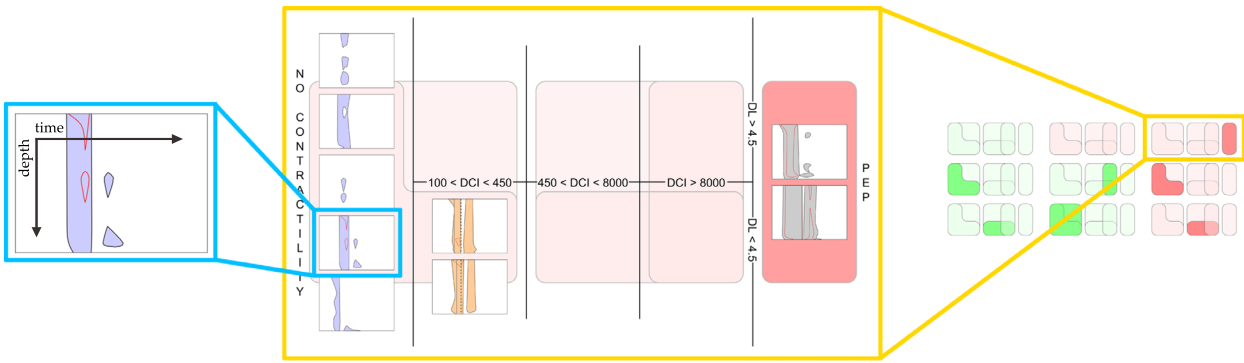


Fig. 1: High-resolution manometry provides large sets of spatio-temporal pressure data along the esophagus during swallowing. The pressure data is required for the diagnosis of motility disorders by the application of aggregated metrics. We alleviate this process by the creation of abstract illustrations of single swallows (left) and their arrangement in a novel, visual representation of a diagnostic decision graph (middle) based on these aggregated parameters. A multitude of patients can also be simultaneously visualized by the usage of an even more simplified version of this decision graph and the creation of small multiples (right).

Abstract— High-resolution manometry is an imaging modality which enables the categorization of esophageal motility disorders. Spatio-temporal pressure data along the esophagus is acquired using a tubular device and multiple test swallows are performed by the patient. Current approaches visualize these swallows as individual instances, despite the fact that aggregated metrics are relevant in the diagnostic process. Based on the current Chicago Classification, which serves as the gold standard in this area, we introduce a visualization supporting an efficient and correct diagnosis. To reach this goal, we propose a novel decision graph representing the Chicago Classification with workflow optimization in mind. Based on this graph, we are further able to prioritize the different metrics used during diagnosis and can exploit this prioritization in the actual data visualization. Thus, different disorders and their related parameters are directly represented and intuitively influence the appearance of our visualization. Within this paper, we introduce our novel visualization, justify the design decisions, and provide the results of a user study we performed with medical students as well as a domain expert. On top of the presented visualization, we further discuss how to derive a visual signature for individual patients that allows us for the first time to perform an intuitive comparison between subjects, in the form of small multiples.

Index Terms—Small multiples, manometry, chicago classification

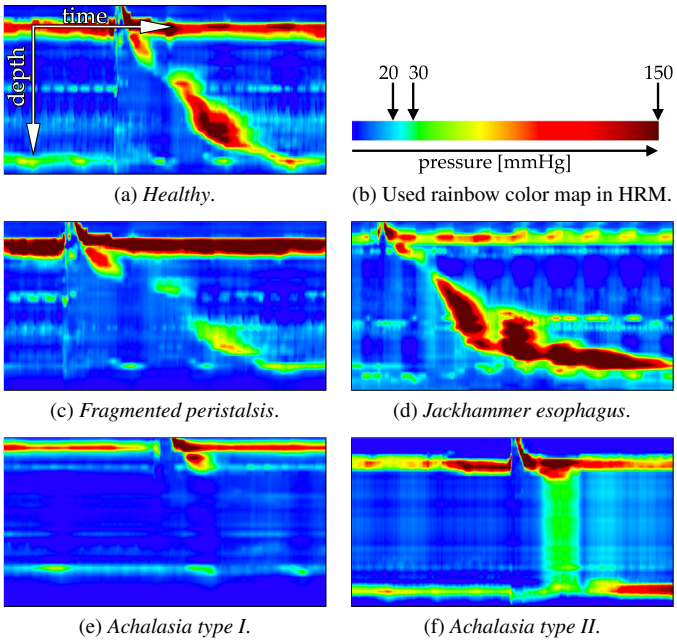
1 Introduction

The esophageal tube presents a connection between the oral cavity and the stomach performing the bolus-transit by well-synchronized propulsive motility during the swallowing process. Imbalance in this system results in dysphagia symptoms and diseases like *achalasia*, *distal esophageal spasm* or *ineffective esophageal motility* [13]. As the frequency of these diseases vary, if untreated, the dysphagia symptoms can lead to a significant weight loss and malnutrition, which makes an early detection essential. Esophageal manometry is the standard-of-care procedure for the diagnosis of such esophageal motility disorders. In today's high-resolution manometry (HRM), electronic solid-state catheters with circumferential conduction of pressure waves allow for the assessment of swallow-activities in the entire esophagus at a high spatial and temporal resolution [16]. The Chicago Classification of motility disorders [13] has been proposed as the gold standard for guiding physicians during the diagnosis.

Within this paper we propose a novel visualization technique for HRM diagnosis based on the Chicago Classification v3.0 (CC3). The proposed approach was motivated by the needs of our medical partners who were unsatisfied with their current diagnosis workflow which is time consuming and requires full concentration. Whereas the CC3 is based on aggregated metrics, current diagnosis systems visualize the individual swallows one-by-one and the medical doctor needs to mentally derive the required aggregates. This proceeding clearly violates the widely accepted concept of *eyes-beat-memory* [21, pg.131]. Therefore, our visualization technique directly communicates these cumulated values to the medical doctor and thus allows for a more effective diagnosis. The visualization of aggregates often implies that the underlying data is more difficult to access, therefore we exploit small multiples to also visualize the individual swallows together with the aggregates. We achieve this by hierarchically subdividing the screen real estate to correlate with the decision nodes of an optimized decision graph, which we have derived from the CC3. To obtain this graph, we have performed a technical analysis of the original decision graph representing the CC3. Since this analysis revealed that the graph does not correctly capture the intended data analysis procedure, the modified decision graph has been developed in a mutual effort of visualization researchers and experienced physicians. It is important to point out, that the new decision graph does not alter, but instead correctly captures the data analysis procedure, as it is verbally described in the CC3 [13]. Based on the modified graph, we later derive the hierarchical screen real estate layout and the embedding of the individual swallows. Furthermore, as

- *Julian Kreiser (Student Member, IEEE) and Timo Ropinski (Member, IEEE) are with the Visual Computing Group, Ulm University. E-mail: julian.kreiser@uni-ulm.de and timo.ropinski@uni-ulm.de.*
- *Alexander Hann and Eugen Zizer are with the Department of Internal Medicine I, Ulm University. E-mail: alexander.hann@uniklinik-ulm.de, eugen.zizer@uniklinik-ulm.de.*

Manuscript received 31 Mar. 2017; accepted 1 Aug. 2017.
Date of publication 28 Aug. 2017; date of current version 1 Oct. 2017.
For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org, and reference the Digital Object Identifier below.
Digital Object Identifier no. 10.1109/TVCG.2017.2744299



(a) *Healthy*. (b) Used rainbow color map in HRM.

(c) *Fragmented peristalsis*. (d) *Jackhammer esophagus*.

(e) *Achalasia type I*. (f) *Achalasia type II*.

Fig. 2: HRM images of characteristic swallows for different disorders. The coordinate frame of the HRM data sets is shown in (a). The widely used rainbow color map, which introduces artificial contrasts, not aligned with important isovalues, is illustrated in (b).

the CC3 implicitly incorporates a severity measure of the diseases to be diagnosed, we were able to design the resulting visualization such, that it perceptually guides the medical doctor during diagnosis based on the importance of the visual entities.

While the proposed visualization has initially been designed with the goal to make the diagnosis of individual patients more effective, we will also show how it can be used to visually compare groups of patients, which was not possible before. To reach this objective, we have developed a simplified visual signature based on the proposed visualization, that - given enough training time - shall enable medical doctors to detect and compare disease indicators among multiple patients by means of a small iconic representation. Thus, we effectively propose a hierarchical visual analysis, which supports the physicians in seamlessly navigating between single swallows, single patients, and groups of patients. Figure 1 illustrates this hierarchical visual analysis.

In the remainder of this paper, we first give a brief introduction into esophageal manometry, where we provide the necessary medical background and discuss the relevant data parameters as well as their relation to certain diseases (see Section 2). Afterwards, in Section 3 we will discuss work related to our approach, before deriving the diagnostic decision graph from the CC3 in Section 4. Our novel visualization, which is based on this decision graph, is proposed in Section 5. There we also justify our design considerations with respect to the data and task at hand. Afterwards we provide technical details about the exploited algorithms in Section 6. To obtain indications about the diagnostic effectiveness of the proposed visualization, we have performed two expert evaluations, as described in Section 7. One evaluation is of a more quantitative nature, where we have asked seven medical students to perform a diagnosis with the original and our visualization. To get also insights from external experts, we have further conducted a formal interview with a medical doctor having more than 10 years of experience in HRM diagnosis. These conducted studies strongly indicate the effectiveness of the proposed visualization, yet they have also brought up a few limitations which we discuss in Section 8. Finally the paper concludes in Section 9.

2 Esophageal Manometry Diagnosis

During a HRM recording, a patient gets an electronic solid-state catheter inserted into the esophageal tube, before he/she has to perform several swallows. Whereas the CC3 states that 10 swallows shall be recorded, in practice it is often more, in order to compensate for data loss or other patient-related factors of influence. Figure 2 shows the recording of such swallows as visualized by today's HRM diagnosis software systems. The x-axis represents time and the y-axis depicts the depth along the esophageal tube, as illustrated in Figure 2a. The color assigned to each pixel encodes the pressure intensity measured in millimeter of mercury, measured in $mmHg$, at each individual time/space coordinate. The individual quantities are color coded with respect to the classical rainbow color scale.

As shown in Table 1, diagnosing HRM data involves the inspection of derived and mostly aggregated data values, points of interest, and structural patterns. The major values defining distinct motility disorders are integrated resting pressure (IRP), intrabolus pressure (IBP), distal contraction integral (DCI), and distal latency (DL) [13]. The diversity of parameters enables physicians to differentiate between several motility disorders in a sensitive and specific manner. The importance of this aspect is emphasized by the fact that different motility disorders require different treatments. According to the severity of the symptoms there are three groups of dysphagia disorders [13] and Figure 3: *achalasia* and *esophageal gastric junction (EGJ) outflow obstruction*, *major disorders of peristalsis* (clinically significant dysphagia disorders), and *minor disorders of peristalsis* (clinical significance is unclear). A clear differentiation between *EGJ outflow obstruction* and *achalasia* has not been defined yet, but this discrimination is important because of different strategies in treatment [24]. The major peristalsis disorders are *hypercontractile esophagus* (or *jackhammer esophagus*), *absent contractility*, and *distal esophageal spasm* (DES). The clinical significance of the minor motility disorders, *ineffective esophageal motility* (IEM), and *fragmented peristalsis*, still remains unclear and an according treatment does not exist.

The visualizations shown in Figure 2 depict individual swallows typical for different disorders, whereas HRM diagnoses requires the incorporation of all swallows recorded during one session, i.e. a median IRP value.

3 Related Work

Within this section, we discuss the work related to our approach. First, we will briefly refer to the current state of the art in manometry visualization, before discussing other related visualization techniques.

3.1 Visualizing Manometry Data

The first established manometry technique was presented in 1968 [7] and utilized up to 12 water-perfused channels with pressure transducers attached to the end, visualized in an ECG-like fashion. Early devices were of low-resolution and their results prone to information loss [10]. With the development of 36-channeled catheters, swallow activity could be recorded in the entire esophageal tube at once [16].

The appearance of electronic solid-state catheters with circumferential conduction of pressure waves has finally led to high-resolution manometry (HRM) [16]. As a result, the swallow-activity in the whole esophagus is captured with less information loss. The high spatial and temporal resolution also results in a more involved diagnostic process, which already today heavily relies on visualization.

The gold standard for HRM diagnosis is defined by the CC3 [13], which defines a workflow for an effective medical diagnosis. Besides this gold standard also alternative approaches exist. A more linearized version of the CC3 has been proposed by Lin et al. [18]. Today, several software systems[1,2] are available which help medical doctors to perform a diagnosis on HRM data. These software systems are extensively used in the medical routine, however, they have to be critically perceived from a visualization point of view due to several reasons. First, they visualize individual swallows as main entities, while aggregates are needed during the diagnosis. Therefore, aggregated values need to be mentally derived. Second, the computed data characteristics are not visualized together with the individual swallows, which hampers validation. Third, all reviewed systems employ the often criticized rainbow color scale [4] to depict pressure in the individual swallows. Despite

[1] ViMeDat, Standard Instruments GmbH, http://www.stdi.de/en/products-service/vimedat

[2] Solar GI HRM, Medical Measurement Systems GmbH, http://www.mmsinternational.com/usa/420/gastroenterology-hrm-products-solar-gi-hrm

	Parameter/condition	Definition	Normal values/interpretation of pathological values
1	IRP [mmHg] (integrated relaxation pressure)	Mean value of the four seconds of maximal deglutitive relaxation in the ten-seconds-window beginning with the relaxation in upper esophageal sphincter.	dependent on specific manometry hardware used (e.g. Sierra device median IRP$> 15mmHg$: outflow obstruction/achalasia/impaired relaxation of the lower esophageal sphincter)
2	DCI [mmHg*s*cm] (distal contractile intergral)	Amplitude*duration*length of the distal esophageal contraction exceeding 20mmHg from the transition zone to the proximal margin of the lower esophageal sphincter.	DCI$<$ 100: failed contraction 100 $<$DCI$<$ 450: weak contraction 450 $<$DCI$<$ 8000: normal contraction DCI$>$ 8000: hypercontractile esophagus
3	DL [s] (distal latency)	Time-period between relaxation of the lower esophageal sphincter and the contractile deceleration point.	DL$<$ 4.5s: premature contraction
4	CDP (contractile deceleration point)	Inflection point on the 30$mmHg$ isobaric line at which propulsion velocity decreases. The CDT-position must be within 3cm proximal the lower esophageal sphincter.	
5	PEP (panesophageal pressurization)	Panesophageal pressure $> 30mmHg$ beginning in the upper esophageal sphincter to the lower esophageal sphincter.	
6	Fragmentation	Large break ($> 5cm$) in the 20mmHg isobaric line with DCI$>$ 450mmHg*s*cm.	

Table 1: Derived data values (rows 1-3), points of interest (row 4), and structural patterns (rows 5 and 6), that need to be considered during a HRM diagnosis. The left column shows the abbreviations used throughout the paper, while the middle column provides a brief explanation. For the derived data values. The right column shows the value ranges critical for the diagnosis.

these flaws, little work has been done by the scientific community in order to improve HRM diagnosis. Parks has introduced a visualization for the analysis for single swallows [22]. The described approach basically follows the visualization of temporal ECG data, whereby differently colored rows are used to depict the signals of the individual sensors. More recently, Parks and Son have extended the visualization by also taking into account the circumferential dimension [23].

The modified barium swallow (MBS), also referred to as videofluoroscopy, and the functional endoscopic evaluation of swallowing (FEES) are two other imaging methods to examine the swallow function in the oropharyngeal area. Brady and Donzelli [5] characterize and contrast these two standard-of-care methods. Aung et al. [1] present a spatio-temporal visualization for the bolus transit during swallowing, generated from profile pictures captured with MBS. Their motivation is the measurement of the time, a swallowed bolus needs to travel past anatomical landmarks. However, these two procedures are solely used in the diagnosis of oropharyngeal dysphagia, our work concentrates on esophageal dysphagia. Due to low sensitivity of the radiologic and endoscopic diagnostics, HRM represents the current gold standard to diagnose achalasia [13].

3.2 Related Visualization Concepts

The HRM visualizations presented within this paper exploit some well established visualization concepts to eliminate the identified downsides of existing approaches. First, we follow the *eyes-beat-memory* paradigm [21, pg.131] to lower the burden of the medical doctors during the diagnosis process. To do so, we exploit small multiples [3], which visually enforce "comparisons of changes, of the differences among objects, of the scope of alternatives" [28]. Small multiples have previously been applied to communicate graph data structures [6, 9], as well as to various other application-related visualization problems, e.g., [11, 20]. Furthermore, it could be shown that they require users to need fewer steps when integrated as a data exploration method [29]. Thus, small multiples often produce a more effective visualization than integrating all data into a single plot [12]. Using small multiples further gives the visualization designer the opportunity to choose an appropriate layout. In general it is beneficial to align them with respect to a given axis, as it allows for a direct comparison of the actual data values. However, also alternative concepts for the layout of small multiples have been presented. Kehrer et al. have investigated the hierarchical nature of small multiples by focusing on different attributes [15]. Thus, they achieve layouts which support comparison based on the attributes of interest. Our approach is similar in that we use the priority of attributes derived from a decision graph to define the major axis of our visualization space. Liu et al. have presented *Correlated Multiples* as an alternative approach, where the plots are placed such that distances in the visualization space reflect their dissimilarities [19]. While this technique is also used to transform a graph, i.e., a proximity graph into a 2D layout, our approach has a more hierarchical nature, as it is based on decision graphs. There are similarities to the space-filling visualization of treemaps [25], but the major difference is that we aim

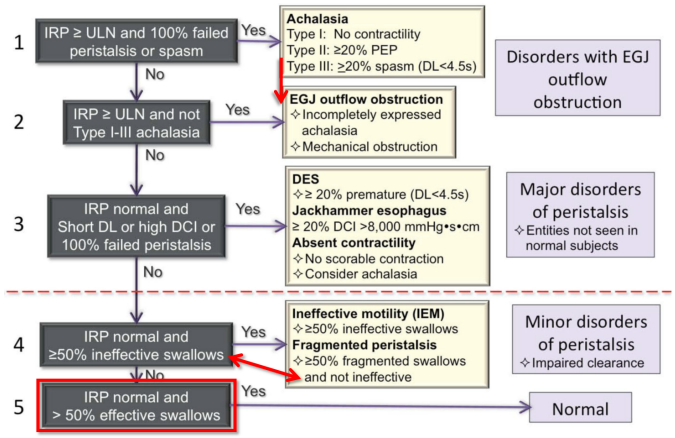


Fig. 3: The official decision graph introduced as part of the Chicago Classification v3.0 [13]. The red arrows depict the identified shortcomings: duplicated reference to IRP, ineffectiveness conflict, and redundant decision node (*from top to bottom*).

at a more sparse representation as we implicitly use white space to encode information. Therefore, our approach is more inline with decision tree visualization techniques [2, 17], which have been shown to be an important instrument when interactively communicating data analysis processes [27]. Furthermore, our visualization has been designed with the goal of interactivity in mind, as it allows for the inspection of individual data sets, while at the same time presenting the aggregated measures in an overview. This makes it comply with the *overview first, details on demand* paradigm [26].

4 Diagnostic Decision Graph

To convey a medical algorithm, decision graphs or trees are often utilized and the same applies to esophageal manometry. However, to be an efficient tool, such representations should be succinct and explicit. Within this section we discuss the currently used hierarchical analysis model of the CC3 and show how it can be optimized to follow a more workflow-centered structure.

The current analysis graph for HRM as provided by the CC3 is a hierarchical diagram shown in Figure 3. At each level from top to bottom, the severity of the potential disorder reduces. Outflow obstructing disorders are considered more critical than peristalsis related disease patterns and have to be identified first. To investigate the conciseness of the CC3 decision graph, we have conducted a technical analysis with visualization researchers and medical doctors, who perform HRM diagnosis on a regular basis.

Workflow optimization. During our analysis we have identified three branching nodes to not be intuitive in terms of following the workflow

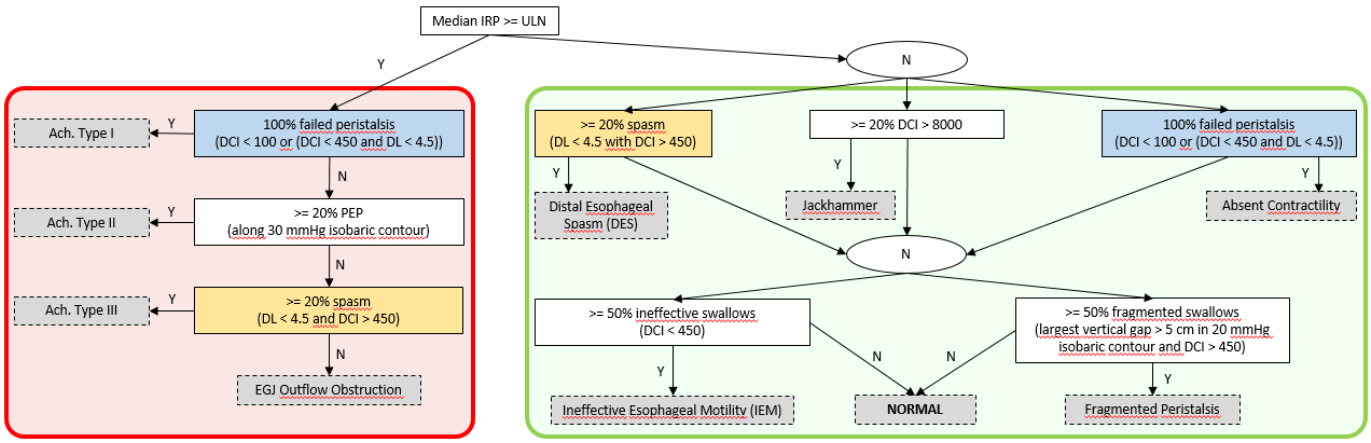


Fig. 4: Our modified decision graph eliminates the problems identified in the original Chicago Classification v3.0 decision graph [13]. Thus, we obtain two main branches based on the IRP (framed *red* and *green* for an elevated and normal IRP respectively) and can guide the physician sequentially, as well as sequentially and parallel, to reflect disease severity within the two main branches. Decision nodes which only differ in the IRP criteria are highlighted in *blue* and *orange* respectively. Resulting diagnoses are accentuated in *gray* with a *dashed* outline.

verbally described by Kahrilas et al. [13]. First, the decision whether the median IRP exceeds the upper limit of normal (ULN) appears in the first and second decision and can be reduced to a single branch-off. Second, the check for *Minor disorders or peristalsis* performed in decision node 4 conflicts with the verbal description of *Fragmented peristalsis*. Whereas decision node 4 requires at least 50% ineffective swallows to proceed to the right, the box on the right in contrast expects that at least 50% swallows are not ineffective. This conflict clearly hampers an intuitive diagnosis process. Third, the last decision node does not have any other alternative than the classification of a patient as normal, or in other words healthy, and thus can be dismissed.

Modified decision graph. Based on the identified shortcomings, we have designed the modified decision graph with respect to the importance of the individual disorders and how parameters affect each branch-off as shown in Figure 4. As defined by the CC3, the three groups of dysphagia disorders are prioritized based on their severity. The median IRP decides between the two major branches in our decision graph. One the left, *achalasia* and *EGJ outflow obstruction*, and on the right, major and minor disorders of peristalsis are separated into two levels of severity. To support readability of the decision graph, we have bundled common paths which is indicated by the elliptical nodes labeled with *N* in the right major branch.

We would like to point out that we do not prioritize within each level in the right major branch, as no explicit ordering is given beside the general severity. An alternative representation is given by Lin et al. [18]. They present a flowchart with strict prioritization for automated diagnosis. As automated diagnosis is not our goal, we would rather like to enable the physicians to bring in their experience when assessing the data. Therefore, we have dismissed the proposed prioritization proposed by Lin et al.

5 Diagnostic Visualization Design

As opposed to the current workflow in manometry diagnosis, which is sequential and tends to a separate inspection of all performed swallows individually, our proposed visualization concentrates all information necessary for an intuitive decision-making into one single representation. Following the current CC3, we guide the attention of the user to the relevant information and supply all important parameters for an efficient diagnosis. To not overload the user with secondary information, abstractions are introduced at different levels of our visualization. Thus, medical doctors can perform a hierarchical, visual analysis which enables them to perform their diagnosis based on aggregates. At the same time it is possible to inspect the individual data sets in detail. Such a coarse to fine approach alleviates the process of diagnosing and simultaneously gives access to every piece of data available. Therefore, our proposed visualization presents an overview of the entire data set of a patient and at the same time allows the user to perform a visualization-guided decision.

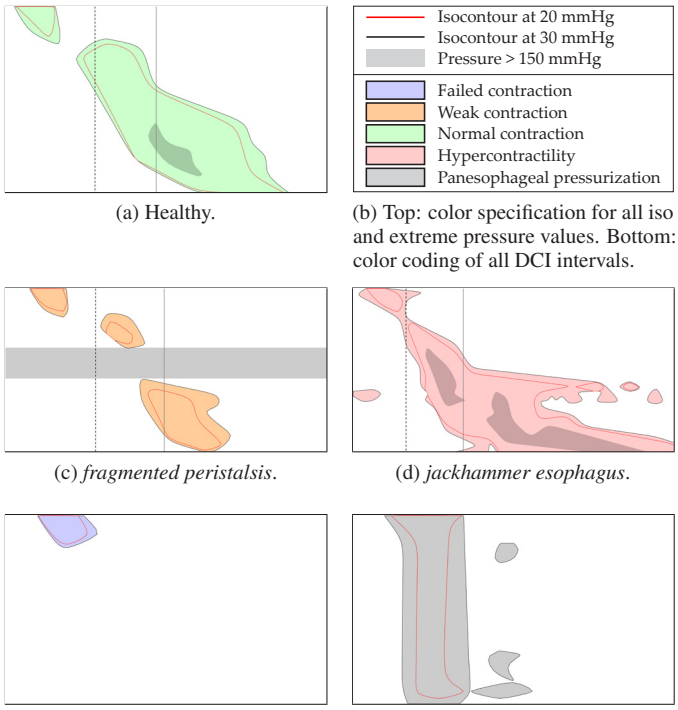
As the CC3 requires analyzing data aggregates, our visualization is based on small multiples, whereby each small multiple represents a single swallow. Thus, we have to integrate multiple swallows into one visualization while avoiding cognitive overload. To do so, we introduce a visual abstraction of each individual swallow, which conserves its main features and still results in a more lightweight depiction. In this section, we will initially discuss our design decisions with respect to this abstraction. What follows is the transformation of the decision graph, presented in Section 4, into a visual representation which acts as a comprehensible tool for the evaluation of a patient, based on the aggregated measures. Finally, for the comparison and display of multiple studies at the same time, a visual signature is derived based on the previously introduced visualization.

5.1 Single Swallow Abstraction

The amount of data accumulating during a HRM procedure is too large to be inspected in detail and at the same time deriving a general idea of the patient's overall condition. To deal with this challenge, we introduce an abstraction of each individual swallow visualization, which conserves typical patterns' characteristics and reduces the user's cognitive overload. While this simplifies the visualization, it still acts as a reminder wrt. the underlying data, as it communicates the general shape of the pressure distribution. This algorithm enables us to create much smaller, but still recognizable versions of each swallow ready to use as small multiples. We will justify the design decisions for the proposed abstraction algorithm within this section. The technical realization of this procedure is detailed in Section 6.

For each of the recorded swallows the derived parameters considered in the CC3 do matter. These are DCI, DL with CDP, PEP, and the fragmentation. The area between the upper and lower esophagus sphincter, starting with the beginning of the swallow, is used to obtain these values. This area also serves as starting point for our visual abstraction. It is also crucial that these parameters remain perceivable, as they may influence the diagnosis. For a detailed description on how we have preserved these values, despite their visual abstraction, we would like to refer to Section 6.

In general, the appearance of a single swallow is determined by its overall shape, defined through various extracted isobaric contours, as detailed in Table 1. This is not only reflected in the CC3, but has also to be visually prominent through the used color scale. In contrast, the exact pressure distribution inside this contour is only of secondary importance. Keeping the visual impression close to familiar structures, users are quickly able to identify common patterns for different types of disorders, even in a more abstracted representation. To create a high contrast with the light background colors used for the DCI, as explained in the next paragraph, we exploit black and red as line colors respectively. Additionally, the area enclosing all values above 150 mmHg is rendered as a shaded version of the base color without an explicit



(a) Healthy.

(b) Top: color specification for all iso and extreme pressure values. Bottom: color coding of all DCI intervals.

(c) *fragmented peristalsis*.

(d) *jackhammer esophagus*.

(e) *Achalasia type I*.

(f) *Achalasia type II*.

Fig. 5: Illustrated versions of the in Figure 2 presented HRM images. The coordinate system is oriented the same way as given in Figure 2a. A key for the colors of all isovalues and DCI intervals is given in (b).

edge surrounding it, as this area is frequently observable in cases of hypercontractility (Figure 5d). The line and color coding used for all isovalues is illustrated in Figure 5a. Examples of four typical swallows for different diseases and one healthy swallow can be seen in Figure 5.

As DCI and DL are most relevant, we communicate this through prominent visual variables. Thus, the DCI defines the color we assign to the entire area inside all isobaric contours. Here, we intentionally avoid using the rainbow color map widely applied in manometry images due to the known shortcomings [4]. In this specific case, especially the introduction of artificial contrasts and features are problematic, as these usually do not align with the relevant isobaric contours, as shown in Figure 2b. Therefore, the rainbow color scheme of the map is not adjusted to match the relevant isovalues used in the manometry diagnosis, i.e., no relevant threshold lies directly between two different hues in the map to distinguish these particular borders.

In contrast, we restrict the color for each individual swallow to a single hue. Thus, the individual DCI values are immediately perceivable, see Figure 5 for examples and a key for what colors mean. Our visualization uses the following colors, all desaturated, to communicate this parameter. To mimic the mostly blue appearance in manometry images in the case of none or low contractility, we use blue as well. An orange tint is utilized in the case of weak contractions as a slight warning color that the overall pressure is too low but at least some motility is present. For normal contractility, we use green as a typical indication that the DCI is in the desired interval. In cases of hypercontractility, red signals a value over the upper boundary of a normal pressurization which is also again similar to the regular manometry image. In cases of PEP, when no DCI should be calculated, we use gray as a neutral alternative to all other colors.

The DL, which serves as the second most important data derived value, is included using two vertical lines. As DL encodes the time from the start of the swallow towards the characteristic CDP point, it is naturally associated with the temporal axis of the plot. As the y-coordinate of this point is not relevant, we have decided to exploit vertical lines. This result in a better visibility and comparison when aligning multiple swallows vertically. Therefore, again for contrast reasons, we use one dashed, black line at the reference value from the beginning of the swallow and a continuous, dark gray line at the actual DL value. The clear depiction of the DL gives the opportunity to judge if the DL is borderline or clearly below or above the given threshold, as exemplified in Figure 5.

The third most important data derived attribute associated with each individual swallow is the fragmentation (see reference Table 1). To communicate this value prominently without interfering with DCI and DL, we introduce vertical gray bars. We have chosen a light gray, as it is well visible, but not too prominent, and thus nicely reflects the relevance of fragmentation with respect to the CC3. In the original manometry images, not involving the display of the vertical gap, the horizontal offset between two pressure zones is also visually perceived, however, fragmentation is only evaluated along the spatial dimension.

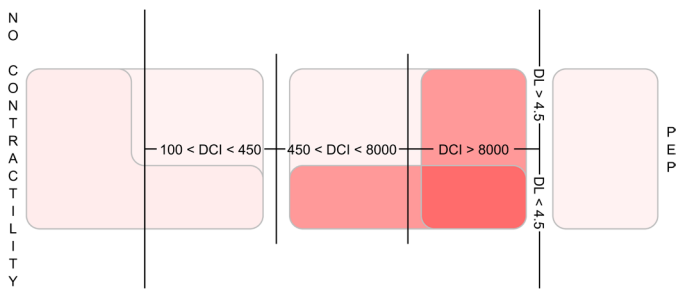
We currently do not include the IRP of each individual swallow in our abstracted representation because it does influence the appearance of the visualization used to communicate the aggregated data values, as explained in the following subsection.

5.2 Small Multiple Embedding

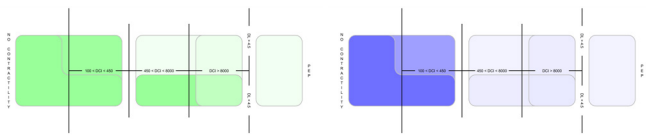
Having an abstracted plot for each individual swallow enables us to integrate multiple swallows into a single visualization, while at the same time still leaving some visual variables to communicate the aggregated values. Thus, the diagnosing physician can focus instead of on the data point inspection on a more complex task to solve, i.e., the diagnosis based on the aggregates as defined in the CC3. We achieve this through an enriched visualization of our diagnostic decision graph proposed in Section 4, whereby we follow the *eyes-beat-memory* principle. The main idea is to embed our abstracted swallow representations into the visualization space, such, that the required aggregate values are visually prominent. Thus, in contrast to conventional small multiples, our visualization exploits the layout and also the white space as main visual communication pillars. We will justify our design decisions with respect to this aggregate visualization within this section, whereby we first detail the layout and its visual parameters before we discuss real-world use cases. The final result is shown in Figure 7.

5.2.1 Layout Overview

Again, our spatial layout is mainly based around the two parameters DCI and DL which appear most frequently within the decision nodes of our decision graph. To effectively communicate these parameters, we have chosen an embedding which uses DCI and DL as the two axis of the visualization space. Within this space, we position each abstracted swallow illustration to represent its particular parameter set (DCI, DL).

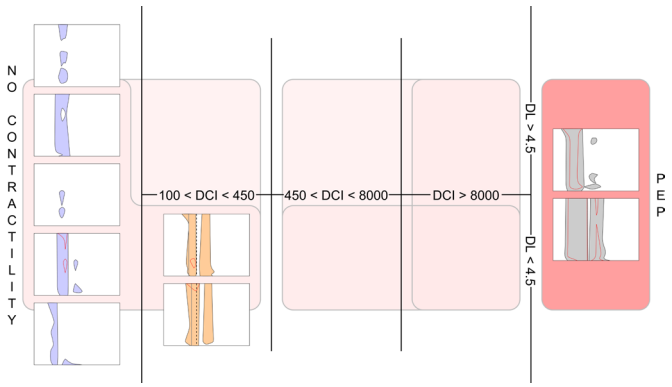


(a) Patient overview with an elevated median IRP, noticeable through the overall red background color. Overlapping decision areas create a stronger intensity in these regions to remain distinguishable

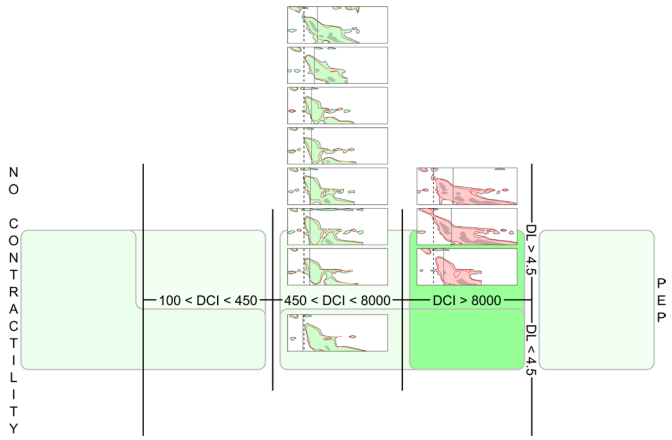


(b) Comparison of the green and blue color selection for users with red-green blindness. On the left, no overlapping decision areas, while on the right the incremental color intensity is noticable and communicates that two separate decision areas are positively evaluated.

Fig. 6: Different overlapping decision areas with recognizable nuances in color intensity to visually separate all involved regions.



(a) Example case of an *achalasia type II* patient. The PEP decision area indicates that at least 20% of all performed swallows fulfill this criteria together with the red background, signalizing an elevated median IRP.



(b) Example case of a *hypercontractile esophagus* patient. The median IRP is normal which is communicated through the overall green decision areas. Three out of ten swallows fulfill the necessary criteria for a *jackhammer esophagus* and therefore, the intensity of the corresponding decision is higher compared to all others.

Fig. 7: Decision graph visualizations of the Chicago Classification v3.0. [13] for two different esophageal motility disorders – *achalasia type II* in (a) and *jackhammer esophagus* in (b). The positively evaluated decision areas in the background are clearly recognizable. The hue of the decision areas is accordingly set, as highlighted in Figure 4. Small, abstracted versions of each individual swallow provide more detailed information about the patient's condition and severity of the case. Labels are provided to communicate all relevant parameter ranges.

Furthermore, we add a spatial region to represent the PEP swallows, for which DCI and DL is undefined. Thus, we obtain a subdivision of the visualization space, which we detail below. The final spatial arrangement with all grid lines and annotations is shown in Figure 6.

Spatial arrangement. Along the horizontal axis, we arrange the varying DCI intervals as given in Table 1, from left to right with ascending values. We have chosen the horizontal axis, as we have more relevant intervals for DCI and thus better fit with common screen aspect ratios. As swallows with PEP should have no assigned DCI value to them, they don't fit directly into one of both axis, either DCI or DL. However, because PEP and DCI are both related to pressure data, we assign the region beyond the largest interval of DCI to the PEP cases. Swallows with no or weak contractility are combined at the most left end of the DCI axis.

Vertically, our visualization space is divided in half, based on the threshold criteria of the DL. The axis proceeds from bottom to top with ascending DL. In cases with no contractility and PEP, the determination of the DL is also not possible as the relevant CDP can not be extracted. Therefore, these cases do not belong to any distinct position along the vertical direction in our arrangement and the horizontal separation line is not shown in these areas, forming one larger, combined cell.

To have a clear separation in our visualization, we introduce grid lines with according annotations to delimit the relevant parameter ranges in certain areas. We clarify that cases of no contractility and PEP do not include information about DL by drawing no horizontal grid lines in these two areas. To further separate them from the grid cells in which the DCI is measurable, vertical lines are drawn further up compared to the center area.

Decision graph encoding. Our proposed decision graph in Section 4 can now be integrated into this established grid layout. Dependent on the criteria a decision node holds, a connected combination of adjacent grid cells is defined, to which we refer as decision areas. These decision areas are represented in the background of our visualization under each corresponding set of grid cells. Different types of disorders can also share such an decision area, as highlighted in Figure 4 using orange and blue inside the respective decision nodes. We have decided to not extend these decision areas over the whole vertical space in our visualization to maintain all necessary information in the field of view of the user. All decision areas displayed under the separating grid lines and annotations are shown in Figure 6. To depict, whether a decision criteria is fulfilled or to which disorder it corresponds, we utilize explicit color coding, as explained in the following paragraph.

Color coding. To communicate the important median IRP in our visualization, we globally color each of the decision areas either red or green as typical severity indicators for an elevated or normal median IRP respectively. We use the median IRP as global color variable since it affects only the first decision between the left and right main branch in our modified decision graph and implicitly illustrates the severity of possible disorders. Figure 4 reflects this color usage as well.

We are aware of the relatively high possibility for red-green color blindness within our users and that the determination of the median IRP would in such a case be hampered. To address this issue, we considered exchanging green with blue (see Figure 6b), but opted for the more self-explanatory and common color pallet of red and green, as it nicely encodes the semantics of *normal* and *too high*. This option can of course be easily changed and set specifically for each user individually.

The intensities of these global IRP hues are determined as follows. If a decision node in our graph evaluates positively, the corresponding decision area gets the global hue combined with a saturation of 75%, whereas in case of a negative evaluation, the saturation is reduced to 25%. This lower limit of 25% of saturation is necessary in the situation when no decision node evaluates to a positive result, the median IRP then communicates either EGJ outflow obstruction or a healthy patient, which has to be signaled to the user. Otherwise all decision areas would be white and thus, the visualization inconclusive. We use 75% as upper limit to blend overlapping areas together with the result of darker shades in the shared region. The result of such blending can be seen in Figure 6. This is especially necessary when one region is a subset of another as illustrated in Figure 6a and 6b. The user needs to distinguish if either the superordinate area is exclusively filled, or additionally to it the embedded subset.

As shown in our decision graph in Figure 4, multiple disorders can also have the same criteria, solely distinguished by the median IRP (*achalasia type I* and *absent contractility*, or *EGJ Outflow Obstruction* and DES). While this results in the same decision area of our aggregated visualization, the overall hue of red or green helps to distinguish the respective diseases.

As positively evaluated decision nodes indicate potential disorders, the stronger saturation in these regions also acts as a focus of attention. When a user has adapted to our layout and can determine the decision node from the corresponding decision area, the evaluation of a patient can be accelerated as only these interesting nodes need to be taken into account to obtain a diagnosis. The same applies to the case in which no decision node is saturated, then the particular background color of

our visualization guides the user to the corresponding diagnosis, either EGJ outflow obstruction or a healthy patient.

5.2.2 Swallow Data Arrangement

We now have detailed how the global layout of our aggregated visualization is computed and how our decision graph can be encoded using spatial and color coding features. One could solely rely on this stage of abstraction and form a diagnosis based on the provided information. However, we also aim at giving the medical doctor direct access to the underlying data, such that he/she sees the whole picture and thus arrives at a more certain diagnosis.

Thus, to enrich this overview of a single patient for the medical doctor, we incorporate the individual swallow illustrations introduced in Subsection 5.1 into this layout. Each swallow of a patient is sorted into the grid cell of which it belongs to, defined by decision areas for the individual values of DCI and DL. If multiple swallows aggregate in one cell, they are stacked on top of each other, sorted by their DL. In the case of a swallow exhibiting no contractility or in cases of PEP, the corresponding illustration is placed around the vertical center in the respective cells without any particular ordering. To not waste too much screen real estate, we have the option of collapsing stacked swallows into each other to compress the vertical extent of our visualization, as shown in Figure 8.

As the *eyes-beat-memory* principle describes, our visualization allows for having all information available at once and thus reduces the cognitive load of the user. In particular, no previously inspected swallow images have to be remembered for a final diagnosis when all of them are available in a compact, single view.

Through the display of all swallows in our overview, the percentage threshold of decision nodes can immediately be identified when following the CC3 evaluation scheme of ten performed swallows per patient. Either 20%, 50%, or 100% thresholds occur in the decision graph. If two, five, or all swallow illustrations reside in one of the grid cells of our visualization is easy to identify. As usually not exactly ten swallows are performed due to complications or an accustomization phase of the patient, usually a mental translation would need to be performed. However, in our visualization the additional color accentuation of each decision area supports, especially in larger or odd numbers of swallows, the diagnosing process by indicating if a decision node has been positively evaluated.

5.3 Visual Signature Design

Nowadays, the HRM diagnosis is possible for single patients, but it is not possible to compare several patients at once. Within this section, we will show how our proposed aggregate visualization can be modified to support such a comparison.

We already communicate the outcome at each decision node in the CC3 graph as graphical backdrop in our proposed visualization, using hue and saturation, depending on each decision. Separating the visual representation of the CC3 decision graph from the additional information of the individual swallows, compact visual signatures can be generated to describe a patient in a very compact, icon-like manner, as shown in Figure 9. To embed all necessary information for a diagnosis into such a small picture, we use the same layout and color scheme of the decision areas as in our aggregate visualization. We do so in the hope that a frequent usage of our visualization helps medical doctors to learn the patterns and thus enables them to perform a diagnosis solely based on the simplified visual signatures. Thus, it would be possible to match distinct visual pattern to a set of potential disorders. This allows users to learn the general diagnosis algorithm using our visualization and simultaneously being able to understand these even more abstract illustrations.

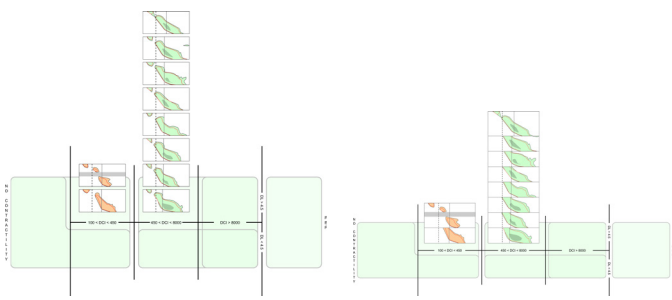
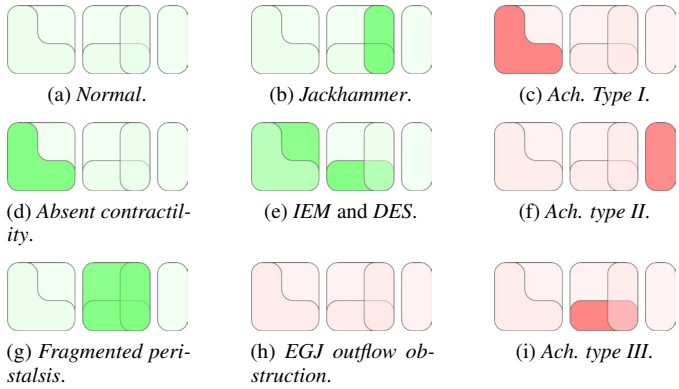


Fig. 8: Side by side illustration of the same data set from a healthy person. On the left side with regular spacing between individual swallows and on the right side with 40% overlap to save screen real estate.



(a) *Normal*. (b) *Jackhammer*. (c) *Ach. Type I*.
(d) *Absent contractility*. (e) *IEM* and *DES*. (f) *Ach. type II*.
(g) *Fragmented peristalsis*. (h) *EGJ outflow obstruction*. (i) *Ach. type III*.

Fig. 9: Small multiple arrangement of our visual signatures. All possible disorders are demonstrated, together with a combination in (e).

These visual signatures are self-contained and can be used to quickly communicate a patient's condition or to generate an aggregated view over many individuals, again in a small multiple arrangement, as demonstrated in Figure 9. If further information is requested, more detailed views on the patient's data can always be supplied in form of our aggregated visualization discussed in Section 5, or even the actual underlying data with full manometry images. Thus, we enable for the first time a hierarchical visual analysis of HRM data.

This level-of-detail approach is very flexible, depending on the situation and availability of electronic devices in the workplace. Mobile versions with step-wise zoom functions between abstractions are imaginable to support doctors. These visual signatures can also be arranged in space and sorted based on other attributes of patients, e.g. age or gender, which could reveal some sort of correlations when clusters of the same type of signature emerge.

6 Implementation

In this section we will provide details for the proposed abstraction algorithm applied to individual swallows. During this abstraction, the processing of each individual swallow consists of three steps, described in the following subsections. The whole pipeline, including the final swallow illustration, is shown in Figure 10.

6.1 Isobaric contour extraction

As isobaric contours play an important role in the classification process, our goal was to preserve the most important contours in each swallow. Different levels of isobaric contours are relevant for the determination of all diagnosis parameters as discussed in Section 2. We have realized the extraction of these by executing marching squares with the respective isovalues, as shown in Figure 10b.

6.2 Area-based contour filtering

Artifacts in manometry images are, amongst others, mainly produced by cardiac activity or respiration. As these usually small patches of pressure fields are not part of the actual peristalsis, we exclude all isobaric contours with a low total pressure from further processing. The threshold value $t_{area} = 50$ in the swallow's spatio-temporal domain has proven to be suitable to sort out these small patches. What has to be considered is that this threshold should not exceed the area of the upper, major pressure zone, otherwise this distinctive feature would also vanish. However, when external forces exert too much pressure, artifacts are preserved. The result after this processing step is illustrated in Figure 10c.

6.3 Feature-aware contour simplification

In order to maintain the general shape and key landmarks of a swallow, we exclude some feature points along each contour from the simplification process, which otherwise would change the later visual appearance of parameters. Therefore, all extrema along the spatial and temporal axis are fixed together with the CDP which defines the DL. This fixation along the vertical axis ensures that gaps between the two major pressure zones, defining fragmented peristalsis, remain unchanged. Usually, the CDP lies along a visually recognizable bend in the peristaltic wave which we also conserve. Overall, the extrema are most of the time located at the upper-left and lower-right of a swallow in addition to the CDP. In conjunction, these key landmarks naturally preserve the general form of the peristalsis, excluding all PEP cases with a pillar-like appearance and often no distinct temporal minimum and maximum in all contour levels. An example of extracted fixation points can be seen in Figures 10c, 10d, and 10e.

Line simplification. To get rid of high-frequency features along each contour which are not part of the general shape, the Douglas-Peucker line simplification algorithm [8] is applied, as shown in Figure 10d.

Uniform sampling. Having the reduced set of points after the line simplification available, we sample this polyline in a uniform fashion which serves as the initial state for the following smoothing step. The uniform distribution of points is beneficial because the subsequent smoothing algorithm works on direct neighbors along the contour and even distances between points prevent point-wise biasing when filtering discrete neighbors.

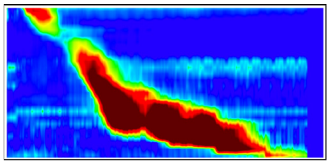
Active contour energy minimization. The simplified version of each contour u consists of an arbitrarily formed shape at this stage without any constraints applied to it. We utilize an active contour model, alternatively called snake, using the physical membrane and thin plate energy terms, $E_{membrane}(u) = \alpha \frac{\partial^2 u}{\partial l^2}$ and $E_{plate}(u) = -\beta \frac{\partial^4 u}{\partial l^4}$, to ensure that our contour continuously anneals to the actual data and simultaneously adopts an overall smooth curvature. To optimize the contour, we apply gradient-descend energy minimization [14]. We include no balloon forces to this optimization as the contour could expand over the previously fixed extrema which would modify the vertical distance between two independent pressure zones, important to determine the severity of fragmentation.

The initial contour u_0 is used as optimization target. We iterate

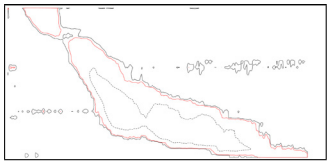
$$u_{t+1} = u_t + \kappa_l \cdot \Delta t \cdot \left((u_0 - u_t) + \alpha \frac{\partial^2 u}{\partial l^2} - \beta \frac{\partial^4 u}{\partial l^4} \right) \quad (1)$$

until the maximum root-mean-square error between the current and previous step, u_{t+1} and u_t, is below the threshold $\varepsilon = 10^{-8}$ to ensure no further changes are happening. For landmarks along the contour which should be fixed in place, we set $\kappa_l = 0$, otherwise $\kappa_l = 1$, to prevent displacement of these features. All points with fixed key landmarks in their neighborhood still take them into account during each optimization step. The time step $\Delta t = 0.1$ denotes the amount of progress between two iterations. The constants α and β are used to weight the influence of the membrane and thin plate energy respectively. Figure 10e illustrates smoothed contours with kept fixation points.

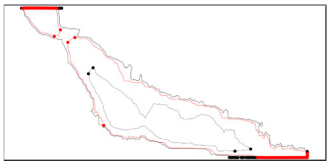
Region clipping. In cases when contours of different isoline levels are too close to each other, the results of the adaptive smoothing of these polylines can overlap. To fix this issue, we clip the contour of the upper isoline against the one directly below. This ensures that regions of lower pressure always surround higher ones.



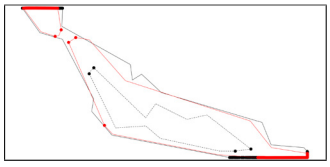
(a) Input data set of a typical swallow for the *jackhammer esophagus* disease pattern.



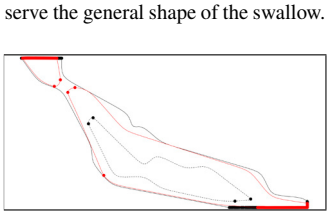
(b) Extracted contour lines, 20 mmHg (*black, continuous*), 30 mmHg (*red, continuous*), 150 mmHg (*black, dashed*).



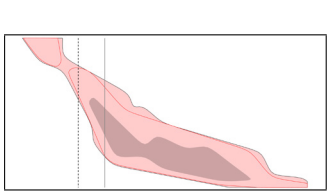
(c) Filtered contours and initial fixation points (*red* and *black* dots) to preserve the general shape of the swallow.



(d) Contours after Douglas-Peucker line simplification.



(e) Smoothed contours after active contour energy minimization.



(f) Final illustration, including DL reference lines and DCI dependent colorization.

Fig. 10: Abstraction pipeline starting witch a HRM data set (a) which simplifies its relevant contours while preserving selected feature points. The final abstracted swallow is shown in (f).

7 Expert Evaluation

In order to evaluate our visualization design, a group of advanced medical students and an independent and experienced expert for manometry diagnosis were consulted. We performed a comparative study to evaluate the capabilities of our visualization and collected qualitative feedback.

7.1 Diagnostic Effectiveness Analysis

To investigate the diagnostic effectiveness of the presented aggregate visualization, we have conducted a quantitative study. We compared our visualization design with the current diagnosis based on the individual swallows, as well as a table showing the derived values for all swallows. The study was designed as within-subject and the task order was balanced using Latin square. Half of the participants started the evaluation using the current diagnosis system, the other half had to use our proposed visualization. We had 7 medical students (of which were 4 male and 3 female) conducting our study, aged 22-32 with an average of 26.2, and from semester 8-13 with an average of 9.8. Due to the fact that they are close to the end of their studies and have acquired all necessary skills, they are going to be introduced to such diagnosis systems more frequently and are therefore well suited as our target group. As the number of participants is below a quantity to be able to perform statistically significant tests, we omit these and informally report the made findings.

All participants got an introduction into both visualizations – despite some of them already knew about the traditional visualization but have not actively learned the diagnostic process itself, it was just mentioned in their studies that manometry diagnosis is performed to record the activity of the esophagus. The introduction for the traditional method was given by a manometry expert, whereupon we introduced our new visualization scheme. Cases of all types of swallows and all final diagnostic decisions were exemplified using both, the old and new decision graphs. During the introduction, questions could be asked at any time by the participants.

We have chosen cases were at least one swallow had a parameter set close to the defined thresholds, given by the CC3, to test if the participants actually pay close attention to each single value. Patients with a number of swallow events different from the officially specified ten were the majority of the cases to examine. As these are not CC3 conform numbers which, nonetheless often occur in real world diagnostic situations, they were used to test the capability of the decision areas as supportive, visual features were decisions based on percentages are needed.

The participants have been confronted with 10 cases to be diagnosed with these techniques, whereby they had to diagnose 5 cases with each technique (*achalasia type I* and *II*, *jackhammer esophagus*, *absent*

contractility, and *healthy*). Based on the obtained diagnosis results, we had to discard the results of one participant, as he has misdiagnosed the majority of all cases for both visualizations. Among the remaining participants, we have analyzed the number of correct diagnosis as well as the timings. Among all participants only one has correctly diagnosed all cases using the traditional visualization. In contrast when using our proposed visualization, three out of six have diagnosed all cases correctly. The one having all correct in the traditional case had also everything correct in our case. Thus, two people, who were correct in our case, made a wrong diagnosis using the traditional approach. When comparing the timings, using our visualization unfortunately takes longer, whereby each case required 46.2 s on average with the traditional and 73.3 s on average with our visualization. Especially, when considering the fact that 3 out of the 6 already knew the traditional method, and that with our visualization we had fewer wrong diagnoses, we believe that this is neglectable. Also considering that 4 out of 6 participants have commented that they believe that the cognitive load when using our visualization for some time would further reduce, we see potential for the presented visualization. Therefore, in general, we believe that the obtained results indicate that our presented visualization has some diagnostic power and that it is helpful for HRM diagnosis.

7.2 Qualitative Evaluation

We consulted an independent expert for manometry diagnosis to gather further information about how our visualization design is perceived. Our goal was to identify which aspects are useful and what has to be adjusted to potentially find its way into everyday clinical practice. In the matter of experience, he conducted at least 330 esophageal manometry procedures in the last decade of which not all have been with the HRM catheter.

Study setup. First, we have generally introduced the expert to our visualization and demonstrated via example cases how to apply it to actual patient data for diagnosing. In a dialog scenario, while conducting additional test cases, we discussed the major topics about: the general layout, the representation of the CC3, the level of abstraction, the look-and-feel, and possible improvements.

Study feedback. Regarding the general layout, we received mixed feedback. Positively commented on was the option to collapse all stacked swallow illustrations, "no scanning of a large area is necessary to get an overview, everything can be overlooked at a glance", but contrary to this statement we were told, that "the plot is far away from the actual anatomy".

When we discussed about the transformation from the CC3 to our visualization, it was stated that "the CC3 is sufficiently represented", and "there is no need to go back to the underlying data". After having seen how to interpret the different decision areas, he mentioned that "it is very obvious how to read the visualization after a couple of introductory cases if someone is familiar with the CC3 itself".

In terms of having an abstract view onto the data, it was said that "an initial hint to a certain diagnosis is helpful, but should not be blindly trusted". Especially in swallow cases which are ambiguous, "he likes it to have to opportunity to revise single swallows on demand and that borderline cases are detectable".

General comments on our visualization were all centered around the point that medical algorithms do not take the ailments of a patient into account. It was said that "additional information beyond the criteria of the CC3, such as an accustoming to the manometry probe or the symptomatology, would be nice to have as an optional overlay".

Towards the end of our evaluation, the expert made statements like "the longer I look at it, the better it gets" and "it seems better than the current raw images, it's very compact". He also pointed out that "he would like to use our novel visualization in his next diagnosis sessions, only when actively using it for an actual diagnosis, the advantages and disadvantages will be truly exposed."

8 Limitations and Discussion

Our conducted evaluations have indicated that our visualizations are favorable as compared to the state-of-the-art, but we have identified limitations which we would like to discuss in this section.

As mentioned before in Subsection 5.2, users with red-green color blindness may experience problems when determining the IRP value in our graph as we employ red and green as common colors to code severity. Currently, we prioritize the merits of expressiveness over the merits of perception. However, alternatively for users which such a condition, green can be switched to blue to counteract this effect. Obviously, this would require awareness about this deficiency, which is not always there.

The screen real estate is in cases when all swallows fall in the same category not optimally used, as large areas of empty space appear. During our qualitative evaluation, experts did not see any problems with this fact. The collapsed version of a column of swallows was perceived as comfortable and it is at the same time a strong indication for one type of disorder. We believe that this unoptimized screen real estate usage is the price we pay for the expressiveness of the proposed visualizations, as clusters of the small multiples directly communicate certain diseases.

Based on the discussion we had with the expert, one issue got exposed rather quickly. Following a given algorithm too closely, a patients's symptomatology is left out of the diagnostic process. Also following the CC3, measurements are only valid if the execution of the study is perfectly controlled (no double swallow, fully lying patient, no external interference), which is not always possible. We currently do not include individual, patient-related information, which might be relevant. Some people just don't get used to the manometry probe and have trouble to swallow only once per measurement. In cases of painful swallow activity, patients sometimes grab themselves at the chest which creates additional pressure not originating from the swallow, which results in artifacts. Such cases could be marked in our visualization to accentuate such data sets. Currently, we do not support such annotations, but do not see why it cannot be done in the future. In general, per swallow annotations would enrich the overall picture as diagnosis should always be a patient-centered process.

Similar to a patient's immediate feedback, the time sequence of all swallows can be of importance. Sometimes, patients have to accustom to the manometry probe and early measurements might show features which typically would not appear in everyday life. Drawing conclusions out of these cases can lead to a wrong diagnosis. Thus, an indication of the order of swallows would be helpful.

9 Conclusion and Future Work

In this paper we have introduced a novel visualization approach for HRM data. The presented visualization is based on the CC3 decision process, which is the gold standard for HRM diagnosis. After clarifying the CC3 decision graph, we were able to generate an aggregate visualization, which directly represents this graph. Thus, in contrast to current diagnostic HRM visualizations, we are for the first time able to communicate the aggregate values relevant in the CC3. Through the introduction of a novel data abstraction technique, we were further capable to integrate the actual data instances into this aggregated visualization. This gives the medical doctor access to individual swallows without obscuring the visualization. Finally, we could show how this aggregated visualization could be transformed into a visual signature which communicates the HRM data for individual patients in a compact manner. Thus, for the first time, it becomes possible to compare the HRM data of multiple patients. Throughout the paper we have justified our design decision, which went into the proposed visualizations, and have demonstrated the applicability based on real-world data sets. Additionally, to analyze the effectiveness of the presented visualizations, we have conducted two expert evaluations.

In the future we would like to further extend the proposed visualizations by incorporating changes reflecting the identified limitations. Foremost, we plan to integrate the described annotation mechanism and thus enable medical doctors to enrich the individual swallows with patient-related data. Additionally, we intend to apply the proposed visual signature to a pool of HRM data available from the last ten years. We will investigate how this large scale visualization can help to spot correlations of the individual data sets to layout the visual signatures. Finally, we plan to conduct a much larger expert evaluation, in which we also investigate learning effects when using our visualization.

Acknowledgments

The authors wish to thank all participants who have performed in the study and all reviewers for their helpful comments.

REFERENCES

[1] M. S. Aung, J. Y. Goulermas, S. Hamdy, and M. Power. Spatiotemporal visualizations for the measurement of oropharyngeal transit time from videofluoroscopy. *IEEE Transactions on Biomedical Engineering*, 57(2):432–441, 2010.

[2] S. T. Barlow and P. Neville. Case study: Visualization for decision tree analysis in data mining. In *infovis*, pp. 149–152, 2001.

[3] J. Bertin. *Graphics and graphic information processing*. Walter de Gruyter, 1981.

[4] D. Borland and R. M. T. Ii. Rainbow color map (still) considered harmful. *IEEE computer graphics and applications*, 27(2), 2007.

[5] S. Brady and J. Donzelli. The modified barium swallow and the functional endoscopic evaluation of swallowing. *Otolaryngologic Clinics of North America*, 46(6):1009–1022, 2013.

[6] M. Burch and D. Weiskopf. A flip-book of edge-splatted small multiples for visualizing dynamic graphs. In *Proceedings of the 7th International Symposium on Visual Information Communication and Interaction*, p. 29. ACM, 2014.

[7] J. Butin, A. Olsen, H. Moersch, and C. Code. A study of esophageal pressures in normal persons and patients with cardiospasm. *Gastroenterology*, 54(4):Suppl–773, 1968.

[8] D. H. Douglas and T. K. Peucker. Algorithms for the reduction of the number of points required to represent a digitized line or its caricature. *Cartographica: The International Journal for Geographic Information and Geovisualization*, 10(2):112–122, 1973.

[9] M. Farrugia and A. Quigley. Effective temporal graph layout: A comparative study of animation versus static display methods. *Information Visualization*, 10(1):47–64, 2011.

[10] M. Fox, G. Hebbard, P. Janiak, J. Brasseur, S. Ghosh, M. Thumshirn, M. Fried, and W. Schwizer. High-resolution manometry predicts the success of oesophageal bolus transport and identifies clinically important abnormalities not detected by conventional manometry. *Neurogastroenterology & Motility*, 16(5):533–542, 2004.

[11] N. Gehlenborg, S. I. O'donoghue, N. S. Baliga, A. Goesmann, M. A. Hibbs, H. Kitano, O. Kohlbacher, H. Neuweger, R. Schneider, D. Tenenbaum, et al. Visualization of omics data for systems biology. *Nature methods*, 7:S56–S68, 2010.

[12] J. Heer, M. Bostock, and V. Ogievetsky. A tour through the visualization zoo. *Commun. Acm*, 53(6):59–67, 2010.

[13] P. J. Kahrilas, A. Bredenoord, M. Fox, C. Gyawali, S. Roman, A. Smout, and J. Pandolfino. The chicago classification of esophageal motility disorders, v3.0. *Neurogastroenterology & Motility*, 27(2):160–174, 2015.

[14] M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active contour models. *International journal of computer vision*, 1(4):321–331, 1988.

[15] J. Kehrer, H. Piringer, W. Berger, and M. E. Gröller. A model for structure-based comparison of many categories in small-multiple displays. *IEEE transactions on visualization and computer graphics*, 19(12):2287–2296, 2013.

[16] B. F. Kessing, P. W. Weijenborg, A. J. Smout, S. Hillenius, and A. J. Bredenoord. Water-perfused esophageal high-resolution manometry: normal values and validation. *American Journal of Physiology-Gastrointestinal and Liver Physiology*, 306(6):G491–G495, 2014.

[17] R. Kohavi and J. D. Tesler. Method, system, and computer program product for visualizing a decision-tree classifier, Aug. 21 2001. US Patent 6,278,464.

[18] Z. Lin, P. J. Kahrilas, S. Roman, L. Boris, D. Carlson, and J. E. Pandolfino. Refining the criterion for an abnormal integrated relaxation pressure in esophageal pressure topography based on the pattern of esophageal contractility using a classification and regression tree model. *Neurogastroenterology & Motility*, 24(8):e356–e363, 2012.

[19] X. Liu, Y. Hu, S. North, and H.-W. Shen. Correlatedmultiples: Spatially coherent small multiples with constrained multi-dimensional scaling. In *Computer Graphics Forum*. Wiley Online Library, 2015.

[20] A. MacEachren, D. Xiping, F. Hardisty, D. Guo, and G. Lengerich. Exploring high-d spaces with multiform matrices and small multiples. In *Information Visualization, 2003. INFOVIS 2003. IEEE Symposium on*, pp. 31–38. IEEE, 2003.

[21] T. Munzner. *Visualization analysis and design*. CRC Press, 2014.

[22] T. Parks. Analysis and visualization methods using manometry data, Jan. 5 2006. US Patent App. 11/129,030.

[23] T. Parks and J. Son. Manometry probe and data visualization, Jan. 22 2009. US Patent App. 11/596,837.

[24] M.-T. Pérez-Fernández, C. Santander, A. Marinero, D. Burgos-Santamaría, and C. Chavarría-Herbozo. Characterization and follow-up of esophagogastric junction outflow obstruction detected by high resolution manometry. *Neurogastroenterology & Motility*, 28(1):116–126, 2016.

[25] B. Shneiderman. Tree visualization with tree-maps: 2-d space-filling approach. *ACM Transactions on graphics (TOG)*, 11(1):92–99, 1992.

[26] B. Shneiderman. The eyes have it: A task by data type taxonomy for information visualizations. In *Visual Languages, 1996. Proceedings., IEEE Symposium on*, pp. 336–343. IEEE, 1996.

[27] S. T. Teoh and K.-L. Ma. Paintingclass: interactive construction, visualization and exploration of decision trees. In *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 667–672. ACM, 2003.

[28] E. R. Tufte. Envisioning information. *Optometry & Vision Science*, 68(4):322–324, 1991.

[29] S. van den Elzen and J. J. van Wijk. Small multiples, large singles: A new approach for visual data exploration. In *Computer Graphics Forum*, vol. 32, pp. 191–200. Wiley Online Library, 2013.