Magnostics: Image-based Search of Interesting Matrix Views for Guided Network Exploration

Michael Behrisch, Benjamin Bach, Michael Hund, Michael Delz, Laura von Rüden,
Jean-Daniel Fekete and Tobias Schreck

Abstract— In this work we address the problem of retrieving potentially interesting matrix views to support the exploration of networks. We introduce Matrix Diagnostics (or MAGNOSTICS), following in spirit related approaches for rating and ranking other visualization techniques, such as Scagnostics for scatter plots. Our approach ranks matrix views according to the appearance of specific visual patterns, such as blocks and lines, indicating the existence of topological motifs in the data, such as clusters, bi-graphs, or central nodes. MAGNOSTICS can be used to analyze, query, or search for visually similar matrices in large collections, or to assess the quality of matrix reordering algorithms. While many feature descriptors for image analyzes exist, there is no evidence how they perform for detecting patterns in matrices. In order to make an informed choice of feature descriptors for matrix diagnostics, we evaluate 30 feature descriptors—27 existing ones and three new descriptors that we designed specifically for MAGNOSTICS—with respect to four criteria: pattern response, pattern variability, pattern sensibility, and pattern discrimination. We conclude with an informed set of six descriptors as most appropriate for MAGNOSTICS and demonstrate their application in two scenarios; exploring a large collection of matrices and analyzing temporal networks.

Index Terms—Matrix Visualization, Visual Quality Measures, Quality Metrics, Feature Detection/Selection, Relational Data

◆

1 INTRODUCTION

In this article, we focus on finding interesting adjacency matrix visualizations for relational data, i.e., networks. Searching and analyzing are key tasks in the process of making sense of large data sets. A widely used approach to implement search and analysis for data relies on so-called feature descriptors (FDs), capturing certain relevant data properties to compute similarity scores between data elements according to these features. Descriptor-based similarity functions are hence a basis for many important exploration tasks, e.g., ranking data elements by similarity or for computing data clusters according to features.

Yet, the choice of feature vectors and similarity functions is a main research challenge; it often requires knowledge of the application context, and sometimes even the user. To date, a large number of feature extraction methods have been proposed for different types of structured data [31, 37]. However, the descriptors these methods use are often defined in a heuristic way, and yield rather abstract information, which are difficult to interpret and leverage by non-expert users in search and analysis tasks. Consequently, it remains difficult to decide which descriptor to choose for a retrieval and analysis problem at hand.

Recently, image-based features have been used to characterize the visual representation of data [10, 22] with the goal to guide the user in the exploration based on the visual representation. Influential for this field is the work of Tukey who formulates the problem that –as the number of plots to interactively inspect increase– exploratory data analysis becomes difficult and time consuming [46]. Tukey proposes to automatically find the "interesting" plots and to investigate those first. To that end, Wilkinson et. al. [46] present a set of 14 measures for the quantification of distribution of points in scatter plots, called Scagnostics. Each measure describes a different characteristic of the data and helps, for example, to filter the views with different Scagnostics measures than the majority. The underlying scatter plots are likely to exhibit informative relations between the two data dimensions. Besides static ranking tasks, image-based data descriptions can also form a basis for dynamic training of classifiers to identify potentially relevant views [5]. This is particularly useful for cases in which a given (static) description and selection heuristic may not fit some user's requirements.

We propose a set of six FDs, called MAGNOSTICS features, which quantify the presence and salience of six common visual patterns in a matrix plot, which are the result of a particular matrix ordering (Fig. 2). Each pattern refers to a topological graph motif, such as clusters, central nodes, or bigraphs. MAGNOSTICS are similar to Scagnostics features describing e.g., the degree of stringyness, clumpiness and outlyingness as relevant patterns in Scatterplots.

Unlike statistical graph measures, which allow describing global graph characteristics, such as density and clustering coefficient, MAGNOSTICS represent interpretable visual features for matrix displays. This is of great importance, because the order or rows and columns in the matrix influences which type of information is visible or hidden from the viewer [4], just like in a 2D layout for node-link representations. Quantifying for one given ordering how well the information is represented in terms of visual patterns helps to assess the visual quality.

MAGNOSTICS can be used for a large variety of tasks, such as finding *good* orderings for visual exploration, finding matrices with specific patterns in a large network data set, analyzing a collection of varied networks, or series of stages in an evolving network (e.g., brain functional connectivity data).

While many FDs for image analysis exist, there is no evidence how they perform for detecting patterns in matrices. In order to make an informed choice of FDs for MAGNOSTICS, we evaluate 30 FDs, including three new descriptors that we specifically designed for detecting matrix patterns. Using a set of 5,570 generated matrix images, we evaluated each FD with respect to four criteria: pattern response, pattern variability, pattern sensibility, and pattern discrimination. For each of the FDs that are part of MAGNOSTICS, we provide a more detailed description, showcasing its performance on real-world data sets. We demonstrate MAGNOSTICS on two application scenarios (Sect. 7). Firstly, querying a large database by example (query-by-example) and via a sketch interface (query-by-sketch). The second scenario analyses a network evolving over time based on time-series of MAGNOSTICS.

- *M. Behrisch, M. Hund, M. Delz are with University of Konstanz, Germany. E-mail: {michael.behrisch, michael.hund, michael.delz}@uni-konstanz.de.*
- *B. Bach is with Microsoft Research-Inria Joint Centre, Saclay, France. E-mail: benj.bach@gmail.com.*
- *L. v. Rüden is with Capgemini and has been with RWTH Aachen University E-mail: E-mail: laura.von.rueden@rwth-aachen.de.*
- *J.-D. Fekete is with Inria, Saclay, France. E-mail: jean-daniel.fekete@inria.fr.*
- *T. Schreck is with the Graz University of Technology, Austria. E-mail: tobias.schreck@cgv.tugraz.at.*
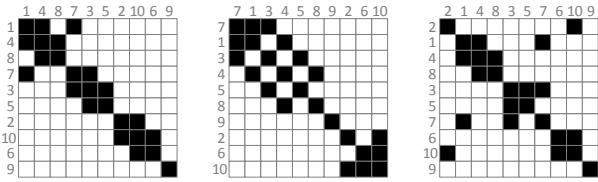
Fig. 1: Different matrix orderings for the same matrix data.

2 RELATED WORK

2.1 Finding Relevant Views in Large Data Spaces

Visual data analysis methods need to be able to handle increasingly large data sets. However, not only the data size grows, but also the possible visualization space for this data. This problem gets even worse when the number of view parameters is taken into consideration. In the case of the analysis of an *n*-dimensional data set with scatter plots, $(n \times (n-1))/2$ two dimensional projections can be produced [40]. If the same data set is visualized with a Parallel Coordinates Plot even $n!$ possible column orderings exist [10]. Similar problems arise in visualization of adjacency matrices, which comprise $(n! \times n!)$ valid row/column orderings. However, all of these visualization approaches have in common that only few view configurations lead to relevant or non-redundant information. Hence, intelligent methods for compressing and filtering data for potential patterns of interest are researched.
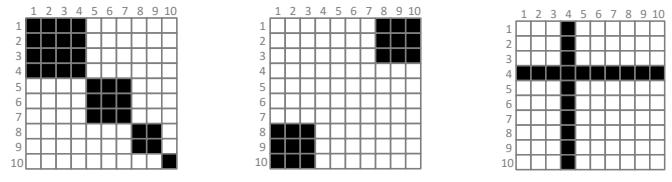
General approaches to support the identification of relevant views in large view spaces include *clutter reduction* [11, 29] and *dimensionality reduction* [20, 41]. Besides fully-automated approaches, others explore interactivity, empowering the user. For example, in [5] an interactive scatter plot exploration approach using a classifier to learn the notion of interestingness from user feedback is proposed. A visual query interface for multivariate data using regressional features is presented in [32]. Alternatively, sketching can be used to express patterns of interest in a large scatter plot space [35]. MAGNOSTICS is meant to be used in similar interactive interfaces, while focusing on quantifying the view relevancy.

2.2 Methods Based on View Quality
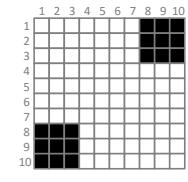
Quantifying the interestingness of visualizations typically requires heuristic feature-based approaches that respond to the (potentially) interesting structural characteristics of a visualization. These methods try to mimic human perception in that they distinguish one or more visual patterns from noise. Several previous works exist, tailored towards specific patterns for certain visualization techniques.

For scatter plots, Wilkinson et al. [46] introduce *Scagnostics* (scatter plot diagnostics), using graph-theoretic measures to describe point distributions. Their feature vector consists of nine interpretable characteristics which are important in the analysis of scatter plots. By using one of these measures, an analyst could make assumptions about inherent information of the described scatter plot. Scagnostics features are global features, describing a whole scatter plot at once. Recently, Shao et al. [36] proposed usage of local features to rank scatter plots. The approach first applies a density-based segmentation of local scatter plot patterns, and then identifies relevant views by an interest measure defined over local patterns.
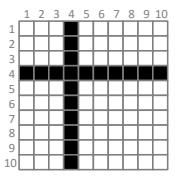
Similarly, Dasgupta et al. [10] propose *Pargnostics* for Parallel Coordinate Plots with the goal to optimize the axis layout so that user's preferences are met. Pargnostics introduces several statistical and image-space measures to quantify e.g., the number of line crossings, crossing angles, convergence, or overplotting measures, all being candidates to rank relevant or informative views. For dense pixel displays, Schneidewind et al. proposed Pixnostics [33], a set of statistical measures in pixel-oriented visualizations. The entropy of an image is measured and shows to be useful to distinguish structured views from noisy ones, reducing the interactive search time for pattern retrieval tasks. In line of this work, Albuquerque et al. present the Noise-Dissimilarity measure for Jigsaw Maps [1], which we also adapted and tested for our matrix pattern analysis scenario. For high-dimensional data analysis,
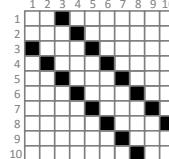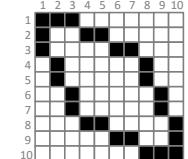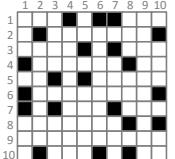

Fig. 2: A Selection of Patterns and Anti-patterns in Matrices.

Bertini et al. [6] proposed a conceptual model for assessing the quality in image spaces and to integrate view quality into the visual exploration process. Finally, Seo et al.'s [34] rank-by-feature framework makes use of correlation and entropy measures to find an appropriate order within histograms, boxplots and scatter plot views.

MAGNOSTICS follows a line of previous works applying view quality selection based on visual view properties. We extend this methodology to matrix visualizations and their visual patterns. Our approach is based on image features extracted for ordered matrices, which are used as a selection or ranking measure to identify relevant views.

3 MATRICES AND VISUAL PATTERNS IN MATRICES

An adjacency matrix $M = (n \times n)$ is a two-dimensional vector with $n = |N|$ with N being the nodes in an undirected network. A matrix *cell* c_{ij}; $0 \leq i, j < n$ denotes the existence of a link between node i and j. In the most simple case $c_{ij} = 1$ if i and j are connected, otherwise $c_{ij} = 0$. Depending on the type of data, c_{ij} can carry particular attributes, such as a strength (e.g., degree of friendship in a social network), or relation type (e.g., friend vs. college) resulting in a weighted adjacency matrix. Here, we focus on *symmetric binary* matrices.

A *visualization* of a matrix is an $n \times n$ plot with each cell representing the underlying value in the matrix using shades, color, or glyphs. Crucial in using matrix visualization and to extract meaningful information is a *good* ordering of rows and columns, indicating the presence of *topological structure* in the data. Fig. 1 shows the same data (matrix), using different orderings for rows and columns.

3.1 Matrix Ordering

The problem of finding a good *matrix ordering* is similar to the problem of finding a layout in node-link diagrams that makes, e.g., clusters and outliers visible pre-attentively. In matrices, visual patterns emerge by re-ordering rows and columns, so that neighboring rows and columns are visually similar. Matrix reordering surveys have been presented by Liiv [23] and more recently, by Behrisch et al. [4].

For a single data set, multiple orderings can be created, but not all ordering algorithms may be equally well suited to make certain structures in the data visible. Little research has been conducted to understand, which algorithm is best suited to reveal certain data characteristics. Hence, it appears useful to generate several matrix orderings and chose the ordering with the most suited results, for the task at hand.

3.2 Visual Patterns in Matrices

A *visual pattern* is visual structure in the matrix that reveals information about the underlying graph topology. Some work has been conducted [4, 27, 45] to collect the main visual patterns in matrix plots, along with their graph-theoretic interpretations. While there is an infinite amount of visual patterns, for MAGNOSTICS we refer to the selection detailed in [4] (Blocks, Off-Diag. Blocks, Star/Lines, Band), where the used visual patterns are also related to user tasks.

Visual patterns are of importance, for example, to biologists searching for cycles in biological process networks, or to sociologists studying social interactions and how people form and connect to groups.

Fig. 2 summarizes the patterns we want to retrieve in matrices. For each of these patterns as shown in Fig. 2, we assume an *optimal* ordering of matrix rows and columns; permuting rows and columns in these examples can make the pattern less evident, though the topological information in the network, i.e. the motif, is still there.

- **Block-Diagonal Pattern:** A Block-diagonal pattern consists of a block (at least 2×2 cells) on the diagonal of the matrix. A diagonal block refers to a densely connected clique in a graph where every node is connected to every other node. Clique motifs can be found in social networks representing groups of mutual friends. A slightly weaker version of a clique, i.e., not every node is connected to every other node, is called a cluster and appears as a block with holes or less sharp boundaries.
- **Off-Diagonal Block Pattern:** An off-diagonal block is visually similar to a diagonal block, but placed at the corners of the matrix not touching the main diagonal (top-left/bottom right in our definition). Off-diagonal blocks refer to bi-cliques in the networks that consist of two sets of nodes where each node from the one set is connected to each node of the second set. Bi-cliques can be found in 2-mode networks, e.g. *persons* related to *cities* they have visited.
- **Star Pattern:** Star patterns in matrices consist of one horizontal and one vertical closed line. However, a line does not have to span the entire matrix. A star pattern corresponds to the star graph motif, i.e., a highly connected node in the network, such as an important airport in a traffic network, or a famous person in a friendship network.
- **Band Pattern:** A band pattern consists of lines parallel to the matrix diagonal, but not being placed *on* the diagonal. A band indicates a path in the network, i.e., a sequence of links from one to another node. Paths can represent a possible way of information transmission in a social network or a sequence of reactions in a biological networks.

While patterns indicate motifs and hence information in the network, we are also interested in patterns conveying algorithm artifacts. Detecting artifacts can help to assess the appropriateness of matrix ordering algorithms and to explicitly hide matrices from the user.

- **Bandwidth Anti-pattern:** The bandwidth pattern is a typical artifact of graph-based matrix reordering algorithms, such as the Cuthill-McKee algorithm. While these algorithms are generally fast, their visual performance greatly depends on the existing/salient ordering within the formed envelope. See [4] for more information.
- **Noise Anti-pattern:** By noise we refer to a random distribution of black cells in the matrix. The more noise a matrix contains, the less information it shows. Throughout the paper we use noise matrices as base-line patterns to assess the quality of a feature descriptor.

4 FEATURE DESCRIPTOR ANALYSIS METHODOLOGY

After the conceptual analysis of relevant structures in matrix views, we next detail on our FD analysis methodology, as depicted in Fig. 3.

4.1 Selecting an Initial Set of Feature Descriptors

The literature on image analysis provides an abundance of image FDs, including such based on color, texture, shape, structure, among other properties. These FDs are traditionally developed and used for processing of real-world images. Our goal is to make an informed selection of candidate FDs appropriate for responding to patterns in matrix visualizations shown in Fig. 2. To that end, we started with an initial set of 27 existing, well-known image measures described in the literature. The main selection criteria are the suitability to respond to our set of visual patterns in matrices, availability, and stability of the respective implementations. We also made sure to include two to five FDs from every different image analysis subdomain [31, 37]: texture descriptors, (localized) color-, edge- and line descriptors, shape-, structure- and contour descriptors, interest point descriptors, and noise descriptors. We include three additional FDs that we designed specifically to respond to patterns in matrix views which by their nature show different
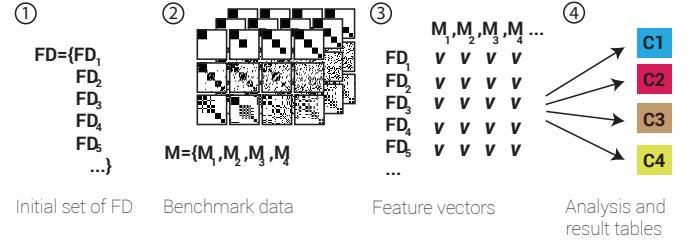

Fig. 3: FD evaluation methodology and main concepts: After a selection of candidate FDs and the generation of a pattern retrieval benchmark dataset for matrices, we evaluated all FDs with respect to four distinct evaluation criteria (C1-C4).

properties than real-world images. For example, our BLOCKS FD not only responds to blocks, but is able to return their number and density (Fig. 2a). NOISE_STATISTICAL_SLIDINGWINDOW quantifies the (local) amount of noise in a matrix (Fig. 2f), while the PROFILE FD is designed to respond to matrices with many lines (Fig. 2c). Table 1 summarizes the full list of FDs we considered for our analysis.

Feature Descriptor	Group	Reference
GLOBAL_COLOR_HISTOGRAM	Color	[31]
AUTO_COLOR_CORRELOGRAM		[19]
FUZZY_HISTOGRAM		[14]
FUZZY_OPPONENT_HISTOGRAM		[42]
COLOR_OPPONENT_HISTOGRAM		[42]
THUMBNAIL		[13]
MPEG7_COLOR_LAYOUT	Color Layout	[21]
LUMINANCE_LAYOUT		[24]
CEDD		[8]
FCTH		[9]
JCD		[8, 13]
EDGEHIST	Edge	Java
MPEG7_EDGE_HISTOGRAM		[28]
HOUGH		[18]
SURF	Point of Interest	[3]
FAST		[30]
BLOCKS	Shape	Java
COMPACTNESS		[26]
ECCENTRICITY		[49]
ADAPTIVE_GRID_RESOLUTION		[48]
JPEG_COEFFICIENT_HISTOGRAM	Structure	[24, 25]
PROFILE		Java
FRACTAL_BOX_COUNTER		[38]
PHOG		[7]
HARALICK	Texture	[15]
GABOR		[24]
TAMURA		[39]
LOCAL_BINARY_PATTERN		[16]
NOISE_STATISTICAL_SLIDINGWINDOW		Java
NOISE_DISSIMILARITY		[1]
GRADIENT		[43]

Table 1: Overview over all tested feature descriptors (FDs). FD names are hyperlinks to access an interactive FD profile page with e.g., a distance-to-noise and a distance-to-base ranking.

4.2 Creating a Matrix Pattern Benchmark Data Set

We tested each FD against the same data set, i.e., visual matrices. An appropriate data set must contain patterns and variations thereof, and different degrees of pattern quality. Moreover, we need multiple samples for the same pattern and its variations in order to account for the variability in the individual data samples. We decided to create an artificial controlled benchmark data set to control for the presence, variation, and quality of patterns, as well as to create as many data samples as necessary. Fig. 4 shows example matrices from the data set. The complete benchmark data set is available online.

Patterns and Variations

For each of the five patterns in Fig. 2, we generated *prototypes* of 30 × 30 matrices and *variations* thereof (Fig. 4 first line). A prototype is an image with only the pattern in an otherwise empty matrix. Any purposeful FD must respond to this pattern. A variation of a prototype is a variation of the general characteristics of the pattern, mainly number, size, and position. For example, a variation of the block pattern varies the number of blocks; a variation of the line patterns changes the line width or the amount of lines, and so forth. The goal is to assess which type of variations a FD is able to differentiate.

Gradual Pattern Degeneration

To measure a FD's ability to respond to unclear and noisy patterns, we generated a gradual degeneration schema for each pattern and its variation, i.e., we gradually decreased the quality of the pattern by introducing "noise" into the matrix. We generated matrices with the following *degeneration* functions, making sure that the last steps of every degradation resulted in a completely random matrix.

a) Point-Swap—A first type of noise function introduces structural noise into the graph, by randomly swapping cells in the matrix (Fig. 4b)). Any cell-swap randomly exchanges one black and one white cell in the matrix (symmetry is preserved by a second corresponding cell-swap in each degeneration step). We can consider the number of cell-swaps as a quality measure how salient a pattern is expressed in the data. For example, a complete block without holes represents a clique in the graph, while the presence of holes indicates a less dense cluster. For our data set, we generated pattern degenerations with the following numbers of cell-swaps: 0, 1, 2, 4, 8, 16, 32 percent of the data, with the last step (32%) showing no evidence of any pattern.

b) Index-Swap—With a second noise function, we iteratively swap rows and columns in the matrix (Fig. 4c)). An index swap preserves the topology of the graph, by randomly exchanging two rows (and columns). Degrading a pattern with an index-swap is similar to using a matrix reordering algorithm that fails to show a topological structure, even though it is present in the data. For our data set, we generated data samples from 0 to 10 index-swaps. We can use the number of index-swaps as quality measure of the respective pattern.

c) Masking—Lastly, we gradually add additional black points (noise) to the matrix to mask the pattern (Fig. 4d)). Thereby, we simulate situations in which the visual pattern is overlapping and intervenes with other patterns, e.g., overlapping clusters, closely connected nodes. The applied noise is exponentially increasing (0% to 16%).

In total we used 23 pattern variations (4 ◨ + 4 ◧ + 6 ▦ + 5 ◪ + 4 ◩) and 24 degeneration types (7 point-swap + 11 index-swaps + 6 maskings). For each condition, we created ten samples, resulting in a set of 5,520 benchmark matrix images with patterns. On top of that, we added 50 pure random noise images with a varying noise density between 1% and 16%, leaving us with a total number of 5,570 MAGNOSTICS benchmark matrix images. The full data set can be downloaded from our website: `http://magnostics.dbvis.de/#/benchmark`.

4.3 Generating Experiment Conditions

For each one of our 5,570 matrix images m_i we run each of our 30 FDs. This resulted in 167,100 trials $FD_j(m_i)$, each one returning an n-dimensional feature vector $v = FD_j(m_i)$. For most FD (except, BLOCKS, TAMURA and some others), vector dimensions may not have a specific human interpretation. In order to analyze and compare all feature vectors we have to rely on *distances* between individual feature vectors $d(v_i, v_j)$. Our distance function is the Euclidean distance between n-dimensional vectors.

Two FDs (ADAPTIVEGRIDRESOLUTION and HOUGH) had to be excluded from the analysis process, since their implementations did not reliably return comparable feature vectors.

4.4 MAGNOSTICS Evaluation Criteria

To inform a selection of FDs, we evaluate each FD according to four criteria. Generally, our criteria selection is inspired by the requirements imposed on information/image retrieval- and computer vision

systems [12]. However, some criteria (esp. C2 and C4) have to be considered as initial research approaches selected specifically to tackle a visual pattern retrieval in matrix data. The individual results are reported in Sect. 5:

C1: **Pattern Response**—*To which specific matrix pattern(s) does the FD respond?* An appropriate FD must distinguish patterns from noise, i.e., a matrix with random distribution of black and white cells. For every pattern in Fig. 2, we can generate a set of prototype patterns and measure the performance by precision and recall for every FD on our entire benchmark data set.

C2: **Pattern Variability**—*How much variation in the pattern can the FD detect?* Patterns in matrices vary, mainly in size or number; for example, there can be one or more blocks, and each of the blocks can have a different size. To measure variability, we can generate variations for every pattern in Fig. 2 and calculate how much the FD response varies/discriminates, using Euclidean distance.

C3: **Pattern Sensitivity**—*How sensitive is the FD to pattern degeneration, e.g., noise?* Patterns are rarely encountered in a prototypical form. For example, blocks may show holes (less dense clusters), or less sharp boundaries (e.g., overlapping clusters). For every pattern in Fig. 2 and its variations, we gradually degrade the pattern until eventually returning a noise matrix (randomly distributed black cells). We develop and derive a pattern sensitivity measure to quantify how well a FD is able to cope with the degeneration of a pattern.

C4: **Pattern Discrimination**—*How discriminatory can the FD differentiate between distinct patterns?* An effective FD should yield discriminative results for different patterns. Otherwise, it does not allow to correctly interpret the FD's response. We measure the pattern discrimination for the FDs by analyzing the differences between vectors returned by the FD.

For MAGNOSTICS, we consider C1 a *decisive* criterion, meaning that we do not want to include FDs into MAGNOSTICS if they do not respond properly to any patterns. Results for C2 and C3 are *descriptive* in that depending on the final use case, a more variable (C2) and/or sensitive (C3) FD may be preferred. Both C1, C4 are considered decisive criteria as we want FDs to discriminate different patterns. In the following, we describe which initial FDs are included in our analysis, how we generated the benchmark data, and how we analyze the returned feature values for each defined experiment.

5 ANALYSIS PROCESS AND QUALITY METRICS

Each criteria (C1-C4) requires an individual analysis of the feature vectors. We report on each criteria individually.

5.1 Pattern Response (C1)

Our pattern response criteria refers to a FD's ability to respond to a specific pattern in the data. As a measure of effectiveness, we use precision and recall measures [2]. High precision reflects a FD's ability to rank a larger number of correct answers at the early ranking positions. High recall means that the FD is able to retrieve a large fraction of all correct answers (matrices) from the target data set. The weighted harmonic mean is an aggregate that combines both measures, but prefers recall over precision (F_2 score, with $\beta = 2$). It is given in Equation 1 and will be used to assess the capability of the FDs to identify the sought matrices.

$$F_\beta = (1 + \beta^2) \cdot \frac{\text{precision} \cdot \text{recall}}{(\beta^2 \cdot \text{precision}) + \text{recall}} \quad (1)$$

Separate precision, recall and F_1 tables can be accessed online[1].

Our experiment setup looks as follows and is repeated for all 30 feature descriptors: First, we derive an appropriate classifier training data set, which consists of all slightly degenerated pattern images ($\leq 4\%$ for point-swap and noise and ≤ 6 index swaps). Second, we conduct a 10-fold cross validation, where each repetition involves the random partitioning of the data into complementary subsets: the training set

---

[1] `http://magnostics.dbvis.de/#/evaluation/patternresponse`
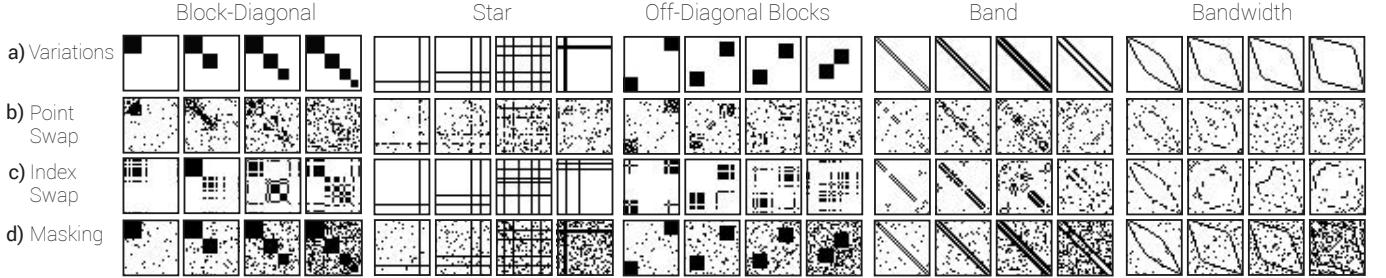
Fig. 4: Examples of matrices as used in benchmark. Benchmark data set can be downloaded at our MAGNOSTICS website.

(70% of the data), which is used to train a classifier, and the validation set (30% of the data) to derive our performance measures. The validation results are averaged over the rounds to reduce the variability. To mitigate the impact of the classifier algorithm we run all experiments with three distinct classifiers (Random Forest, Naïve Bayes, Support Vector Machine). Since the results are comparable we decided to just report on the Random Forest Classifier[2].

We present the results of our pattern response experiments in Fig. 8(a). The F_2 scores are ranging from 0 (low; white) to 1 (high; saturated). As shown in the table, most of the FDs perform with an F_2 score of 0.9 or higher, meaning that these FD are able to differentiate patterns from noise. Generally, texture descriptors yield the highest overall F-scores. None of the patterns were specifically a problem for any of our selected FDs, which means that altogether, our preselected set of FDs is able to respond to each of our six base patterns. Some measures yield higher scores for certain patterns while scoring less for others. For example, MPEG7_EDGE_HISTOGRAM yields lower scores for both block patterns. These results confirm the purpose of an algorithm designed to respond to edges in images. In contrast, BLOCKS performs well for block patterns, but worse for lines.

We will report on the choice of each feature descriptor for each pattern separately in Sect. 6.

5.2 Pattern Variability (C2)

Pattern variability corresponds to how sensitive a FD is with respect to variations in the pattern. Therefore, we decided to take only proto-typical patterns without any additional degeneration (e.g., noise) into consideration for these experiments. Since no scalar measure exists to describe the pattern variability, i.e., there is so far no vector dimension on the number of blocks or lines, we need to look at pair-wise distances between feature vectors and analyze if they allow us to discriminate the pattern variations.

Fig. 5 illustrates all pairwise distances of the test patterns between two selected feature vectors in a distance matrix (Descriptor Distance Matrix, DDM). One row and column represents each pattern variation. The last row contains the noise matrix as point of reference for a distance normalization. Black values mean low distances, red represents medium distances, and white reflects high distances. The example in Fig. 5 shows a desired distribution of distances for the selected FDs (THUMBNAIL and COMPACTNESS), i.e., pairwise distances become gradually higher for more variation of the pattern. Hence, we consider these FDs as good candidates to measure gradually the variability of matrix patterns by distance. For better quantification and comparison, we also report on the standard deviation of the normalized distance scores, as shown below.

$$\sigma_{\text{FD}_i}(\text{PV}) = \sqrt{\frac{1}{|\text{PV}|} \sum_{1 \leq i,j \leq |\text{PV}|} \big(ndist(\text{PV}_i, \text{PV}_j) - \bar{x}\big)^2} \quad (2)$$

Where, PV_i corresponds to the ith pattern variation from a base pattern variation set (e.g., Block1-4), $ndist()$ represents a normalized Euclidean distance and $\bar{x}$ corresponds to the average of all distance

[2]All other experiment result tables can be found in the Appendix.
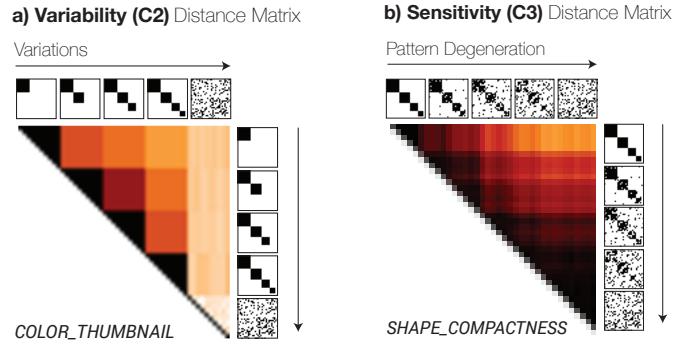

Fig. 5: Distances between two selected feature vectors (THUMBNAIL and COMPACTNESS) of different pattern variations.

combinations. We excluded the reference noise column from this calculation, since it has only illustration purposes.

Other than for precision and recall, poor performance on variability is not a sufficient reason to exclude a FD from the selection. A FD with low variability supports application cases where *any* expression of a pattern is of importance, e.g., finding matrices with block patterns. A FD with high variability supports applications where more detail is necessary, such as finding matrices most similar to a given one.

We report the results of our pattern variability experiments in Fig. 8(b). The standard deviation scores are ranging from 0 (low = white) to 0.5 (high = dark red). From the table we see that the only FD with no variability at all is HARALICK. Other FDs score consistently high for our variability measure meaning that they are able to discriminate between the pattern variations. We also found that variations for the off-diagonal and block pattern are lower for most FDs, which might be due to the fact these coherent rectangles are not recognized independently of their location in the matrix plot (c.f., translation invariance). Line patterns seem to yield generally high variability, suggesting that our FDs are sensitive to changes in the expression of lines.

5.3 Pattern Sensitivity (C3)

Pattern sensitivity refers to how sensitive a FD is to pattern degenerations (e.g., noise), implying a visually less salient pattern. High sensitivity means that only visually salient patterns can be detected, while a low sensitivity allows even less salient patterns to be detected.

Our analysis approach here is similar to the one for C2; we analyze the Euclidean distances between feature vectors of different degeneration levels (Sect. 4.2). The assumption is that with an increase in degeneration, the distances between feature vectors and the feature vectors for the non-degenerated base pattern follow a *monotonic increase*.

Fig. 5(b) shows a DDM for a variation of the Block-Diagonal Pattern (4 blocks) and examples of increasing degeneration with the PointSwap modification function. The matrix shows a monotonic increase in distances from the proto-pattern to the degenerated matrix. For our analysis of sensitivity, we only report a one pattern variation for every base pattern. Tests with the other pattern variations showed only small difference. Eventually, we obtained 150 DDMs, one for every combination of the five base patterns and the 30 FDs.
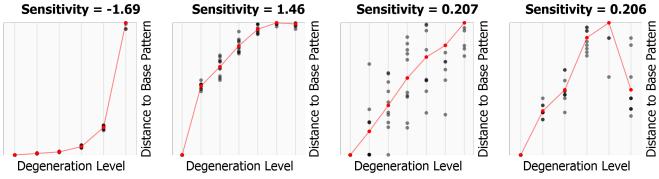
Fig. 6: Relations between degeneration levels and mean distance (red) to base patterns.

In analyzing the monotony we (1) assess if there is any monotonic increase or not. If the increase in not monotonic the respective FD is very sensitive to any degeneration, i.e., the FD cannot accurately rank matrices according to their amount of noise. (2) For monotonic plots, we can then quantify the sensitivity of the FD: highly sensitive (fast increase of distances) or tolerant (slow increase of distances).

Fig. 6 shows examples of monotonic increases of distances. The *x*-axis shows the degeneration level, the *y*-axis shows the distance to the average of one base-pattern vector. Each point in the plot represents the distance of a sample. The red points and line represent average distances. Fig. 6 (a) shows low sensitivity, (b) high sensitivity, (c) medium sensitivity but with a high variation in the distances, and (d) non-monotonically increasing distances.

To quantify monotony of the 150 DDMs and to better inform our final choice, we calculated a quality metric for the monotony of a feature descriptor FD_i and a base pattern P_j as follows

$$monotony(FD_i,P_j) = \begin{cases} 1, & (mean_{l+1} - mean_l + 0.05) > 0; \\ & \forall l \in noise\_levels. \\ 0, & otherwise \end{cases}$$

where $mean_l$ indicates the average of feature vectors retrieved for the noise level k. For monotonic increases, we then calculated the sensitivity as the *signed difference* between the averaged vector for each degeneration level, and the normalized equality function $x = y$. $sensitivity(FD_i,P_j)$ is high for $sensitivity > 0$ (fast increase), and low (slow increase) for $sensitivity < 0$.

$$sensitivity(FD_i,P_j) = \sum_{l=0}^{|noise\_levels|-1} \left((mean_k - (\frac{l}{|noise\_levels|})) \right)$$

We show the results of our pattern sensitivity analysis for the point-swap modification function in Fig. 8(c). Cross signs mean that the respective FD did not yield a monotonic increase of distances. The entire overview of our results for all degeneration functions and base patterns, as well as every distance distribution, can be explored on our interactive website[3].

In summary, most feature vectors show low sensitivity (i.e., are tolerant) to our masking function. However, for structural degeneration (point-swap and index-swap), there are significant differences in the sensitivity of all FDs. In most cases, a FD is tolerant for all or for none base pattern. In our analysis, we found that MPEG7_COLOR_LAYOUT and GABOR performed best across all modification functions.

5.4 Pattern Discrimination (C4)

Our last criteria focuses on how much an individual FD is able to discriminate different patterns, i.e., blocks, lines, etc. High discrimination means that an FD allows distinguishing between patterns. The respective DDM must show overall high values.

To measure discrimination for a given FD, we investigate pairwise distances between all feature vectors v for every FD, across all patterns. In order to obtain a single measure for all the pairwise differences per FD, we report on the median values of distances (we cannot assume distances to be normally distributed), as well as their scatter. High median distances indicate high distances among feature vectors for different patterns, suggesting a high ability to discriminate patterns. High scatter means that this FD can discriminate only a few patterns.

[3] `magnostics.dbvis.de/#/evaluation/patternsensitivity`
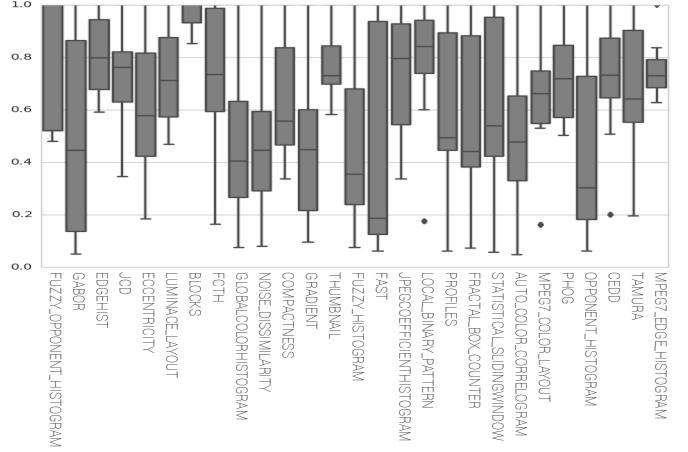
Fig. 7: Pattern Discrimination (C4): High mean values indicate high distances between feature vectors for the individual patterns (CI=95%).

Fig. 7 shows box plots with 25% and 75% quartiles for all FDs. A median distance of 0.0 would mean that FDs, across patterns, look all the same and hence this particular FD cannot represent the difference between patterns. A value of 0.5 means that FDs for different patterns are different by 50%. Our results suggest that most measures are capable of producing different vectors for different patterns. High discrimination is achieved with BLOCKS, EDGEHIST, and LOCAL_BINARY_PATTERN. Small spreads are found for BLOCKS, THUMBNAIL and MPEG7_EDGE_HISTOGRAM. FDs with smaller discrimination and high spread (e.g. FAST or GABOR) may still show differences between individual patterns.

6 FEATURE DESCRIPTORS FOR MATRIX PATTERNS

From our experiments, summarized across all criteria in Fig. 9, we can now choose a purposeful subset of FDs that allows us to describe our matrix pattern selection. The exact choice of FDs required for a specific application may vary, but starting with our set of selected FDs –the MAGNOSTICS set– is expected to be a suitable fit. We disregarded FDs, which did not perform well with respect to our criteria C1-C4 or were outperformed by another FD for the same pattern.

We generally found that some patterns, e.g., blocks or off-diagonal blocks, appear to be easier to detect for our FDs than others, e.g., bands and bandwidths. This may be due to the fact that these patterns are (mostly) positioned around the diagonal/counter-diagonal. Especially, for lines our FDs allowed for little translational invariance (shifts in the x-/y- positions). For our selection we also found that we should incorporate trade-off considerations between pattern response (C1) and the pattern variability (C2), since for some patterns more variations can be expected then for others.

In summary, we selected six FDs for MAGNOSTICS that we found most purposeful according to our experimental comparison in the preceding section. For each FD we briefly describe its name, functionality, which patterns it detects, and report on variability (C2), stability (C3), discrimination (C4).
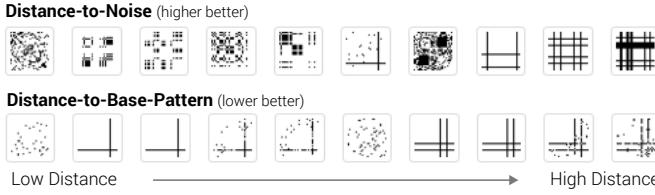
Block Descriptor ◼

We designed the BLOCKS descriptor as a heuristic to measure the *block-iness* of matrix plots *along* the diagonal, thus allowing us to retrieve matrices with a Diagonal-Block matrix form ◼. The BLOCKS FD, shows to be good in all the experiments (C1-4) with a F_2 score of 0.98, a high variability of 0.48 and a good sensitivity score of 0.68 for all degeneration functions.
In a sliding window approach, the descriptor linearly scans for rectangles of a minimum size (width/height) and density (black to white ratio). Found blocks are iteratively enlarged if either the horizontal, vertical and diagonal direction leads to a block density increase. In a post-processing step we remove all blocks, which are fully covered by other blocks. In contrast to the minimum density factor, we also calculate a factor describing the separatedness/distinctiveness of a block

a) Pattern Response (C1)

	SURF	FAST	STATISTICALSLIDINGWINDOW	HARALICK	EDGEHIST	GLOBALCOLORHISTOGRAM	HOUGH	BLOCKS	COMPACTNESS	ECCENTRICITY	ADAPTIVEGRIDRESOLUTION	JPEGCOEFFICIENTHISTOGRAM	PROFILES	DISSIMILARITY	GRADIENT	AUTO_COLOR_CORRELOGRAM	CEDD	FCTH	FUZZY_HISTOGRAM	FUZZY_OPPONENT_HISTOGRAM	GABOR	JCD	LUMINACE_LAYOUT	MPEG67_COLOR_LAYOUT	MPEG67_EDGE_HISTOGRAM	OPPONENT_HISTOGRAM	TAMURA	THUMBNAIL	LOCAL_BINARY_PATTERN	FRACTAL_BOX_COUNTER	PHOG
Block	0.00	0.40	0.98	0.97	0.96	0.98	0.00	0.98	0.96	0.69	0.00	0.76	1.00	0.00	0.00	0.78	0.75	0.75	0.97	0.70	0.83	0.88	0.99	1.00	0.80	0.66	0.95	1.00	0.92	0.64	0.92
OffDiagBlock	0.00	0.23	0.96	0.96	0.95	0.93	0.00	0.89	0.95	0.60	0.00	0.71	1.00	0.00	0.00	0.84	0.73	0.73	0.93	0.78	0.82	0.86	0.78	1.00	0.72	0.70	0.96	1.00	0.94	0.70	0.88
Band	0.00	0.45	0.93	0.94	0.89	0.93	0.00	0.00	0.94	0.70	0.00	0.87	0.95	0.00	0.00	0.87	0.89	0.34	0.94	0.74	0.83	0.92	0.81	0.96	0.92	0.72	0.99	0.99	0.94	0.70	0.76
Lines	0.63	0.54	0.97	0.95	0.94	0.96	0.63	0.65	0.97	0.68	0.63	0.88	1.00	0.00	0.00	0.86	0.90	0.68	0.96	0.28	0.86	0.89	0.71	0.99	0.97	0.74	0.97	0.98	0.97	0.80	0.82
Bandwidth	0.00	0.37	0.96	0.98	0.90	0.94	0.00	0.00	0.98	0.00	0.00	0.84	0.96	0.00	0.00	0.94	0.81	0.44	0.94	0.00	0.87	0.83	0.96	0.97	0.97	0.81	0.99	0.99	0.94	0.73	0.79
Noise	0.00	0.00	0.62	0.83	0.08	0.91	0.00	0.00	0.92	0.00	0.00	0.01	0.41	0.00	0.00	0.29	0.48	0.00	0.90	0.00	0.35	0.56	0.04	0.45	0.22	0.45	0.77	0.41	0.52	0.55	0.00

b) Pattern Variability (C2)

	SURF	FAST	STAT.SW	HARALICK	EDGEHIST	GLOBALCOLOR	HOUGH	BLOCKS	COMPACT.	ECC.	ADAPT.GRID	JPEGCOEFF	PROFILES	DISSIM.	GRADIENT	AUTO_COLOR	CEDD	FCTH	FUZZY_HIST	FUZZY_OPP	GABOR	JCD	LUMINACE	MPEG67_COLOR	MPEG67_EDGE	OPPONENT	TAMURA	THUMBNAIL	LBP	FRACTAL	PHOG
OffDiagBlock	0.40	0.43		0.35			0.38	0.46	0.37			0.37	0.35			0.47	0.48	0.49	0.33			0.41	0.45	0.37	0.46	0.45	0.44	0.38	0.46	0.35	0.37
Bandwidth	0.45	0.40		0.41	0.35			0.34				0.44	0.37	0.35	0.45	0.34	0.48	0.48	0.35		0.39	0.42	0.42	0.36	0.42	0.33	0.36	0.47	0.36	0.43	0.45
Block	0.34	0.34		0.37	0.35		0.49	0.35	0.36			0.43	0.36	0.36	0.45	0.35	0.39	0.38	0.35	0.37		0.35	0.35		0.39		0.35	0.34	0.36	0.34	0.41	0.42
Band		0.30		0.32	0.31			0.30	0.46			0.35	0.31	0.31	0.31	0.33	0.37	0.35	0.31		0.31	0.34	0.34	0.33	0.42	0.31	0.31	0.33	0.34	0.31	0.36
Lines	0.43	0.31		0.33	0.31		0.45	0.32	0.33			0.31	0.28	0.29	0.35	0.35	0.38	0.44	0.33	0.41	0.31	0.35	0.38	0.30	0.36	0.34	0.31	0.31	0.30	0.28	0.32

c) Pattern Sensitivity (C3) Point Swap

	SURF	FAST	STAT.SW	HARALICK	EDGEHIST	GLOBALCOLOR	HOUGH	BLOCKS	COMPACT.	ECC.	ADAPT.GRID	JPEGCOEFF	PROFILES	DISSIM.	GRADIENT	AUTO_COLOR	CEDD	FCTH	FUZZY_HIST	FUZZY_OPP	GABOR	JCD	LUMINACE	MPEG67_COLOR	MPEG67_EDGE	OPPONENT	TAMURA	THUMBNAIL	LBP	FRACTAL	PHOG
OffDiagBlock	X	0.61	0.97	0.83	0.84	X	X	1.15	0.62	2.49	X	X	0.58	X	0.75	0.77	1.09	1.03	X	2.26	0.03	1.10	0.63	0.10	1.23	0.26	0.88	0.91	0.92	1.62	0.35
Bandwidth	X	X	X	1.20	1.45	X	X	X	1.21	X	X	X	2.18	1.89	X	X	X	X	X	X	0.41	X	1.32	1.09	1.50	X	X	1.31	X	1.69	X
Block	X	0.90	1.04	NA	1.25	X	X	0.68	0.86	2.28	X	2.15	0.72	X	0.21	1.02	1.34	1.30	X	2.19	0.11	1.38	1.01	0.11	1.17	X	1.04	1.07	1.17	1.97	0.75
Band	X	X	1.58	NA	1.39	X	X	X	X	X	X	1.43	X	X	0.64	X	X	X	X	X	X	1.23	1.40	X	X	1.68	1.53	X	1.61	1.16	
Lines	X	X	1.59	NA	X	X	X	X	1.46	X	X	1.18	X	X	1.57	X	X	X	X	X	X	1.65	0.79	0.94	1.87	X	1.58	1.46	1.85	1.92	X

Fig. 8: Analysis result overview for all FD and patterns: (a) C1: numbers present the F2-score; higher scores (darker) indicate better response to the pattern, (b) C2: high values (darker) indicate higher variability between patterns, (c) C3: low values (darker) indicate lower sensitivity.

Fig. 9: Result table showing all values. Colors correspond to criteria (blue=C1 (response), red=C2 (variability), brown=C3 (sensitivity)). Black dots in each colored rectangle indicate our subjective ranking; 3 dots represent high ranks, 0 dots indicate low ranks.

and retain only those blocks which are perceptively distinct from the surrounding.

Distance-to-Noise (higher better)

Distance-to-Base-Pattern (lower better)

Low ⟶ High

Local Binary Pattern Descriptor

The LOCAL_BINARY_PATTERN (LBP) texture descriptor [16] is classically used for background-foreground detection in videos. In our experiments the descriptor showed to be responding to the *off-diagonal block patterns*. In comparison to our BLOCKS FD, which is designed for blocks along the matrix diagonal, LBP adds an additional off-diagonal component to the MAGNOSTICS feature vector. It performs well for the C1 and C2 experiments, with F2 score of 0.9 and a high variability of 0.46, but appears to be sensitive to index swaps. An alternative choice for Off-Diagonal Blocks can be the TAMURA textual FD, which is significantly more sensitive to noise.

In a sliding window approach the FD constructs histograms of pixel intensities, called local binary patterns (LBP) for a central pixel to N neighboring pixels.

Distance-to-Noise (higher better)

Distance-to-Base-Pattern (lower better)

Low Distance ⟶ High Distance

Profile Descriptor

We designed the PROFILE descriptor with the aim to describe *lininess* characteristics (many/few short/long lines) in matrix plots. In our experiments the FD responds with a perfect F_2 score of 1.0, and distinguishes clearly between the base pattern variations, thus leading

to a quite low variability score of 0.28. As all other FDs it reacts moderately to noise. However, C1 and C2 make this FD especially suited for query-by-example search tasks. The PROFILE FD computes two axis-aligned histograms of the plot, where every matrix row, respectively column, represents one histogram bin and the bin's value corresponds the number of black pixels within the respective row. In order to achieve translation invariance (i.e., an otherwise empty matrix with just one row/column line should be equally scored independent of the line's location) we are computing a standard deviation from the profile histogram with the intuition that matrix plots with many lines will show high values, while nearly empty matrices or highly blocky matrices will show low values (few jumps).

Distance-to-Noise (higher better)

Distance-to-Base-Pattern (lower better)

Low Distance → High Distance
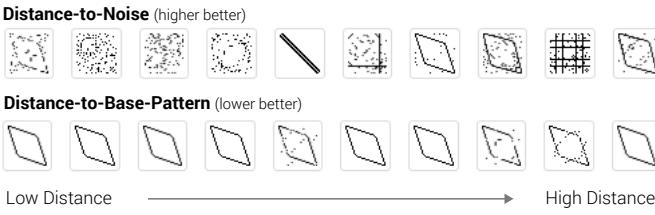
MPEG7 Edge Histogram ⬘

The MPEG7 Edge Histogram FD [28] can be used to retrieve lines along the diagonal ⬘. It responds most often to the band pattern (F_2 score of 0.92) and allows for some variability (0.42). As all other FDs, MPEG7_EDGE_HISTOGRAM FD has problems to deal with pattern degenerations, which might be due to the high specificity of the pattern. The FD subdivides an image into 4×4 sub-images. From each sub-image an edge histogram (5 bins with vertical, horizontal, 45-degree diagonal, 135-degree diagonal and non-directional edge types) is extracted [47].

Distance-to-Noise (higher better)

Distance-to-Base-Pattern (lower better)

Low Distance → High Distance

Color and Edge Directivity Descriptor (CEDD) ⬙

The CEDD descriptor [8] showed a good response to the bandwidth pattern. It incorporates color and texture information in a histogram form. While the texture information uses the same MPEG7_EHD implementation as described above, the color histogram is constructed from an adaptive binning of the HSV color space.

Distance-to-Noise (higher better)

Distance-to-Base-Pattern (lower better)

Low Distance → High Distance

Although, CEDD does not have the best scores in our MAGNOSTICS evaluation (F_2 of 0.81, Variability of 0.48), we decided to include the CEDD FD. Similar to the closely related FCTH, CEDD outperforms the other FDs in terms of variability, which is an important factor for bandwidth patterns (⬙). These patterns are often the result of graph-based reordering methods (e.g., the Cuthill-McKee matrix reordering), which enumerate row-/column permutations in a breath-first search leaving an envelope shape behind.

Haralick Texture Descriptor ▦

The HARALICK FD [15] is one of the classical texture descriptors for images. It responds quite reliably to the noise anti-pattern ▦. For
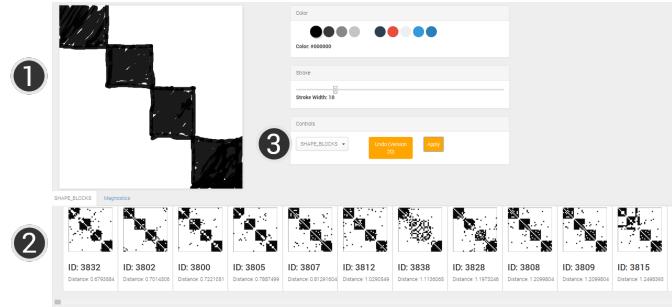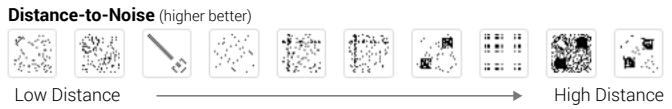

Fig. 10: Query-By-Sketch interface for exploring large collections of matrix plots. The user can sketch in the canvas (1) an approximated matrix pattern and retrieve a ranked result list (2) according to a selected MAGNOSTICS FD (3).

this pattern we conducted only the C1 experiments, since C2-C4 are not meaningful. We decided to include Haralick even though its F_2 score is only 0.83, which is less accurate than three other color intensity approaches (GLOBAL_COLOR_HISTOGRAM,
FUZZY_HISTOGRAM,
COMPACTNESS). However, Haralick is generally more expressive and reliable allowing to quantify a greater variety of noise levels.

We also experimented with our own STATISTICAL_ SLIDINGWIN-DOW noise descriptor, which regards the sliding window values as a time series of differences for subsequent regions in the image. On this time series we calculated average, variance, and standard deviation.

Distance-to-Noise (higher better)

Low Distance → High Distance

7 APPLICATION SCENARIOS

In this section, we report on two use cases that apply the MAGNOSTICS FDs to support exploration of large matrix view data, specifically searching in a database of networks and matrices, and analyzing network changes over time. In both cases, MAGNOSTICS helps to retrieve information (matrices, or changes in a series of matrices) relevant to the respective task.

7.1 Searching in Collections of Networks

Large collections of networks may occur in many applications. As an example, von Rüden reports in [44] about existing matrix collections from the high-performance computing domain with more than 290,000 matrix plots with the analysis goal to retrieve matrices with similar patterns. Another example of large matrix collections is the results of the cross-product between networks/tables and their matrix reordering algorithms as presented in [4].

For the retrieval of similar patterns we developed a prototypical query-by-sketch interface, depicted in Fig. 10, which can also be accessed online[4]. The query-by-sketch interface allows the user to describe intuitively the expected visual patterns by drawing a sketch of a sought matrix pattern (Fig. 10 (1)), and receiving a ranked list of similar images (Fig. 10 (2)) according to one or multiple selected feature descriptors (Fig. 10 (3)). For each individually selected FD the feature vectors of the sketch image and the database images are compared (Euclidean distance) to retrieve a ranking score. The MAGNOSTICS FD set is also available: Here, a derived six-dimensional feature vector is constructed by calculating the Distance-to-Base-Pattern for every individual MAGNOSTICS FD. The query image feature vector is subsequently compared with the Euclidean distance to the image feature vectors in the database.

Fig. 10 shows an example using the BLOCKS descriptor for the matrix reordering data set collection with 4,313 matrices [4]. Matrices show several blocks for the drawn sketch.

---

[4]http://magnostics.dbvis.de/#/sketch
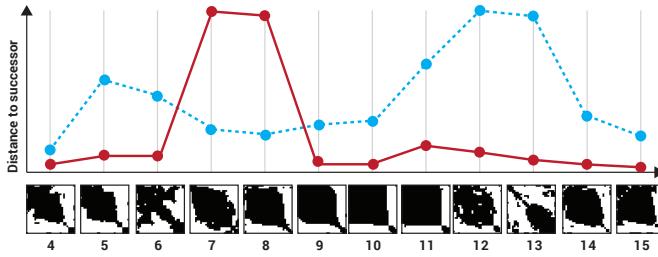
Fig. 11: Detail from a dynamic network representing brain connectivity (300 time records).

7.2 Dynamic Network Analysis

Another challenge in the realm of large data collections are networks changing over time; every time step represents the network at a different time instance. Fig. 11 shows 11 matrices from a dynamic brain connectivity network, usually comprising around 300 time points (individual matrices). Brain connectivity refers to the co-variance (connection) between activity in brain regions (regions of interest, ROI). Neuroscientists are specifically interested in clusters (Blocks) and their evolution, as well as in identifying noisy time periods. Here, a time point represents 2-seconds, and dark cells indicate high connectivity (binarized to obtain un-weighted network). Rows and columns in the matrices are ordered to optimize the visual patterns *independently* for each matrix.

To find time points with major changes with respect to noise and blocks, we first calculate a feature vector v_i for every matrix (time point) t_i with HARALICK (noise) and another one with BLOCKS. Then, we calculate the difference for each feature vector v_i and the following time step v_{i+1} (separately for both HARALICK and BLOCKS). High distances indicate high changes between two consecutive time points, low distances indicate little change (with respect to noise).

Fig. 11 shows both distances plotted (red=BLOCKS, blue=HARALICK). For noise we can observe major differences between time points 11-14, while the type of blocks changes most between time points 6-8. Both changes are observable in the respective matrices. Further, we observe, between 9 and 15, a constant change (constant distances) for the BLOCKS FD (red), indicating a salient trend; a possible explanation could be the change from two separate blocks (clusters) to one larger block.

8 DISCUSSION AND EXTENSIONS

Quantifying patterns in visualizations is an active research field with varying approaches. While pattern measures can also be derived from the data space, in our case by graph analysis, MAGNOSTICS follows the idea to take advantage of screen-space measures. This has the advantage that a direct correspondence to the human perceptual system can be constructed.

With MAGNOSTICS we propose an initial set of feature descriptors which showed appropriate based on certain desirable criteria (see Sect. 4.4) and showed useful to support ranking, searching and clustering of matrices (see Sect. 7).

8.1 Defining and Retrieving Interesting Matrix Views

While we have focused on six specific patterns and their variations, we believe our criteria and methodology from Sect. 4 will allow for the selection of FDs for different patterns including directed and weighted networks, as well as application specific patterns. Also, higher-level visual *characteristics* of matrices, such as the degree of "clumpy-ness" could be investigated with our MAGNOSTICS evaluation approach.

Further criteria for an effective MAGNOSTICS FD could also be considered. For example, the human users' assessment of feature descriptor quality in light of matrix exploration tasks could influence the FD selection process in a user-centric fashion. For example, lines may still be visible to humans, though the algorithms may not detect them and vice versa. It may also be interesting to link MAGNOSTICS analysis with interactive matrix exploration techniques. For example,

the NodeTrix approach [17] includes node-link views aside the matrices, hence we would also need to include FDs for node-link views to jointly rank interesting views.

A scenario we have not followed further, but which is described in Scagnostics [46], is to present matrix views that are most *different* from the others, i.e., outlier views. Since MAGNOSTICS is based on the distances between feature vectors, it is straightforward to cluster matrix views based on the feature-vector distances and to report on more common views and outliers. Finally, MAGNOSTICS can be embedded in interactive approaches, in that FDs get weighted either explicitly by the user, or implicitly by a function analyzing an user's preferences while navigating the matrix space.

Given that FDs and their corresponding distance functions are the basis for many data analysis tasks, more application scenarios as the presented can be supported. In this work we considered ranking and searching tasks. Other potentially useful applications include the classification of matrix views according to expected classes, according to user interest or the clustering of matrix views.

8.2 Novel Image Measures for Matrix Diagnostics

Another open question remains the development of novel feature descriptors for matrices. As described in the previous section, there may be more visual patterns important to users of matrix analysis. While our FDs are *global*, a natural extension would be to consider *local* FDs, focussing on specific subnetworks. Furthermore, we considered adjacency (binary) matrices, but MAGNOSTICS could be extended with FDs also for continuous matrices (e.g., representing weighted networks).

Finally, we mention that the MAGNOSTICS framework can easily be applied to evaluate existing matrix reordering algorithms in terms of the visual patterns they produce. Our approach could even inform the design of new reordering algorithms which aim to optimize matrix views for certain user-dependent patterns of interest.

9 CONCLUSION

We introduced MAGNOSTICS, an experimentally validated set of view descriptors aiming to guide exploration of large network collections, represented as matrix views. Starting from a set of 30 feature descriptors, including three novel specifically designed FDs, we identified a set of six useful FDs for the analysis of visual patterns in matrices. The selection of descriptors was guided by a structured and explorative methodology, using a novel large matrix benchmark data set and based on four quantifiable criteria: (1) pattern response, (2) pattern similarity, (3) pattern sensitivity and (4) pattern discrimination. We demonstrated how MAGNOSTICS can be applied in an interactive ranking and searching tasks, and to analyze time series of networks. MAGNOSTICS complements the set of previous feature-based analysis frameworks in the context of Visual Analytics tools.

REFERENCES

[1] G. Albuquerque, M. Eisemann, D. J. Lehmann, H. Theisel, and M. Magnor. Improving the visual analysis of high-dimensional datasets using quality measures. In *Visual Analytics Science and Technology (VAST), 2010 IEEE Symposium on*, pp. 19–26. IEEE, 2010.

[2] R. A. Baeza-Yates and B. A. Ribeiro-Neto. *Modern Information Retrieval - the concepts and technology behind search, Second edition*. Pearson Education Ltd., Harlow, England, 2011.

[3] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool. Speeded-up robust features (surf). *Computer Vision and Image Understanding*, 110(3):346–359, June 2008. doi: 10.1016/j.cviu.2007.09.014

[4] M. Behrisch, B. Bach, N. H. Riche, T. Schreck, and J.-D. Fekete. Matrix Reordering Methods for Table and Network Visualization. *Computer Graphics Forum*, 2016. doi: 10.1111/cgf.12935

[5] M. Behrisch, F. Korkmaz, L. Shao, and T. Schreck. Feedback-driven interactive exploration of large multidimensional data supported by visual classifier. In *Proc. IEEE Conference on Visual Analytics Science and Technology*, pp. 43–52, 2014. doi: 10.1109/VAST.2014.7042480

[6] E. Bertini, A. Tatu, and D. A. Keim. Quality Metrics in High-Dimensional Data Visualization: An Overview and Systematization. *IEEE Symp. on Information Visualization (InfoVis)*, 17(12):2203–2212, Dec. 2011.

[7] A. Bosch, A. Zisserman, and X. Munoz. Representing shape with a spatial pyramid kernel. In *Proc. of the 6th ACM International Conference on Image and Video Retrieval*, pp. 401–408. ACM, 2007.

[8] S. A. Chatzichristofis and Y. S. Boutalis. Cedd: color and edge directivity descriptor: a compact descriptor for image indexing and retrieval. In *Computer vision systems*, pp. 312–322. Springer, 2008.

[9] S. A. Chatzichristofis and Y. S. Boutalis. Fcth: Fuzzy color and texture histogram - a low level feature for accurate image retrieval. In *Image Analysis for Multimedia Interactive Services, 2008. WIAMIS '08. 9th Int. Workshop on*, pp. 191–196, May 2008. doi: 10.1109/WIAMIS.2008.24

[10] A. Dasgupta and R. Kosara. Pargnostics: Screen-space metrics for parallel coordinates. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):1017–1026, 2010.

[11] G. Ellis and A. Dix. A taxonomy of clutter reduction for information visualisation. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1216–1223, 2007. doi: 10.1109/TVCG.2007.70535

[12] R. Gonzalez and R. Woods. *Digital Image Processing*. Pearson/Prentice Hall, 2008.

[13] F. Graf. Jfeaturelib v1.6.3, Sept. 2015. doi: 10.5281/zenodo.31793

[14] J. Han and K.-K. Ma. Fuzzy color histogram and its use in color image retrieval. *Image Processing, IEEE Transactions on*, 11(8):944–952, 2002.

[15] R. Haralick, K. Shanmugam, and I. Dinstein. Textural features for image classification. *Systems, Man and Cybernetics, IEEE Transactions on*, SMC-3(6):610–621, Nov 1973. doi: 10.1109/TSMC.1973.4309314

[16] M. Heikkl and M. Pietikinen. A texture-based method for modeling the background and detecting moving objects. *IEEE transactions on pattern analysis and machine intelligence*, 28(4):657–62, 2006.

[17] N. Henry, J.-D. Fekete, and M. J. McGuffin. NodeTrix: a hybrid visualization of social networks. *IEEE transactions on visualization and computer graphics*, 13(6):1302–9, 2007. doi: 10.1109/TVCG.2007.70582

[18] P. Hough. Method and means for recognizing complex patterns, Dec 1962.

[19] J. Huang, S. R. Kumar, M. Mitra, W.-J. Zhu, and R. Zabih. Image indexing using color correlograms. In *Computer Vision and Pattern Recognition, Proc. of IEEE Conference on*, pp. 762–768, Jun 1997. doi: 10.1109/CVPR.1997.609412

[20] S. Ingram, T. Munzner, V. Irvine, M. Tory, S. Bergner, and T. Möller. Dimstiller: Workflows for dimensional analysis and reduction. In *Proc. of the IEEE Conference on Visual Analytics Science and Technology, IEEE VAST,*, pp. 3–10, 2010. doi: 10.1109/VAST.2010.5652392

[21] E. Kasutani and A. Yamada. The mpeg-7 color layout descriptor: a compact image feature description for high-speed image/video segment retrieval. In *Image Processing, 2001. Proceedings. 2001 Int. Conference on*, vol. 1, pp. 674–677 vol.1, 2001. doi: 10.1109/ICIP.2001.959135

[22] D. J. Lehmann, F. Kemmler, T. Zhyhalava, M. Kirschke, and H. Theisel. Visualnostics: Visual guidance pictograms for analyzing projections of high-dimensional data. In *Computer Graphics Forum*, vol. 34, pp. 291–300. Wiley Online Library, 2015.

[23] I. Liiv. Seriation and matrix reordering methods: An historical overview. *Statistical analysis and data mining*, 3(2):70–91, 2010.

[24] M. Lux and S. A. Chatzichristofis. Lire: Lucene image retrieval: An extensible java cbir library. In *Proceedings of the 16th ACM International Conference on Multimedia*, MM '08, pp. 1085–1088. ACM, New York, NY, USA, 2008. doi: 10.1145/1459359.1459577

[25] M. Lux, A. Pitman, and O. Marques. Callisto: Tag recommendations by image content. *WISMA 2010*, p. 87, 2010.

[26] B. S. Morse. Lecture 9: Shape description (regions).

[27] C. Mueller, B. Martin, and a. Lumsdaine. Interpreting large visual similarity matrices. *2007 6th International Asia-Pacific Symposium on Visualization*, pp. 149–152, Feb. 2007. doi: 10.1109/APVIS.2007.329290

[28] D. K. Park, Y. S. Jeon, and C. S. Won. Efficient use of local edge histogram descriptor. In *Proceedings of the 2000 ACM Workshops on Multimedia*, MULTIMEDIA '00, pp. 51–54. ACM, New York, NY, USA, 2000. doi: 10.1145/357744.357758

[29] W. Peng, M. O. Ward, and E. A. Rundensteiner. Clutter reduction in multi-dimensional data visualization using dimension reordering. In *Information Visualization. INFOVIS 2004. IEEE Symp.*, pp. 89–96. IEEE, 2004.

[30] E. Rosten, R. Porter, and T. Drummond. Faster and better: A machine learning approach to corner detection. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 32:105–119, 2010. doi: 10.1109/TPAMI.2008.275

[31] Y. Rui, T. S. Huang, and S.-F. Chang. Image retrieval: Current techniques, promising directions, and open issues. *Journal of visual communication and image representation*, 10(1):39–62, 1999.

[32] M. Scherer, J. Bernard, and T. Schreck. Retrieval and exploratory search in multivariate research data repositories using regressional features. In *Proceedings of the 11th annual international ACM/IEEE joint conference on Digital libraries*, pp. 363–372. ACM, 2011.

[33] J. Schneidewind, M. Sips, and D. A. Keim. Pixnostics: Towards measuring the value of visualization. In *Visual Analytics Science And Technology, 2006 IEEE Symposium On*, pp. 199–206. IEEE, 2006.

[34] J. Seo and B. Shneiderman. A rank-by-feature framework for unsupervised multidimensional data exploration using low dimensional projections. In *Information Visualization, 2004. INFOVIS 2004. IEEE Symposium on*, pp. 65–72, 2004. doi: 10.1109/INFVIS.2004.3

[35] L. Shao, M. Behrisch, T. Schreck, T. von Landesberger, M. Scherer, S. Bremm, and D. A. Keim. Guided Sketching for Visual Search and Exploration in Large Scatter Plot Spaces. In M. Pohl and J. Roberts, eds., *Proc. EuroVA International Workshop on Visual Analytics*. The Eurographics Association, 2014. doi: 10.2312/eurova.20141140

[36] L. Shao, T. Schleicher, M. Behrisch, T. Schreck, I. Sipiran, and D. Keim. Guiding the exploration of scatter plot data using motif-based interest measures. In *IEEE Int. Symposium on Big Data Visual Analytics*, 2015.

[37] A. W. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain. Content-based image retrieval at the end of the early years. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 22(12):1349–1380, 2000.

[38] T. Smith, G. Lange, and W. Marks. Fractal methods and results in cellular morphologydimensions, lacunarity and multifractals. *Journal of neuroscience methods*, 69(2):123–136, 1996.

[39] H. Tamura, S. Mori, and T. Yamawaki. Textural features corresponding to visual perception. *IEEE Transactions on Systems, Man, and Cybernetics*, 8(6):460–473, June 1978. doi: 10.1109/TSMC.1978.4309999

[40] A. Tatu, G. Albuquerque, M. Eisemann, P. Bak, H. Theisel, M. A. Magnor, and D. A. Keim. Automated analytical methods to support visual exploration of high-dimensional data. *IEEE Transactions on Visualization and Computer Graphics*, 17(5):584–597, 2011.

[41] A. Tatu, F. Maas, I. Farber, E. Bertini, T. Schreck, T. Seidl, and D. Keim. Subspace search and visualization to make sense of alternative clusterings in high-dimensional data. In *Visual Analytics Science and Technology (VAST), 2012 IEEE Conference on*, pp. 63–72. IEEE, 2012.

[42] K. van de Sande, T. Gevers, and C. Snoek. Evaluating color descriptors for object and scene recognition. *Pattern Analysis and Machine Intelligence, IEEE Trans. on*, 32(9):1582 –1596, sept. 2010. doi: 10.1109/TPAMI.2009.154

[43] L. von Rüden. Visual analytics of parallel-performance data: Automatic identification of relevant and similar data subsets. Master's thesis, RWTH Aachen University, April 2015.

[44] L. von Rüden, M.-A. Hermanns, M. Behrisch, D. A. Keim, B. Mohr, and F. Wolf. Separating the wheat from the chaff: Identifying relevant and similar performance data with visual analytics. In *Proc. of the 2nd Workshop on Visual Performance Analysis (VPA), Supercomputing Conference 2015*, pp. 4:1–4:8. ACM, 2015. doi: 10.1145/2835238.2835242

[45] L. Wilkinson. *The Grammar of Graphics (Statistics and Computing)*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2005.

[46] L. Wilkinson, A. Anand, and R. Grossman. Graph-theoretic scagnostics. In *Proceedings of the Proceedings of the 2005 IEEE Symposium on Information Visualization*, INFOVIS '05, pp. 21–. IEEE Computer Society, Washington, DC, USA, 2005. doi: 10.1109/INFOVIS.2005.14

[47] C. S. Won. *Adv. in Multimedia Information Processing - PCM 2004: 5th Pacific Rim Conference on Multimedia, Tokyo, Japan, 2004. Proceedings, Part III*, chap. Feature Extraction and Evaluation Using Edge Histogram Descriptor in MPEG-7, pp. 583–590. Springer, 2005. doi: 10.1007/978-3-540-30543-9_73

[48] M. Yang, K. Kpalma, and J. Ronsin. *A Survey of Shape Feature Extraction Techniques*. Peng-Yeng Yin, 2008.

[49] I. T. Young, J. E. Walker, and J. E. Bowie. An analysis technique for biological shape. i. *Information and control*, 25(4):357–370, 1974.