# Unsupervised Segmentation of Anomalies in Sequential Data, Images and Volumetric Data Using Multiscale Fourier Phase-Only Analysis

Fabian Bürger and Josef Pauli

Intelligent Systems Group, Faculty of Engineering,
University of Duisburg-Essen, Germany
{fabian.buerger,josef.pauli}@uni-due.de
http://www.is.uni-due.de/

**Abstract.** This paper presents an unsupervised method to detect and segment anomalies and novel patterns in sequential data, images and volumetric data. The proposed Multiscale Phase-Only Transformation (MPHOT) addresses the case when no prior knowledge about the data or even its dimensionality is provided. It is based on the fusion of the Phase-Only Transform (PHOT) in scale space using only one adaptive sensitivity parameter. The PHOT uses the Discrete Fourier Transform (DFT) to remove all regularities while it detects small defects and pattern boundaries. The proposed multiscale extension allows the precise segmentation of large anomalies as well. We present experiments on synthetic and measured data in fields of time series analysis, image processing and volumetric data segmentation to show the universality of our approach.

**Keywords:** Unsupervised Anomaly Detection, Novelty Detection, Texture Segmentation, Volumetric Segmentation, Data Mining.

## 1 Introduction

The detection of "novel" or abnormal patterns in sequence or spatial data has many applications in signal processing and data mining. For one-dimensional time series there are applications in sensor monitoring tasks or audio data processing. In the two-dimensional case there is the huge field of image processing with applications in surface quality inspection as well as the detection of salient objects for e.g. robot vision systems. In the three-dimensional domain there are possible applications in analysis of medical MRI/CT images or meteorological data.

In many applications, sensors collect large amounts of data which mostly contains regions with unremarkable background patterns. However, it is possible that certain regions show significantly different patterns which potentially indicate the need for further analysis. The human vision system can rapidly detect abnormal patterns in images or 2D-visualizations of data without the need to be specially trained. In real-world applications, manual inspection of huge amounts

of data is usually undesired because it is repetitive, time consuming and also subjective and therefore potentially error-prone. Hence, there is a high demand for the automation of such tasks.

Aiger and Talbot proposed the Phase-Only Transform (PHOT) [1] to automatically detect defects in textured images using the phase information from the Discrete Fourier Transform (DFT). Their results are already promising and resemble the human perception of abnormal patterns in images. Motivated by the simplicity and effectiveness of the PHOT, we analyze further properties and also limitations. We present improvements of the PHOT to achieve not only the detection of anomalies but also a precise segmentation. Furthermore, we extend the algorithm to $n$-dimensional data which opens new application fields.

## 2    Related Work

A lot of research has been made in the fields of novelty or anomaly detection. General overviews of state-of-the-art research can be found in [2] and [3]. First, the type of input data is important for choosing a certain approach. Data for novelty detection can be separated into unrelated data, such as feature vectors for classification, and spatial, temporal or graph-based related data. To detect abnormal patterns in time series, grayscale images and volumetric data, we assume that these anomalies can be characterized by their salient patterns which are defined by a significantly different spatial or temporal distribution.

A popular approach for anomaly detection in data streams is the extraction of problem-specific features, such as statistical, structural, filter based or model based features [4]. These features are analyzed with statistical or machine learning methods to detect anomalies. Usually a set of well-chosen features outperforms generalized methods for a particular application. But these features need to be adapted when the patterns in the data may change dynamically. Another aspect is the availability of data labels as methods are generally separated into supervised, semi-supervised and unsupervised approaches. In applications where samples of both normal and abnormal classes can easily be generated, a supervised approach is normally preferred. Since it is often impossible to train all possible normal and abnormal patterns that might occur, unsupervised methods require less maintenance. In case of unsupervised anomaly detection in sequential data, various approaches have been proposed, such as time series bitmaps [5], symbolic representations [6] and time frequency analysis in combination with support vector machines [7].

It is believed that a fast preattentive stage is responsible for the extraordinary pattern recognition capabilities of the human (or mammalian) vision system [8]. This stage processes basic features like color, contrast, orientation and motion to discard irrelevant visual information. Computational models of visual attention, first introduced by Itti and Koch [9], use the fusion of feature channels to derive so-called saliency maps indicating important areas inside of images. In [10], [11] and [12] frequency based models for visual attention are presented which are closely related to our approach.

# 3   The Phase-Only Transform

## 3.1   Definition and Applications

We focus on the case when the data pattern can change significantly and unforeseen anomalies have to be detected robustly when there is no prior training data available. This case is addressed by Aiger and Talbot in [1]. They propose a simple method to detect "patterns that do not appear much" in grayscale images without any training using a spectral whitening process which is basically a normalization of the magnitude in the discrete Fourier domain. The remaining phase contains important parts of the signal [13]. Since all operations can be applied to $n$-dimensional arrays (or tensor data), we present the generalization of the originally proposed PHOT. Furthermore, we added two modifications denoted with an asterisk *). Given an $n$-dimensional input array

$$A(l_1, \ldots, l_n) = A(v) \to \mathbb{R}, \quad 0 \le l_1 < N_1, \ldots, 0 \le l_n < N_n, \quad l_k \in \mathbb{Z} \quad (1)$$

where an index item of the $N_1 \times \ldots \times N_n$ array is denoted as vector $v \in V \in \mathbb{Z}^n$. $V$ is the index set and contains $|V| = N_1 \cdot N_2 \cdot \ldots \cdot N_n$ items.

1. Transform $A(v)$ into the DFT domain: $\widetilde{A}(\widetilde{v}) \leftarrow \mathrm{DFT}_n(A(v))$ where $\widetilde{v} \in \widetilde{V}$ denote the spatial frequencies.
2. For $\forall \widetilde{v} \in \widetilde{V}$ normalize each number by its magnitude:
   $$\widetilde{A}_{\mathrm{norm}}(\widetilde{v}) \leftarrow \widetilde{A}(\widetilde{v})/|\widetilde{A}(\widetilde{v})|, \quad |\widetilde{A}(\widetilde{v})| = \sqrt{\Re(\widetilde{A}(\widetilde{v}))^2 + \Im(\widetilde{A}(\widetilde{v}))^2} > 0$$
3. Perform inverse DFT to obtain: $O(v) \leftarrow \mathrm{IDFT}_n(\widetilde{A}_{\mathrm{norm}}(\widetilde{v}))$.
4. *) Apply the absolute value function to $O(v) \leftarrow |O(v)|$. This step prevents the cancellation of positive and negative values during the following filter operations.
5. Apply an $n$-dimensional Gaussian filter $G_\sigma$ with $\sigma = 3.0$ to obtain:
   $O(v) \leftarrow O(v) * G_\sigma$.
6. Compute the *scalar* mean $\mu$ and standard deviation $\sigma$ of all values in $O(v)$ and apply the Mahalanobis distance: $O_{PHOT}(v) \leftarrow (O(v) - \mu)/\sigma$ for $\forall (v) \in V$ and $\sigma > 0$.
7. *) Cut values below zero: $O_{PHOT}(v) \leftarrow \max(O_{PHOT}(v), 0) \ \ \forall v \in V$ because after the modification in step 4, only positive values of the Mahalanobis distance indicate significant anomalies.

The functions $\mathrm{DFT}_n(\cdot)$ and $\mathrm{IDFT}_n(\cdot)$ denote the $n$-dimensional DFT which can be computed of sequences of one-dimensional DFTs [14]. The Gaussian filter is needed to dampen the high noise level in the $O_{PHOT}(v)$ array due to the emphasis of high frequencies. The resulting evidence array contains the anomaly energy or significance for each item $v$. This array can be thresholded to obtain a binary segmentation in which values $O_{PHOT}(v) \ge T_0$ denote an abnormal item. The authors in [1] suggest a global threshold $T_0 = 4.0$ to detect defective areas in case of 2D image processing.

## 3.2   Properties and Limitations

The PHOT removes any regular pattern at any scale while high filter responses appear near or at boundaries of irregular patterns. This behavior can be explained as follows. Under the assumption of an additive signal model

$$A(v) = A_{periodic}(v) + A_{anomaly}(v) \tag{2}$$

in which the periodic part consists of the superposition of harmonic functions while the abnormal pattern differs significantly from the periodic one. In [1] the authors explain the effect with a non-negative phase integral excursion of the abnormal signal. The way how the DFT realizes local information leads to an alternative explanation. The phase shifts the periodic functions so that their superposition is non-zero at least at pattern boundaries. The magnitude of all periodic waves is normalized which causes the values at these borders to gain significantly more weight. In proximity to these boundaries Gibbs overshoot and ringring phenomena occur, but their intensity is always smaller than the main peak causing them [15]. In fact, in the one-dimensional case the overshoot can be quantified to around 8.9%. The absolute value operation in step 4 (our modification) of the generalized PHOT causes the mean $\mu$ of $O(v)$ to raise so that the influence of the artifacts on the positive Mahalanobis distance is reduced.

An example of this transform can be found in figure 1 in which the original PHOT and our modified PHOT have been applied on an image showing a Brodatz texture collage [16]. Our results provide precise localizations of small defects and much clearer contours of larger defective areas. The proposed modifications dampen the effect of the Gibbs artifacts, though the detected areas appear slightly too large. Furthermore, only the boundaries of large abnormal areas are detected because a regular pattern in the inside is removed by the PHOT as well as the regular pattern outside.
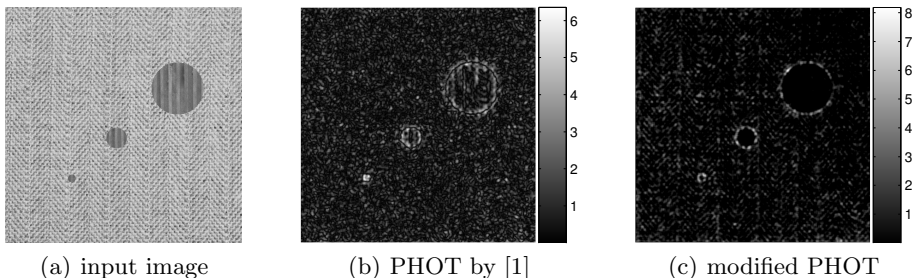


|         (a) input image         |         (b) PHOT by [1]         |         (c) modified PHOT         |

**Fig. 1.** Example application of the PHOT on a collage of Brodatz textures showing 3 abnormal, circular areas

# 4   Multiscale Extension of the PHOT

## 4.1   Motivation for the Multiscale Extension

Many visual attention models, like [9], use multiscale analysis to detect size-invariant salient areas. We transfer this idea to the PHOT in order to achieve a complete or "filled" segmentation rather than a boundary estimation of the anomalies. Using a multiscale approach avoids the need for a shape model (e.g. convex shapes) which is desirable since there is no prior knowledge about the morphology of abnormal areas.

## 4.2   Multiscale Fusion

The idea is to compute a data resolution pyramid, perform the modified PHOT on each scale and merge the results to achieve a scale-invariant anomaly segmentation. Given an $n$-dimensional array $A(v)$, the Multiscale Phase-Only Transform (MPHOT) is defined as follows. First, a set of $m$ scales is determined while we use dyadic Gaussian resolution pyramids $S_m = \{s_1, s_2, \ldots, s_m\}$ with $s_i = 2^{-i+1}$, $1 \leq i \leq m$. Experiments in section 5 show that the best results can be obtained using $S_5 = \{1.0, 0.5, 0.25, 0.125, 0.0625\}$.

1. For $\forall s_i \in S$ use linear interpolation to scale down $A(v)$ to $A_i(v)$. Note that the size of the array is scaled to $\lceil s_i \cdot N_1 \rceil \times \ldots \times \lceil s_i \cdot N_n \rceil$.
2. Perform the $n$-dimensional modified PHOT defined in section 3.1 to all $A_i(v)$ and obtain the scaled PHOT arrays $O_{PHOT,i}(v)$.
3. In order to overcome the Gibbs artifacts at boundaries we suggest the use of a generalized morphological erosion [17] which causes a thinning of areas with high values in the evidence array. We choose a centered structuring element $K$ in shape of an $n$-cube with $k^n$ one-entries while we use an element size of $k = 5$: $O_{PHOT,i}(v) \leftarrow O_{PHOT,i}(v) \ominus K$.
   Note that a "gray value" erosion is meant here which selects the minimum value in the neighborhood defined by $K$.
4. For $\forall s_i \in S$ use linear interpolation to scale up $O_{PHOT,i}(v)$ to $\widehat{O}_{PHOT,i}(v)$ with the size of the input array $A(v)$.
5. To realize the fusion of the results we suggest a nonlinear element-wise weighted summation:

$$O_{MPHOT}(v) = \sum_{i=1}^{m} \left( \frac{\widehat{O}_{PHOT,i}(v)}{T(s_i)} \right)^{\alpha} \quad \forall v \in V, \quad T(s_i) = T_0 \cdot s_i^{\gamma} \quad (3)$$

where $T(s_i)$ is a scale-adaptive evidence threshold function containing the base threshold $T_0$ parameter and the factor $s_i^{\gamma}$. This factor compensates the decay of the evidence values at smaller scales due to alias effects. Experiments show that $\gamma = \frac{1}{3}$ equalizes well the maximum evidence values at

all scales. The division by the scale-adaptive threshold $T(s_i)$ introduces a percentaged metric in which the value 1 means "100% significance" using threshold $T_0$. The fusion exponent $\alpha > 1$ emphasizes values greater than 1 and dampens values below. Experiments show that $\alpha = 2$ works well in the sense of anomaly shape preserving and reduction of false positives.

6. Apply a threshold to the significance values $O_{MPHOT}(v) \geq 1$ (greater or equal 100% of $T_0$) to obtain the final segmentation.

The main parameter of the MPHOT is the significance threshold $T_0$ for the Mahalanobis distance which is similar to the one in [1], but takes the accumulated significance at all scales into account. Example applications on one- and two-dimensional data can be found in figure 2 and 3. The main effect of the multiscale fusion can be seen in both dimensions: The high resolution scale filters detect boundaries which are subsequently "filled" with the responses of the smaller scales. A threshold of $T_0 = 3.0$ is used here which is a good "global" parameter, as our results in section 5 show. The MPHOT extension basically uses PHOT (which relies mainly on DFT and convolution), linear interpolation and mathematical erosion which are available in many programming libraries. The multiscale extension only considers down-scaled PHOT calculations which can independently be processed in parallel on CPU or GPU devices.
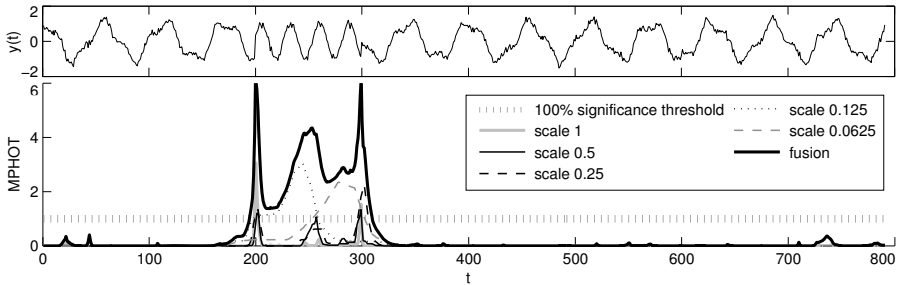


**Fig. 2.** Application of the MPHOT on synthetic time series with an anomaly in the interval $200 \leq t \leq 300$ ($T_0 = 3.0$ and scales $S_5$)
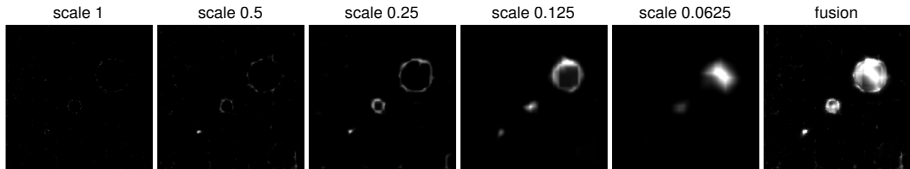


**Fig. 3.** Application of the MPHOT on the texture collage of figure 1 ($T_0 = 3.0$ and scales $S_5$)

# 5   Results

The focus of our experiments is the diversity of applications of the MPHOT rather than extensive investigations on single problems. To show the flexibility of our approach we choose a set of eight diverse applications in which both background patterns and anomalies are different, but can easily be separated by humans. In case of time series we use a synthetic, periodic signal with noise (figure 2), a human electrocardiogram (ECG) as well as a human electroencephalogram (EEG) of an epileptic seizure (both figure 4). In the two-dimensional case we use a texture collage (figure 1), a steel surface with defects and a natural photograph with animals that should be segmented (both figure 5). In the three-dimensional domain we use periodic and random patterns containing anomalies (figure 6). We present combined evaluations for all dimensions to show that the MPHOT does not need prior knowledge about the data or its dimension.

In order to measure the performance of the MPHOT with respect to the threshold parameter $T_0$ and the scale set $S_m$ we evaluate the results elementwise with ground truth data. We derive the true positive rate, false positive rate as well as sensitivity, specificity and accuracy [18]. Our results in table 1 show that the specificity and accuracy values of the segmentation are almost dimension-invariant. The sensitivity measure shows a higher standard deviation which can be explained with the different definitions of the "degree" of anomalies depending on the application context and also due to partly subjective ground truth data. Figure 7 (a) shows the overall average ROC curves (Receiver Operating Characteristic) comparing the proposed MPHOT in combination with different sets of scales with the original PHOT [1]. We also compare two alternative methods, namely the Mahalanobis distance thresholding directly (MD) and in combination with a prior Fourier magnitude thresholding (FMT+MD), in which frequencies with high magnitudes (two times higher than the mean) are removed. The proposed MPHOT using scales $S_5$ performs best in our experiments. Finally in figure 7 (b) the effect of the threshold parameter $T_0$ on the $F_\beta$-score is depicted for our scenarios. The $F_\beta$-score is a weighted combination of precision and sensitivity in which $F_1$ means an equal weight, $F_2$ a double emphasis on sensitivity and $F_{0.5}$ double weight on precision [19]. In our datasets a range for $T_0$ of $[2.0, 3.75]$ performs well depending on the desired $F_\beta$-score. The suggested global threshold $T_0 = 4.0$ of the PHOT in [1] also leads to good results using MPHOT when the $F_{0.5}$-score should be optimized.

**Table 1.** Sensitivity, specificity and accuracy for $T_0 = 3.0$ and scales $S_5$

|                        | sensitivity       | specificity       | accuracy          |
|------------------------|-------------------|-------------------|-------------------|
| 1D scenarios           | $0.774 \pm 0.203$ | $0.985 \pm 0.013$ | $0.938 \pm 0.060$ |
| 2D scenarios           | $0.814 \pm 0.100$ | $0.991 \pm 0.002$ | $0.987 \pm 0.005$ |
| 3D scenarios           | $0.744 \pm 0.184$ | $0.999 \pm 0.000$ | $0.996 \pm 0.002$ |
| overall (1D, 2D, 3D)   | $0.781 \pm 0.142$ | $0.991 \pm 0.009$ | $0.971 \pm 0.042$ |

(a) ECG time series with anomaly [6]    (b) EEG with an epileptic seizure [20]
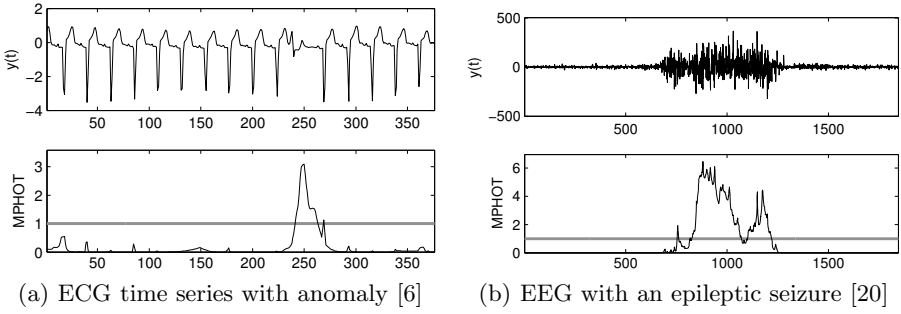
**Fig. 4.** Application of the MPHOT on measured time series with $T_0 = 3.0$ and scales $S_5$



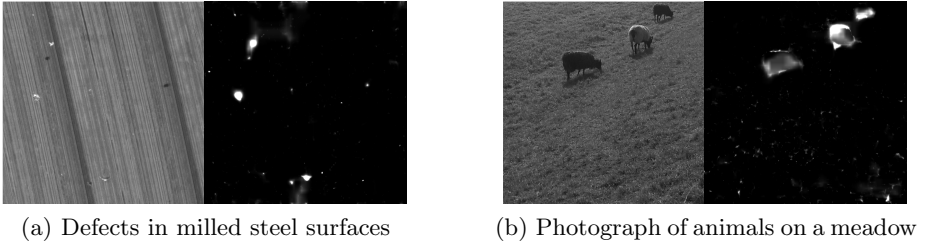(a) Defects in milled steel surfaces    (b) Photograph of animals on a meadow

**Fig. 5.** Application of the MPHOT on images as defect detection (a) and object/background segmentation in natural images (b) with $T_0 = 3.0$ and scales $S_5$
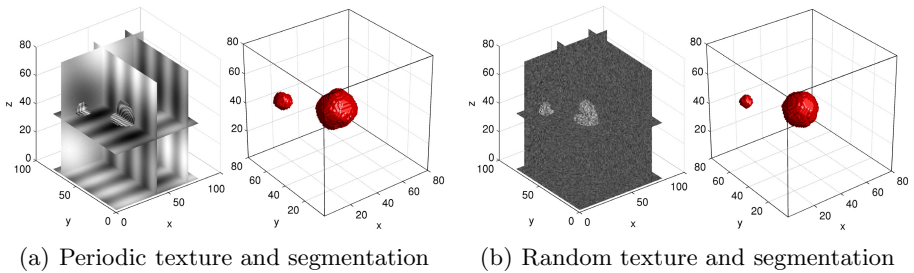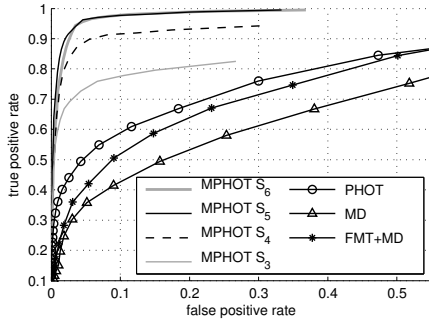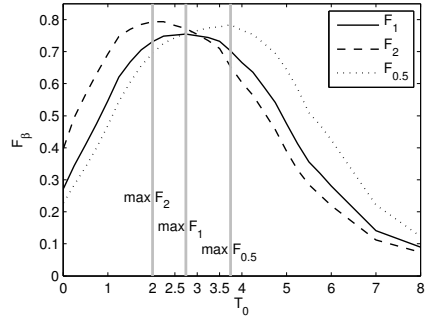


(a) Periodic texture and segmentation    (b) Random texture and segmentation

**Fig. 6.** Application of the MPHOT on volumetric data containing 2 ball-shaped anomalies with the corresponding segmentation using $T_0 = 3.0$ and scales $S_5$. Note that the volumetric data is only shown partly by 5 slices.

(a) Overall ROC curves

(b) Overall $F_\beta$-scores for $T_0$ using MPHOT with scales $S_5$

**Fig. 7.** Influence of the threshold $T_0$ and scales on ROC space and $F_\beta$-score

## 6    Conclusions

We presented a multiscale and multi-dimensional extension to the PHOT which allows an unsupervised detection and segmentation of anomalies in 1D, 2D and 3D data. The proposed MPHOT provides promising results in a huge variety of tasks while having only one main threshold parameter. Furthermore, we analyzed the optimal range for this parameter to achieve good segmentation results without prior knowledge of the pattern inside of the data or its dimension. As only few approaches for unsupervised, dimension-invariant anomaly segmentation exist, future work will cover the comparison to state-of-the-art methods at each dimension. Additionally, new applications of the MPHOT on even higher dimensional data need to be explored.

## References

1. Aiger, D., Talbot, H.: The phase only transform for unsupervised surface defect detection. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 295–302 (2010)
2. Chandola, V., Banerjee, A., Kumar, V.: Anomaly detection: A survey. ACM Computing Surveys 41(3), 1–58 (2009)
3. Markou, M., Singh, S.: Novelty detection: a review – part 1+2. Signal Processing 83(12), 2481–2521 (2003)
4. Xie, X.: A Review of Recent Advances in Surface Defect Detection using Texture analysis Techniques. Electronic Letters on Computer Vision and Image Analysis 7, 1–22 (2008)
5. Wei, L., Kumar, N., Lolla, V., Keogh, E.J., Lonardi, S., Ratanamahatana, C.: Assumption-free anomaly detection in time series. In: Proceedings of the 17th International Conference on Scientific and Statistical Database Management, SSDBM 2005, pp. 237–240. Lawrence Berkeley Laboratory (2005)

6. Keogh, E., Lin, J., Fu, A.: HOT SAX: efficiently finding the most unusual time series subsequence. In: Fifth IEEE International Conference on Data Mining, p. 8 (2005)
7. Davy, M., Godsill, S.: Detection of abrupt spectral changes using support vector machines an application to audio signal segmentation. In: IEEE International Conference on Acoustics, Speech, and Signal Processing, vol. 2, pp. 1313–1316 (2002)
8. Wolfe, J.: Guided Search 2.0 A revised model of visual search. Psychonomic Bulletin & Review 1, 202–238 (1994)
9. Itti, L., Koch, C., Niebur, E.: A model of saliency-based visual attention for rapid scene analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence 20(11), 1254–1259 (1998)
10. Hou, X., Zhang, L.: Saliency Detection: A Spectral Residual Approach. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 1–8 (2007)
11. Wang, Z., Li, B.: A two-stage approach to saliency detection in images. In: IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 965–968 (2008)
12. Guo, C., Zhang, L.: A novel multiresolution spatiotemporal saliency detection model and its applications in image and video compression. IEEE Transactions on Image Processing 19(1), 185–198 (2010)
13. Oppenheim, A., Lim, J.: The importance of phase in signals. Proceedings of the IEEE 69(5), 529–541 (1981)
14. Tolimieri, R., An, M., Lu, C.: Mathematics of multidimensional Fourier transform algorithms, vol. 2. Springer (1997)
15. Hewitt, E., Hewitt, R.: The Gibbs-Wilbraham phenomenon: An episode in fourier analysis. Archive for History of Exact Sciences 21, 129–160 (1979)
16. Brodatz, P.: A Photographic Album for Artists and Designers. Dover Publications (1966)
17. Acharya, R., Laurette, R.: Mathematical morphology for 3-D image analysis. In: International Conference on Acoustics, Speech, and Signal Processing, pp. 952–955 (1988)
18. Fawcett, T.: An introduction to ROC analysis. Pattern Recognition Letters 27(8), 861–874 (2006)
19. Van Rijsbergen, C.J.: Information retrieval. Butterworth-Heinemann (1979)
20. Quiroga, R.Q., Blanco, S., Rosso, O., Garcia, H., Rabinowicz, A.: Searching for hidden information with Gabor Transform in generalized tonic-clonic seizures. Electroencephalography and Clinical Neurophysiology 103(4), 434–439 (1997)