Pattern Matching in Compressed Texts and Images

Pattern Matching in Compressed Texts and Images

Don Adjeroh

West Virginia University
USA
don@csee.wvu.edu

Tim Bell

University of Canterbury
New Zealand
tim.bell@canterbury.ac.nz

Amar Mukherjee

University of Central Florida
USA
amar@eecs.ecf.edu

the essence of knowledge

Boston – Delft

Foundations and Trends® in Signal Processing

Published, sold and distributed by:
now Publishers Inc.
PO Box 1024
Hanover, MA 02339
USA
Tel. +1-781-985-4510
www.nowpublishers.com
sales@nowpublishers.com

Outside North America:
now Publishers Inc.
PO Box 179
2600 AD Delft
The Netherlands
Tel. +31-6-51115274

The preferred citation for this publication is D. Adjeroh, T. Bell and A. Mukherjee, Pattern Matching in Compressed Texts and Images, Foundations and Trends® in Signal Processing, vol 6, nos 2–3, pp 97–241, 2012

ISBN: 978-1-60198-684-9

Foundations and Trends® in Signal Processing
Volume 6 Issues 2–3, 2012

Editorial Scope

Foundations and Trends® in Signal Processing will publish survey and tutorial articles on the foundations, algorithms, methods, and applications of signal processing including the following topics:

- Adaptive signal processing
- Audio signal processing
- Biological and biomedical signal processing
- Complexity in signal processing
- Digital and multirate signal processing
- Distributed and network signal processing
- Image and video processing
- Linear and nonlinear filtering
- Multidimensional signal processing
- Multimodal signal processing
- Multiresolution signal processing
- Nonlinear signal processing
- Randomized algorithms in signal processing
- Sensor and multiple source signal processing, source separation
- Signal decompositions, subband and transform methods, sparse representations
- Signal processing for communications
- Signal processing for security and forensic analysis, biometric signal processing
- Signal quantization, sampling, analog-to-digital conversion, coding and compression
- Signal reconstruction, digital-to-analog conversion, enhancement, decoding and inverse problems
- Speech/audio/image/video compression
- Speech and spoken language processing
- Statistical/machine learning
- Statistical signal processing
  - classification and detection
  - estimation and regression
  - tree-structured methods

Information for Librarians

Foundations and Trends® in Signal Processing, 2012, Volume 6, 4 issues. ISSN paper version 1932-8346. ISSN online version 1932-8354. Also available as a combined paper and online subscription.

Foundations and Trends® in
Signal Processing
Vol. 6, Nos. 2–3 (2012) 97–241

DOI: 10.1561/2000000038

Pattern Matching in Compressed Texts and Images

Don Adjeroh[1,*], **Tim Bell**[2,†]
and Amar Mukherjee[3,‡]

[1] *Lane Department of Computer Science and Electrical Engineering, West Virginia University, Morgantown, WV 26506-6109, USA, don@csee.wvu.edu*

[2] *Department of Computer Science and Software Engineering, University of Canterbury, Christchurch, New Zealand, tim.bell@canterbury.ac.nz*

[3] *School of Electrical Engineering and Computer Science, University of Central Florida, Orlando, FL 32816, USA, amar@eecs.ecf.edu*

Abstract

This review provides a survey of techniques for pattern matching in compressed text and images. Normally compressed data needs to be decompressed before it is processed, but if the compression has been done in the right way, it is often possible to search the data without having to decompress it, or at least only partially decompress it. The problem can be divided into lossless and lossy compression methods, and then in each of these cases the pattern matching can be either exact

* This work was partially supported by grants from the US National Science Foundation IIS-0228370 and IIS-0312484.

† Supported by a grant from the Hitachinaka Techno Center Inc, Japan.

‡ This work has been partially supported by grants from the US National Science Foundation IIS-9977336 and IIS-0207819.

or inexact. Much work has been reported in the literature on techniques for all of these cases, including algorithms that are suitable for pattern matching for various compression methods, and compression methods designed specifically for pattern matching. This work is surveyed in this review. The review also exposes the important relationship between pattern matching and compression, and proposes some performance measures for compressed pattern matching algorithms. Ideas and directions for future work are also described.

Keywords: Compressed pattern matching; text compression; image compression; performance measures; searching.

Contents

1 Introduction **1**

1.1 Data Compression Methods 4
1.2 Compressed Pattern Matching 6
1.3 Compressed Domain versus Transform Domain Analysis 7
1.4 Organization 8

2 Search Strategies **11**

2.1 The Pattern Matching Problem and its Variants 11
2.2 Search Strategies for Text 15
2.3 Search Strategies for Images 21

3 Relationship Between Searching and Compression **25**

3.1 Pattern Matching for Compression 25
3.2 Compression for Pattern Matching 28

4 Searching Compressed Data: Performance Measurement **31**

4.1 Performance Measures for Compression Algorithms 31
4.2 Performance Measures for Pattern Matching Algorithms 34
4.3 Performance Measures for Compressed Pattern Matching 35

5 Searching Compressed Data: Text **39**

5.1 Dictionary Methods 40
5.2 Searching on LZ-encoded Text 43
5.3 Grammar-based Compression 52
5.4 Run-length Encoding 60
5.5 Statistical (symbolwise) Compression 61
5.6 Text Compression by Sorted-contexts 67
5.7 Variable-to-fixed Length Codes 69
5.8 Others 73

6 Searching Compressed Data: Images **77**

6.1 Searching Lossless Compressed Images 79
6.2 Searching Compressed Document Images 89
6.3 Searching on Lossy Compressed Images 92

7 Directions for Further Research **109**

7.1 New Compressed Pattern Matching Algorithms 109
7.2 New Search-aware Compression Algorithms 112
7.3 Compressed Image Descriptors 113
7.4 Performance Measures 114
7.5 Integration and Adaptation 114
7.6 Hardware Implementation 116
7.7 New Applications 116

8 Conclusion **119**

Acknowledgments **121**

References **123**

1

Introduction

Given a text sequence (the database) and a pattern sequence (the query), the pattern matching problem is to search in the text to determine all the locations (if any) where the pattern occurs. Searching for a pattern is an important activity that is performed on a daily basis, with significant applications in a diverse range of areas, including signal processing, computer vision, robotics, pattern recognition, data mining, computational biology, and bioinformatics, to name a few. Today computers are increasingly being used to process text, digitized images, digital video, and various other types of data. However, representing the data can require large amounts of storage space, and operations on the data, such as searching for a pattern, are often time consuming. Even the amount of image or textual data typically available to the ordinary user has witnessed a tremendous growth due to a number of factors, such as improvements in storage and communication technologies (for example, the web, electronic mail, smartphones, general wireless mobile devices), widespread deployment of digital libraries, improved document processing techniques, and the availability of different types of sensors, including cameras and scanners. Apart from the problem of sheer size, the huge amounts of data involved also pose problems for efficient search and retrieval of the required information from the stored data.

Since the digitized data are usually stored using compression techniques, and because of the problem of efficiency (in terms of storage space, computational time, and power consumption), the trend now is to keep the compressed data in its compressed form for as much time as possible. This means that operations such as search and analysis on the data (be it text or images) is ideally performed directly on the compressed representation, without decompression, or at least, with minimal decompression. Intuitively, compared to working on the original uncompressed data, operating directly on the compressed data will require the manipulation of less data, and hence should be more efficient. This also avoids the often time-consuming process of decompression, and the problem of temporary storage space that may be required to keep the decompressed data. The need to search data directly in its compressed form has been recognized by international compression standards such as MPEG-4, MPEG-7 [1, 311] and H.264 [373], where part of the requirement is the ability to search for objects directly in the compressed video.

This review surveys techniques that solve the two basic problems of efficiency (in storage and computation) at the same time. That is, the digitized image or text is stored and searched in a compressed format. Without addressing both problems together, compression and searching work against each other, since a simple system would have to decompress a file before searching it, thus slowing down the pattern matching process. However, there is a strong relationship between compression and pattern matching, and this can be exploited to enable both tasks to be performed efficiently at the same time.

In fact, pattern matching can be regarded as the basis of compression. For example, a *dictionary* compression system might identify the occurrences of an English word in a text, and replaces these with a reference to the word in a lexicon. The main task of the compression system is thus to identify patterns (in this example, words), which are then represented using a compact code. The identified pattern need not be an exact match every time (for example, in lossy compression).

Thus, the strong connection between searching and compression can be traced way back to the early days of lossy compression of signals. For instance, Shannon [303, 304] proposed block source coding with a fidelity criterion (essentially the basis for vector quantization), whereby the encoder uses a reproduction codebook (akin to the dictionary above), and searches in the codebook for the codeword with a minimum distortion to the input vector, and transmits the index for the codeword. If the type of pattern used for compression is the same as the type being used during a later search of the text, then the compression system can be exploited directly to perform a fast search. In the example of the dictionary system, if users wish to search the compressed text for words, then they could look up the word in the lexicon, which would immediately establish whether a search will be successful. If the word is found, then its code could be determined, and the compressed text searched for the code. This will considerably reduce the amount of data to be searched, and the search will be matching whole words rather than a character at a time. In one sense, much of the searching has already been performed off-line at the time of compression.

The potential savings are significant. Text can typically be compressed to less than one-third of its original size, and images are routinely compressed to a tenth or even a hundredth of the size of the raw data. These factors indicate that there is considerable potential to speed up searching, and indeed, systems exist that are able to achieve much of this potential saving. For instance, compressed domain indexing and retrieval is the preferred approach to multimedia information management [210, 234] where orders of magnitude speedup have been recorded over operations on uncompressed data [7, 384].

There are various surveys on the general problem of content-based image and video retrieval [100, 315, 318]. None has focused on the restricted problem of compressed domain image retrieval. Similarly, pattern matching and compression were considered in the survey paper [97]. However, compressed pattern matching was not considered, and various techniques have been developed since then. More importantly, we try to compare and contrast methods that have been proposed for compression, and for compressed pattern matching in text and images. Thus we consider compressed pattern matching for both lossy compression (used for images) and for lossless compression (used for text and images).

1.1 Data Compression Methods

We begin with a brief introduction to data compression methods. More complete introductions to the topic of compression are available [50, 91], and the reader is referred to textbooks for more details [3, 51, 264, 294, 298]. There is also an IEEE conference series on data compression from 1991 to the present — see http://www.cs.brandeis.edu/˜dcc/.

Compression methods are generally classed as *lossless* or *lossy*. Lossless methods enable the original data to be recovered exactly, without any error. Lossless compression, sometimes called *text compression*, is typically used for text, and to a lesser extent on images, such as in medical imaging where exact reconstruction of the original image is important. In contrast, lossy methods allow some deterioration in the original data, and are generally applied in situations where the data have been digitized from an analog source (such as images, video and audio). Usually the level of deterioration is near-imperceptible, yet considerable compression improvement can be achieved because the system is not storing unnecessary detail. Many lossy methods include a lossless method as a sub-component. For example, an image or video compression system might transform the image to a frequency domain, quantize some of the frequency domain coefficients (which is a lossy step), and then encode the coefficients using a lossless method.

Figure 1.1 shows a general model of what happens in data compression. The data transformation stage transforms the input data into a form that will make it easier to compress, for instance by exposing the redundancies or repetitions in the data. Some transformation schemes convert the input data or signal into the frequency domain, with the aim of packing most of the energy in the signal into only a few transform coefficients. The specific transformation performed depends on the type of input data. For text, the transformation could be a simple cyclic permutation of the text sequence (for example, the Burrows–Wheeler transform, or BWT [3, 70]); for images, we could have spatial prediction where the original image is essentially represented as a sequence

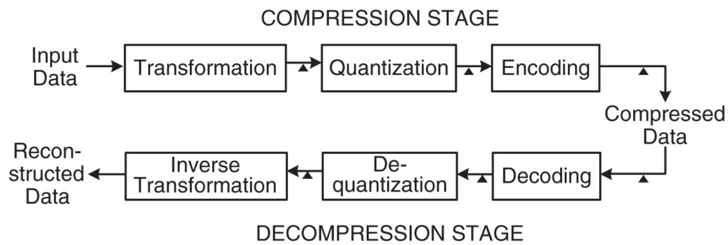

Fig. 1.1 General model for data compression. The triangle markers show the points where pattern matching on the transformed or compressed data can take place.

of prediction errors; for audio signals we could have temporal prediction on the signals; and for video, spatio-temporal prediction using motion vectors can also be viewed as a form of transformation. Standard linear transforms that pack the energy into a few coefficients, such as the Fourier transform or the wavelet transform, are quite common in image, video and audio compression. The particular transformation applied will have an impact on the compression performance, and on the ability to search the compressed data.

The quantization stage is used to reduce the number of distinct values in a signal to a much smaller number. Input data or a signal may contain a large number of distinct values. For analog signals (without digitization), the signal is continuous and the digitization process must quantize it into a set of distinct values, so a digital video has already had some quantization applied. Transforming the digitized signal puts the representation into a domain where the effect of further quantization will be less perceptible. Quantization of the transformed signal reduces the representation to a smaller set of distinct values, each of which can be represented using fewer bits, thus requiring reduced data storage. Quantization however leads to reduced accuracy in the data representation. In fact, the quantization stage represents the major source of compression, but also the major source of data loss (error) in the reconstructed data. In the case of a channel with limited bandwidth, the data rate available will determine the level of quantization needed, and in this case the quantization can be seen as maintaining the best possible fidelity for the capacity available. In some cases for images, video, or audio, it may be possible to keep the loss of accuracy to a level where it isn't observable by a human, given the limitations of human perception. Thus, quantization is common in such applications. Quantization could be performed on either scalar or on vector values, leading to the respective notions of scalar quantization and vector quantization. The encoding stage (also called the coding stage) codes the data to remove further redundancies, often based on the probability distribution of symbols in the data, or symbols in the quantized data for lossy compression.

Decompression involves performing the reverse operations of decoding, de-quantization and inverse transformation. The operations before and after the quantization stage are generally reversible and hence do not introduce any loss or artefacts in the compression. (Here, we ignore errors due to limited arithmetic precision, such as round-off errors during the transformation stage). Quantization, however, is not reversible and thus introduces some error in the compression process. In effect, from the viewpoint of compression models, the major difference between lossless and lossy compression is the quantization stage: lossless compression does not involve any quantization. The quantization stage also accounts for the huge compression ratios often achievable in lossy data compression.

1.2 Compressed Pattern Matching

In traditional pattern matching one is given a text and a pattern, and the problem is to determine whether the pattern occurs in the text. The search could be exact, whereby the pattern matches a substring of the text with no errors, or could be inexact or approximate, whereby some mismatches or errors could be allowed in the match. The result could be simply a binary decision on whether the pattern occurs in the text, a count of the number of occurrences, or possibly a listing of all the locations of occurrences (if any). Variations of the pattern matching problem have been studied, including multiple pattern matching, parameterized pattern matching, and multi-dimensional pattern matching. Compressed pattern matching involves one or more of the pattern matching variants, with the constraint that either the text, the pattern, or both are in compressed form [15, 133]. In general, exact pattern matching is a natural fit for lossless compression, though inexact pattern matching can also be performed on lossless compressed data. On the other hand, given their nature, one can only hope for approximate matches for lossy compressed data. We identify two major categories for compressed pattern matching problems. *Fully compressed pattern matching* is when the text and the pattern are both compressed, and matching involves no form of decompression. Let $\mathcal{C}$ be a compression scheme. Then, given $T_c = \mathcal{C}(T)$, the compressed version of the text T, and $P_c = \mathcal{C}(P)$, the compressed form of the pattern P, the problem of fully compressed pattern matching is to determine all the positions in T, where the pattern P occurred, without first decompressing T_c and P_c. The more general term *compressed pattern matching* is used for the following less restricted problem: Given $T_c = \mathcal{C}(T)$, and the (uncompressed) pattern P, determine all the positions in T, where the pattern P occurred, without first decompressing T_c. Most efficient pattern matching algorithms perform some type of preprocessing (either on the prefixes or the suffixes) of the pattern, P. Thus fully compressed pattern matching is made more difficult, since the usual preprocessing on the pattern is no longer easy to achieve without decompressing P. For text, various algorithms have been proposed for both variants of the compressed pattern matching problem. For compressed-domain image retrieval, the usual assumption is that the query image is not compressed. However, some methods on 2D image pattern matching also consider fully-compressed pattern matching.

1.3 Compressed Domain versus Transform Domain Analysis

Our primary focus in this review is on compressed pattern matching for text and images. Compressed pattern matching is one activity under the general area of compressed domain analysis. In image and video analysis, the terms “compressed domain” and “transform domain” are often used interchangeably. Here we distinguish between the two. With compressed domain analysis, the required analysis is performed on the compressed data with minimal or no decoding, and before the stage of inverse quantization. For transform domain analysis, the required analysis is performed on the transform coefficients, typically

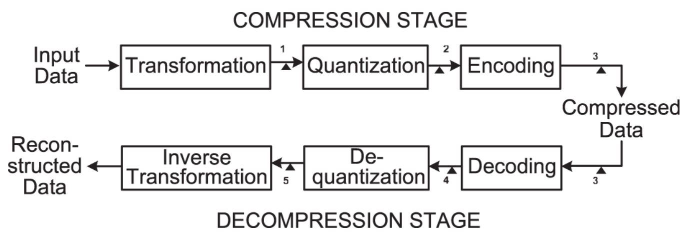

Fig. 1.2 Transform domain versus compressed domain analysis. Whether operating in the transform domain or compressed domain depends on the point(s) in the compression pipeline where the analysis is performed. The type(s) of analyses are characterized as follows: transform domain: points (1), (2), (4), and (5); compressed domain: point (3); compressed domain with partial decoding: points (4) and (5).

after the transformation or quantization stages in the compression pipeline, but before the final encoding stage. Here the objective of the transformation could be simply for efficient analysis, and not necessarily for compression or data storage. This is typical in some signal analysis or image analysis applications. If the data is already compressed, transform domain analysis could be performed after decoding the compressed stream, but before the inverse transformation stage. Figure 1.2 uses the compression pipeline of Figure 1.1 to explain the difference between compressed domain and transform domain analysis. For images or video, compressed (or transform) domain analysis could include general operations, such as image enhancement, noise removal, shape, color or texture-based image retrieval, etc. Recent surveys on general compressed or transform domain analysis of images and video are provided in [254, 343]. From the description above, most video and image retrieval and search operations, especially those for lossy compression, are performed in the transform domain (mainly after the stages of transformation, or quantization in a few cases, but generally before the encoding stage). Some text pattern matching methods operate in the transform domain (with partial decoding), while others operate in the compressed domain, that is, after final encoding, such as using Huffman codes.

1.4 Organization

In this review, we survey compression methods for text and images, especially identifying the search techniques that they use, and how they could be exploited for searching the compressed data later. First, we consider search strategies and pattern matching methods for uncompressed data, to set the scene for more sophisticated systems. Next we explore the interesting relationship between searching and data compression. This is followed by a discussion on performance measurement for compressed pattern matching. The next two sections survey techniques that have been developed for searching compressed data, which is sometimes called *compressed-domain pattern matching*. The first section looks at methods for lossless data compression (as used for text). The next section then surveys methods that have been proposed for pattern matching on compressed images. This is presented in two parts. The first part considers methods for pattern matching on lossless compressed images, which often borrow a lot from methods for searching compressed text. The second part surveys methods where the image has been compressed using lossy methods. The review concludes with a speculation on the likely directions of future work in the area.

Given the breadth of the topics involved, we will focus only on compressed pattern matching for text and images. Video and audio will be mentioned at times, but without much detail. Also, we will only briefly describe other types of signal processing activities (different from pattern matching) that are often performed in the compressed domain. For lossless compression (for text and images), we focus mainly on exact pattern matching. For lossy compression, we focus mainly on compressed domain or transform domain retrieval. Throughout this work, we assume that data from analog sources, such as images, audio and video, have already been digitized. Thus we ignore the potential data loss due to the digitization process.

References

[1] M. Abdel-Mottaleb and S. Krishnamachari, "Multimedia descriptions based on MPEG-7: Extraction and applications," *IEEE Transactions on Multimedia*, vol. 6, no. 3, pp. 459–468, 2004.

[2] T. Acharya and P.-S. Tsai, *JPEG2000 Standard for Image Compression: Concepts, Algorithms and VLSI Architectures*. Hoboken, New Jersey: John Wiley & Sons, 2005.

[3] D. Adjeroh, T. Bell, and A. Mukherjee, *The Burrows-Wheeler Transform: Data Compression, Suffix Arrays and Pattern Matching*. New York, NY: Springer, 2008.

[4] D. Adjeroh and K. V. Bhupathiraju, "On lossless image compression using the Burrows-Wheeler Transform," in *Proceedings, IEEE International Conference on Image Processing*, (B. Macq and P. Schelkens, eds.), pp. 1997–2000, 2011.

[5] D. A. Adjeroh, T. Bell, M. Powell, N. Zhang, and A. Mukherjee, "Pattern matching in BWT-transformed text," in *Proceedings, IEEE Data Compression Conference*, p. 445, 2002.

[6] D. A. Adjeroh, U. Kandaswamy, N. Zhang, A. Mukherjee, M. T. Brown, and T. Bell, "BWT-based efficient shape matching," in *Proceedings, ACM Symposium on Applied Computing*, (Y. Cho, R. L. Wainwright, H. Haddad, S. Y. Shin, and Y. W. Koo, eds.), pp. 1079–1085, ACM, 2007.

[7] D. A. Adjeroh and M. C. Lee, "Robust and efficient transform domain video sequence analysis: An approach from the generalized color ratio model," *Journal of Visual Communication and Image Representation*, vol. 8, no. 2, pp. 182–207, 1997.

[8] D. A. Adjeroh and M.-C. Lee, "An occupancy model for image retrieval and similarity evaluation," *IEEE Transactions on Image Processing*, vol. 9, no. 1, pp. 120–131, 2000.
[9] D. A. Adjeroh and M.-C. Lee, "Scene-adaptive transform domain video partitioning," *IEEE Transactions on Multimedia*, vol. 6, no. 1, pp. 58–69, 2004.
[10] D. A. Adjeroh, M. C. Lee, and I. King, "A distance measure for video sequence similarity matching," *Computer Vision and Image Understanding*, vol. 75, no. 1, pp. 25–45, 1999.
[11] D. A. Adjeroh and F. Nan, "Suffix-sorting via Shannon-Fano-Elias codes," *Algorithms*, vol. 3, no. 2, pp. 145–167, 2010.
[12] A. V. Aho and M. Corasick, "Efficient string matching: An aid to bibliographic search," *Communications of the ACM*, vol. 18, no. 6, pp. 333–340, 1975.
[13] T. Akutsu, "Approximate string matching with don't care characters," *Proceedings, Combinatorial Pattern Matching (LNCS 807)*, pp. 240–249, 1994.
[14] M. Alzina, W. Szpankowski, and A. Grama, "2D-Pattern matching image and video compression: Theory, algorithms, and experiments," *IEEE Transactions on Image Processing*, vol. 9, no. 8, 2000.
[15] A. Amir and G. Benson, "Efficient two-dimensional compressed matching," in *Proceedings, IEEE Data Compression Conference*, (J. A. Storer and M. Cohn, eds.), pp. 279–288, IEEE Computer Society, 1992.
[16] A. Amir, G. Benson, and M. Farach, "Let sleeping files lie: Pattern matching in Z-compressed files," *Journal of Computer and System Sciences*, vol. 52, pp. 299–307, 1996.
[17] A. Amir, G. Benson, and M. Farach, "Optimal two-dimensional compressed matching," *Journal of Algorithms*, vol. 24, pp. 354–379, 1997.
[18] A. Amir and G. Calinescu, "Alphabet independent and dictionary scaled matching," *Proceedings, Combinatorial Pattern Matching, (LNCS 1075)*, pp. 320–334, 1996.
[19] A. Amir, O. Kapah, and D. Tsur, "Faster two-dimensional pattern matching with rotations," *Theoretical Computer Science*, vol. 368, no. 3, pp. 196–204, 2006.
[20] A. Amir, G. Landau, and U. Vishkin, "Efficient pattern matching with scaling," *Journal of Algorithms*, vol. 13, pp. 2–32, 1992.
[21] A. Amir, G. M. Landau, and D. Sokol, "Inplace 2D matching in compressed images," *Journal of Algorithms*, vol. 49, no. 2, pp. 240–261, 2003.
[22] A. Amir, G. M. Landau, and D. Sokol, "Inplace run-length 2D compressed search," *Theoretical Computer Science*, vol. 290, no. 3, pp. 1361–1383, 2003.
[23] A. Amir, O. Lipsky, E. Porat, and J. Umanski, "Approximate matching in the L_1 metric," in *Proceedings, Combinatorial Pattern Matching*, LNCS vol. 3537, (A. Apostolico, M. Crochemore, and K. Park, eds.), pp. 91–103, Springer, 2005.
[24] D. Anastassiou, M. Brown, H. Jones, J. Mitchell, W. Pennebaker, and K. Pennington, "Series/1-based videoconferencing system," *IBM Systems Journal*, vol. 22, no. 1/2, pp. 97–110, 1983.
[25] S. Aouat and S. Larabi, "Indexing binary images using quad-tree decomposition," in *Proceedings, IEEE International Conference on Systems, Man and Cybernetics*, pp. 3074–3080, Istanbul, Turkey, October 2010.

[26] A. Apostolico, P. L. Erdos, and A. Juttner, “Parameterized searching with mismatches for run-length encoded strings,” *Theoretical Computer Science* (in press), 2012.

[27] A. Apostolico and R. Giancarlo, “The Boyer Moore Galil string searching strategies revisited,” *SIAM Journal on Computing*, vol. 15, no. 1, pp. 98–105, February 1986.

[28] A. Apostolico, G. M. Landau, and S. Skiena, “Matching for run-length encoded strings,” *Proceedings, Complexity and Compression of Sequences*, 1997.

[29] A. Areepongsa, N. Kaewkamnerd, S. Syed, and R. K. Rao, *Wavelet based compression for retrieval system*. November 2000.

[30] F. Arnia, M. Fujiyoshi, and H. Kiya, “The use of DCT coefficient sign for content-based copy detection,” in *Proceedings, International Symposium on Communications and Information Technologies (ISCIT ’07)*, pp. 1476–1481, October 2007.

[31] F. Arnia, I. Iizuka, M. Fujiyoshi, and H. Kiya, “DCT sign-based similarity measure for JPEG image retrieval,” *IEICE Transactions*, vol. 90-A, no. 9, pp. 1976–1985, 2007.

[32] F. Arnia, I. Iizuka, M. Fujiyoshi, and H. Kiya, “Fast method for joint retrieval and identification of JPEG coded images nased on DCT sign,” in *Proceedings, IEEE International Conference on Image Processing*, pp. 229–232, 2007.

[33] R. B. Arps, T. K. Truong, D. J. Lu, R. C. Pasco, and T. D. Friedman, “A multipurpose VLSI CAD chip for adaptive data compression of bilevel images,” *IBM Journal of Research and Development*, vol. 32, no. 6, pp. 775–795, November 1988.

[34] M. Atallah, Y. Génin, and W. Szpankowski, “Pattern matching image compression: Algorithmic and experimental results,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, pp. 618–627, 1999.

[35] R. Baeza-Yates and G. H. Gonnet, “A new approach to text searching,” *Communications of the ACM*, vol. 35, no. 10, pp. 74–82, 1992.

[36] R. Baeza-Yates and B. Ribeiro-Neto, *Modern Information Retrieval*. ACP Press/Addison-Wesley, 1999.

[37] C. Bai, K. Kpalma, and J. Rosin, “Analysis of histogram descriptor for image retrieval in DCT domain,” *Intelligent Interactive Multimedia*, vol. SIST 11, pp. 227–235, 2011.

[38] B. S. Baker, “Parameterized pattern matching: Algorithms and applications,” *Journal of Computer and System Science*, vol. 52, no. 1, pp. 28–42, 1996.

[39] T. Baker, “A technique for extending rapid exact-match string matching to arrays of more than one dimension,” *SIAM Journal on Computing*, vol. 7, no. 4, pp. 533–541, November 1978.

[40] H. Bannai, T. Gagie, I. Tomehiro, S. Inenaga, G. M. Landau, and M. Lewenstein, “An efficient algorithm to test square-freeness of strings compressed by straight-line programs,” *Information Processing Letters*, vol. 112, no. 19, pp. 711–714, 2012.

[41] P. Barcaccia, A. Cresti, and S. D. Agostino, “Pattern matching in text compressed with the ID heuristic,” in *Proceedings, IEEE Data Compression Conference*, pp. 113–118, 1998.

[42] H. Bay, A. Ess, T. Tuytelaars, and L. J. V. Gool, "Speeded-up robust features (SURF)," *Computer Vision and Image Understanding*, vol. 110, no. 3, pp. 346–359, 2008.

[43] R. Beal and D. A. Adjeroh, "p-Suffix sorting as arithmetic coding," *Journal of Discrete Algorithms*, vol. 16, pp. 151–169, 2012.

[44] R. Beal and D. A. Adjeroh, "Parameterized longest previous factor," *Theoretical Computer Science*, vol. 437, pp. 21–34, 2012.

[45] R. Beal and D. A. Adjeroh, "Compressed parameterized pattern matching," *Proceedings, IEEE Data Compression Conference*, pp. 461–470, 2013.

[46] T. Bell and D. Kulp, "Longest-match string searching for Ziv-Lempel compression," *Software — Practice and Experience*, vol. 23, no. 7, pp. 757–772, July 1993.

[47] T. Bell and A. Moffat, "A note on the DMC data compression scheme," *The Computer Journal*, vol. 32, no. 1, pp. 16–20, February 1989.

[48] T. Bell, M. Powell, A. Mukherjee, and D. A. Adjeroh, "Searching BWT compressed text with the Boyer-Moore algorithm and binary search," in *Proceedings, IEEE Data Compression Conference*, pp. 112–121, 2002.

[49] T. Bell and I. Witten, "The relationship between greedy parsing and symbolwise text compression," *Journal of the ACM*, vol. 41, no. 4, pp. 708–724, July 1994.

[50] T. C. Bell, "Data compression," in *Encyclopedia of Computer Science*, (A. Ralston, E. Reilly, and D. Hemmendinger, eds.), pp. 492–496, Nature Publishing Group, 2000.

[51] T. C. Bell, J. G. Cleary, and I. H. Witten, *Text Compression*. Englewood Cliffs, NJ: Prentice Hall, 1990.

[52] S. Belongie, J. Malik, and J. Puzicha, "Shape matching and object recognition using shape contexts," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 4, pp. 509–522, 2002.

[53] M. A. Bender and M. Farach-Colton, "The LCA problem revisited," in *LATIN'2000* , LNCS vol. 1776, (G. H. Gonnet, D. Panario, and A. Viola, eds.), pp. 88–94, Springer, 2000.

[54] P. Berman, M. Karpinski, L. Larmore, W. Plandowski, and W. Rytter, "The complexity of two-dimensional compressed pattern matching," Technical Report TR-96-051, Berkeley, CA: International Computer Science Institute, Dec., 1996.

[55] P. Berman, M. Karpinski, L. L. Larmore, W. Plandowski, and W. Rytter, "On the complexity of pattern matching for highly compressed two-dimensional texts," *Journal of Computer and System Science*, vol. 65, no. 2, pp. 332–350, 2002.

[56] S. Berretti and A. D. Bimbo, "Modeling spatial relationships between 3D objects," in *Proceedings, International Conference on Pattern Recognition*, pp. 119–122, 2006.

[57] S. Berretti, A. D. Bimbo, and E. Vicario, "Weighted walkthroughs between extended entities for retrieval by spatial arrangement," *IEEE Transactions on Multimedia*, vol. 5, no. 1, pp. 52–70, 2003.

[58] R. S. Bird, "Two dimensional pattern matching," *Information Processing Letters*, vol. 6, no. 5, pp. 168–170, October 1977.
[59] L. E. Boucheron and C. D. Creusere, "Lossless wavelet-based compression of digital elevation maps for fast and efficient search and retrieval," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 5, pp. 1210–1214, 2005.
[60] R. Boyer and J. Moore, "A fast string searching algorithm," *Communications of the ACM*, vol. 20, no. 10, pp. 762–772, October 1977.
[61] J. Bracamonte, M. Ansorge, F. Pellandini, and P.-A. Farine, "Efficient compressed domain target image search and retrieval," in *Proceedings, International Conference on Image and Video Retrieval*, pp. 154–163, Berlin, Heidelberg, 2005.
[62] A. Bremler-Barr, D. Hay, and Y. Koral, "CompactDFA: Generic state machine compression for scalable pattern matching," in *Proceedings, IEEE INFOCOM*, pp. 659–667, 2010.
[63] A. Bremler-Barr and Y. Koral, "Accelerating multi-patterns matching on compressed HTTP traffic," in *Proceedings, IEEE INFOCOM*, pp. 397–405, 2009.
[64] N. R. Brisaboa, A. Fariña, J.-R. López, G. Navarro, and E. R. Lopez, "A new searchable variable-to-variable compressor," in *Proceedings, IEEE Data Compression Conference (DCC 2010)*, (J. A. Storer and M. W. Marcellin, eds.), pp. 199–208, 2010.
[65] N. R. Brisaboa, A. Fariña, G. Navarro, and J. R. Paramá, "Dynamic lightweight text compression," *ACM Transactions on Information Systems*, vol. 28, no. 3, 2010.
[66] N. J. Brittain and M. R. El-Sakka, "Grayscale true two-dimensional dictionary-based image compression," *Journal on Visual Communication and Image Representation*, vol. 18, no. 1, pp. 35–44, 2007.
[67] H. Bunke and J. Csirik, "An algorithm for matching run-length coded strings," *Computing*, vol. 50, pp. 297–314, 1993.
[68] H. Bunke and J. Csirik, "An improved algorithm for computing the edit distance of run-length coded strings," *Information Processing Letters*, vol. 54, pp. 93–96, 1995.
[69] H. Bunke and A. Sanfeliu, eds., *Syntatic and Structural Pattern Recognition: Theory and Applications*. Singapore: World Scientific, 1990.
[70] M. Burrows and D. J. Wheeler, "A block-sorting lossless data compression algorithm," Technical Report 124, Digital Equipment Corporation, Palo Alto, California, May 1994.
[71] P. J. Burt and E. H. Adelson, "The Laplacian pyramid as a compact image code," *IEEE Transactions on Communications*, vol. 31, no. 4, pp. 532–540, 1993.
[72] CalgaryCorpus, "The Calgary Corpus, ftp://ftp.cpsc.ucalgary.ca/pub/projects/text.compression.corpus," 2000.
[73] M. Calonder, V. Lepetit, C. Strecha, and P. Fua, "BRIEF: Binary robust independent elementary features," in *Proceedings, European Conference on Computer Vision*, LNCS vol. 6314, (K. Daniilidis, P. Maragos, and N. Paragios, eds.), pp. 778–792, Springer, 2010.

[74] A. Cannane and H. E. Williams, "General-purpose compression for efficient retrieval," *JASIST*, vol. 52, no. 5, pp. 430–437, 2001.

[75] A. Cannane and H. E. Williams, "A general-purpose compression scheme for large collections," *ACM Transactions on Information Systems*, vol. 20, no. 3, pp. 329–355, 2002.

[76] CanterburyCorpus, "The Canterbury Corpus, http://corpus.canterbury.ac.nz," 2000.

[77] B. Carpentieri, "Special issue on data compression, communication and processing," *Algorithms*, vol. 5, no. 1, 2010.

[78] CCITT, *Draft Recommendation T.82 & ISO DIS 11544: Coded Representation of Picture and Audio information — Progressive Bi-Level Image Compression*. February 1993.

[79] V. Chandrasekhar, M. Makar, G. Takacs, D. Chen, S. Tsau, N. M. Chung, R. Grzeszczuk, Y. Reznik, and B. Girod, "Survey of SIFT compression schemes," in *Proceedings IMMW: International Mobile Multimedia Workshop, IEEE International Conference on Pattern Recognition*, Istanbul, Turkey, 2010.

[80] V. Chandrasekhar, M. E. Sargin, and D. A. Ross, "Automatic language identification in music videos with low level audio and visual features," in *Proceedings, International Conference on Acoustics Speech and Signal Processing*, pp. 5724–5727, 2011.

[81] V. Chandrasekhar, G. Takacs, D. M. Chen, S. S. Tsai, Y. A. Reznik, R. Grzeszczuk, and B. Girod, "Compressed histogram of gradients: A low-bitrate descriptor," *International Journal of Computer Vision*, vol. 96, no. 3, pp. 384–399, 2012.

[82] S.-F. Chang, "Compressed-domain techniques for image/video indexing and manipulation," in *Proceedings, IEEE International Conference on Image Processing*, vol. 1, pp. 314–317, 1995.

[83] S.-F. Chang, J. Smith, H. J. Meng, H. Wang, and D. Zhong, "Finding images/video in large archives: Columbia's content-based visual query project," *D-Lib Magazine*, February 1997.

[84] T. Chang and C. J. Kuo, "Texture analysis and classification with a tree-structured wavelet transform," *IEEE Transactions on Image Processing*, vol. 2, no. 4, pp. 429–441, October 1993.

[85] W. I. Chang and E. L. Lawler, "Sublinear approximate string matching and biological applications," *Algorithmica*, vol. 12, pp. 327–344, 1994.

[86] D. Chen and A. Bovik, "Visual pattern image coding," *IEEE Transactions on Communications*, vol. 38, no. 12, pp. 2137–2146, December 1992.

[87] M. Christodoulakis, C. S. Iliopoulos, M. S. Rahman, and W. F. Smyth, "Identifying rhythms in musical texts," *International Journal of Foundations of Computer Science*, vol. 19, no. 1, pp. 37–51, 2008.

[88] J. G. Cleary and W. J. Teahan, "Unbounded length contexts for PPM," *The Computer Journal*, vol. 40, no. 2/3, pp. 67–75, 1997.

[89] J. G. Cleary and I. H. Witten, "Data compression using adaptive coding and partial string matching," *IEEE Transactions on Communications COM-32*, pp. 396–402, April 1984.

[90] R. Clifford and C. S. Iliopoulos, "Approximate string matching for music analysis," *Soft Computing*, vol. 8, no. 9, pp. 597–603, 2004.

[91] M. Cohn, "Data compression," in *Encyclopedia of Computer Science and Technology*, 1994.

[92] C. Constantinescu and J. Storer, "Improved techniques for single-pass adaptive vector quantization," *Proceeding of the IEEE*, vol. 82, pp. 933–939, 1994.

[93] G. Cormack and R. Horspool, "Data compression using dynamic Markov modeling," *The Computer Journal*, vol. 30, no. 6, pp. 541–550, December 1987.

[94] T. M. Cover and J. A. Thomas, *Elements of Information Theory, Wiley Series in Telecommunications and Signal Processing*. New York, NY, USA: John Wiley & Sons, 2006.

[95] M. Crochemore, A. Czumaj, L. Gasieniec, S. Jarominek, T. Lecroq, W. Plandowski, and W. Rytter, "Speeding up two string-matching algorithms," *Algorithmica*, vol. 12, no. 4/5, pp. 247–267, 1994.

[96] M. Crochemore, C. Hancart, and T. Lecroq, *Algorithms on Strings*. Cambridge University Press, 2007.

[97] M. Crochemore and T. LeCroq, "Pattern-matching and text compression," *ACM Computing Surveys*, vol. 28, no. 1, pp. 39–41, March 1996.

[98] M. Crochemore, F. Mignosi, A. Restivo, and S. Salemi, "Data compression using antidictionaries," *Proceedings of the IEEE*, vol. 88, no. 11, pp. 1756–1768, 2000.

[99] A. H. Daptardar and J. A. Storer, "VQ based image retrieval using color and position features," in *Proceedings, IEEE Data Compression Conference*, pp. 432–441, 2008.

[100] R. Datta, D. Joshi, J. Li, and J. Z. Wang, "Image retrieval: Ideas, influences, and trends of the new age," *ACM Computing Surveys*, vol. 40, no. 2, 2008.

[101] T. Deselaers, D. Keysers, and H. Ney, "Features for image retrieval: An experimental comparison," *Information Retrieval*, vol. 11, no. 2, pp. 77–107, 2008.

[102] R. A. DeVore, B. Jawerth, and B. J. Lucier, "Image compression through wavelet transform coding," *IEEE Transactions on Information Theory*, vol. 38, no. 2, pp. 719–746, 1992.

[103] D. S. Doermann, "The indexing and retrieval of document images: A survey," *Computer Vision and Image Understanding*, vol. 70, no. 3, pp. 287–298, 1998.

[104] O. Drbohlav and A. Leonardis, "Towards correct and informative evaluation methodology for texture classification under varying viewpoint and illumination," *Computer Vision and Image Understanding*, vol. 114, no. 4, pp. 439–449, (Special issue on Image and Video Retrieval Evaluation), 2010.

[105] F. Drews, J. Lichtenberg, and L. Welch, "Scalable parallel word search in multicore/multiprocessor systems," *Journal of Supercomputing*, vol. 51, no. 1, pp. 58–75, Janaury 2010.

[106] T. Eilam-Tzoreff and U. Vishkin, "Matching patterns in strings subject to multi-linear transformations," *Theoretical Computer Science*, vol. 60, pp. 231–254, 1988.

[107] D. Eppstein, M. T. Goodrich, and J. Z. Sun, "Skip quadtrees: Dynamic data structures for multidimensional point sets," *International Journal of Computational Geometry Appliances*, vol. 18, no. 1/2, pp. 131–160, 2008.

[108] F. Ercal, M. Allen, and H. Feng, "A systolic image difference algorithm for RLE-compressed images," *IEEE Transactions on Parallel and Distributed Systems*, vol. 11, no. 5, pp. 433–443, 2000.

[109] M. Farach and M. Thorup, "String matching in Lempel-Ziv compressed strings," in *Proceedings, Annual ACM Symposium on the Theory of Computing*, pp. 703–712, New York, May 1995.

[110] A. Fariña, N. R. Brisaboa, G. Navarro, F. Claude, Á. S. Places, and E. Rodríguez, "Word-based self-indexes for natural language text," *ACM Transactions on Information Systems*, vol. 30, no. 1, p. 1, 2012.

[111] S. Faro and T. Lecroq, "The exact online string matching problem: A review of the most recent results," *ACM Computing Surveys*, vol. 45, no. 2, p. 13, 2013.

[112] P. Fenwick, "The Burrows-Wheeler Transform for block sorting text compression," *The Computer Journal*, vol. 39, no. 9, pp. 731–740, September 1996.

[113] P. Ferragina and G. Manzini, "Opportunistic data structures with applications," in *Proceedings, IEEE Symposium on Foundations of Computer Science (FOCS, 2000)*, pp. 390–398, 2000.

[114] P. Ferragina and G. Manzini, "An experimental study of an opportunistic index," in *Proceedings, ACM-SIAM Symposium on Discrete Algorithms (SODA, 2001)*, pp. 269–278, 2001.

[115] P. Ferragina and G. Manzini, "Indexing compressed text," *Journal of the ACM*, vol. 52, no. 4, pp. 552–581, 2005.

[116] A. E. Firth, T. Bell, A. Mukherjee, and D. A. Adjeroh, "A comparison of BWT approaches to string pattern matching," *Software — Practice and Experience*, vol. 35, no. 13, pp. 1217–1258, 2005.

[117] Y. Fisher, ed., *Fractal Image Compression*. New York: Springer-Verlag, 1995.

[118] K. Fredriksson and S. Grabowski, "A general compression algorithm that supports fast searching," *Information Processing Letters*, vol. 100, no. 6, pp. 226–232, December 2006.

[119] K. Fredriksson, V. Mäkinen, and G. Navarro, "Rotation and lighting invariant template matching," *Information and Computations*, vol. 205, no. 7, pp. 1096–1113, 2007.

[120] K. Fredriksson and M. Mozgovoy, "Efficient parameterized string matching," *Information Processing Letters*, vol. 100, no. 3, pp. 91–96, 2006.

[121] K. Fredriksson, G. Navarro, and E. Ukkonen, "Sequential and indexed two-dimensional combinatorial template matching allowing rotations," *Theoretical Computer Science*, vol. 347, no. 1–2, pp. 239–275, 2005.

[122] J. Fridrich, "Visual hash for oblivious watermarking," in *SPIE Photonic West Electronic Imaging 2000, Security and Watermarking of Multimedia Contents*, (San Jose, California), pp. 286–294, January, 24–26 2000.

[123] Z. Galil, "On improving the worst case running time of the Boyer-Moore string matching algorithm," *Communications of the ACM*, vol. 22, no. 9, pp. 505–508, September 1979.

[124] Z. Galil and R. Giancarlo, "Parallel string matching with k mismatches," *Theoretical Computer Science*, vol. 51, pp. 341–384, 1997.

[125] Z. Galil and K. Park, "An improved algorithm for approximate string matching," *SIAM Journal of Computing*, vol. 19, no. 6, pp. 689–999, 1990.

[126] M. Garris, J. Blue, G. Candela, P. Grother, S. Janet, and C. Wilson, "NIST Form-based handprint recognition system (release 2.0)," Technical Report CD-ROM 5050, National Institute of Standards and Technology, 1997.

[127] P. Gawrychowski, "Optimal pattern matching in LZW compressed strings," in *Proceedings, Symposium on Discrete Algorithms*, (D. Randall, ed.), pp. 362–372, SIAM, 2011.

[128] P. Gawrychowski, "Pattern matching in Lempel-Ziv compressed strings: Fast, simple, and deterministic," in *Proceedings, European Symposium on Algorithms*, vol. 6942, (C. Demetrescu and M. M. Halldórsson, eds.), pp. 421–432, Springer, 2011.

[129] P. Gawrychowski, "Tying up the loose ends in fully LZW-compressed pattern matching," in *Proceedings, Symposium on Theoretical Aspects of Computer Sciences*, vol. 14, (C. Dürr and T. Wilke, eds.), pp. 624–635, Schloss Dagstuhl — Leibniz-Zentrum fuer Informatik, 2012.

[130] L. Gąsieniec, A. Gibbons, and W. Rytter, "Efficiency of fast parallel pattern searching in highly compressed texts," in *Proceedings, Mathematical Foundations of Computer Science 1999*, (M. Kutylowski, L. Pacholski, and T. Wierzbicki, eds.), pp. 48–58, Berlin/Heidelberg: Springer, 1999. 10.1007/3-540-48340-3_5.

[131] L. Gąsieniec, M. Karpinski, W. Plandowski, and W. Rytter, "Randomized efficient algorithms for compressed strings: The finger-print approach," *Proceedings, Combinatorial Pattern Matching (LNCS 1075)*, pp. 39–49, 1996.

[132] L. Gąsieniec and I. Potapov, "Time/space efficient compressed pattern matching," *Fundamenta Informaticae*, vol. 56, no. 1–2, pp. 137–154, 2003.

[133] L. Gąsieniec and W. Rytter, "Almost optimal fully LZW-compressed pattern matching," in *Proceedings, IEEE Data Compression Conference*, pp. 316–325, 1999.

[134] O. Gerek, A. Cetin, A. Tewfik, and V. Atalay, "Subband domain coding of binary textual images for document archiving," *IEEE Transactions on Image Processing*, vol. 8, no. 10, pp. 1438 –1446, October 1999.

[135] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*. Boston: Kluwa Academic Press, 1992.

[136] R. Giancarlo and R. Gross, "Multi-dimensional pattern matching with dimensional wildcards: Data structures and optimal on-line search algorithm," *Journal of Algorithms*, vol. 24, pp. 223–265, 1997.

[137] R. Giancarlo and F. Utro, "Textual data compression in computational biology: A synopsis," *Bioinformatics*, vol. 25, pp. 1575–1586, 2009.

[138] J. D. Gibson, T. Berger, T. Lookabaugh, D. Lindbergh, and R. L. Baker, *Digital Compression for Multimedia: Principles and Standards*. San Francisco: Morgan Kaufmann, 1998.

[139] J. Gilchrist and A. Cuhadar, "Parallel lossless data compression using the Burrows-Wheeler transform," *International Journal of Web Grid Services*, vol. 4, no. 1, pp. 117–135, May 2008.

[140] B. Girod, V. Chandrasekhar, R. Grzeszczuk, and Y. A. Reznik, "Mobile visual search: Architectures, technologies, and the emerging MPEG standard," *IEEE MultiMedia*, vol. 18, no. 3, pp. 86–94, 2011.

[141] R. Goldberg, I. Taksa, and A. Spink, "Metasearch engines and information retrieval: Computational complexity of ranking multiple search results," in *Proceedings, International Conference on Information Technology: New Generations*, pp. 315–320, 2008.

[142] G. C. Gonzalez and R. E. Woods, *Digital Image Processing.* Reading, MA: Addison-Wesley, 1992.

[143] J. Gregor and M. G. Thomason, "Dynamic programming alignment of sequences representing cyclic patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 15, no. 2, pp. 129–135, 1993.

[144] R. Grossi and J. S. Vitter, "Compressed suffix arrays and suffix trees with applications to text indexing and string matching," *SIAM Journal on Computing*, vol. 35, no. 2, pp. 378–407, 2005.

[145] M. Gu, M. Farach, and R. Beigel, "An efficient algorithm for dynamic text indexing," in *Proceedings, Symposium on Discrete Algorthims*, (D. D. Sleator, ed.), pp. 697–704, ACM/SIAM, 1994.

[146] D. Gusfield, *Algorithms on Strings, Trees and Sequences: Computer Science and Computational Biology.* Cambridge University Press, 1997.

[147] M. Harrison, "Implementation of the substring test by hashing," *Communications of the ACM*, vol. 14, no. 12, pp. 777–779, December 1971.

[148] G. Heidemann and H. Ritter, "Data compression — a generic principle of pattern recognition?," in *Proceedings, Computer Vision and Computer Graphics Theory and Applications*, (A. Ranchordas, H. J. Araújo, J. M. Pereira, and J. Braz, eds.), pp. 202–212, Berlin/Heidelberg: Springer, 2009.

[149] E. E. Hilbert, "Cluster compression algorithm: A joint clustering/data compression concept," Technical Report, Publication 77-43, Jet Propulsion Lab, Pasedena CA, 1977.

[150] D. Hirschberg and D. Lelewer, "Efficient decoding of prefix codes," *Communications of the ACM*, vol. 33, no. 4, pp. 449–459, April 1990.

[151] P. G. Howard and J. S. Vitter, "Design and analysis of fast text compression based on quasi-arithmetic coding," *Information Processings and Managements*, vol. 30, no. 6, pp. 777–790, 1994.

[152] W.-J. Huang, N. Saxena, and E. J. McCluskey, "A reliable LZ data compressor on reconfigurable coprocessor," *Proceedings of Symposium on Field Programmable Custom Computing Machine*, pp. 249–258, April 2000.

[153] D. Huffman, "A method for the construction of minimum redundancy codes," *Proceedings of IRE*, vol. 40, no. 9, pp. 1098–1101, September 1952.

[154] J. J. Hull, "Document image similarity and equivalence detection," *International Journal on Document Analysis and Recognition*, vol. 1, no. 1, pp. 37–42, 1998.

[155] A. Hume and D. Sunday, "Fast string searching," *Software — Practice and Experience*, vol. 21, no. 11, pp. 1221–1248, November 1991.

[156] R. Hunter and A. H. Robinson, "International digital facsimile coding standards," *Proceedings of IEEE*, vol. 68, no. 7, pp. 854–867, July 1980.

[157] D. P. Huttenlocher, G. A. Klanderman, and W. Rucklidge, "Comparing images using the Hausdorff distance," *IEEE Transactions on Pattern Analysis Machine Intelligence*, vol. 15, no. 9, pp. 850–863, 1993.

[158] F. Idris and S. Panchanathan, "Image indexing using wavelet vector quantization," *Proceedings of SPIE: Digital Image Storage Archiving Systems*, vol. 2606, pp. 269–275, October 1995.

[159] F. Idris and S. Panchanathan, "Image and video indexing using vector quantization," *Journal of Machine Vision and Applications*, vol. 10, pp. 43–50, July 1997.

[160] C. S. Iliopoulos and M. S. Rahman, "Indexing circular patterns," in *Proceedings, International Workshop on Algorthims and Computation*, pp. 46–57, 2008.

[161] H. Imura and Y. Tanaka, "Compression and string matching method for printed document images," in *Proceedings, International Conference on Document Analysis and Recognitions*, pp. 291–295, 2009.

[162] S. Inenaga, A. Shinohara, and M. Takeda, "An efficient pattern matching algorithm on a subclass of context free grammars," *Proceedings, Developments in Language Therapy, LNCS 3340*, pp. 225–236, 2004.

[163] A. Ingber, T. Courtade, and T. Weissman, "Quadratic similarity queries on compressed data," *Proceedings, IEEE Data Compression Conference*, pp. 441–450, 2013.

[164] S. Inglis and I. Witten, "Compression-based template matching," in *Proceedings, IEEE Data Compression Conference*, pp. 106–115, 1994.

[165] N. Jayant, J. Johnston, and J. Sefranek, "Signal compression based on models of human perception," *Proceedings of the IEEE*, vol. 81, no. 10, pp. 1385–1422, 1993.

[166] S. Jeong, C. S. Won, and R. M. Gray, "Image retrieval using color histograms generated by Gauss mixture vector quantization," *Computer Vision and Image Understanding*, vol. 94, no. 1–3, pp. 44–66, 2004.

[167] A. Jeż, "Faster fully compressed pattern matching by recompression," in *Proceedings, International Colloquium on Automata, Languages and Programming*, vol. 7391, (A. Czumaj, K. Mehlhorn, A. M. Pitts, and R. Wattenhofer, eds.), pp. 533–544, Springer, 2012.

[168] Y. Jia and E.-H. Yang, "Context-dependent multilevel pattern matching for lossless image compression," *IEEE Transactions on Information Theory*, vol. 49, no. 12, pp. 3169–3184, 2003.

[169] P. Johansen, "Combinatorial pattern recognition, the method and the program package," Technical Report 94/10, DIKU, Department of Computer Science, University of Copenhagen, Denmark, 1994.

[170] S. W. John Baras, "Hierarchical wavelet representations of ship radar returns," Technical Report, ISR Technical Report 93-100, 1993.

[171] M. Johnson, "Generalized descriptor compression for storage and matching," in *Proceedings, British Machine Vision Conference*, (F. Labrosse, R. Zwiggelaar, Y. Liu, and B. Tiddeman, eds.), pp. 1–11, British Machine Vision Association, 2010.

[172] U. Kandaswamy, S. A. C. Schuckers, and D. A. Adjeroh, “Comparison of texture analysis schemes under nonideal conditions,” *IEEE Transactions on Image Processing*, vol. 20, no. 8, pp. 2260–2275, 2011.
[173] J. Kärkkäinen, G. Navarro, and E. Ukkonen, “Approximate string matching Ziv-Lempel compressed text,” *Proceedings, Combinatorial Pattern Matching, LNCS 1848*, pp. 195–209, 2000.
[174] R. M. Karp and M. O. Rabin, “Efficient randomized pattern matching algorithms,” Technical Report TR-31-8, Aiken Computation Lab, Harvard University, 1981.
[175] M. Karpinski, W. Rytter, and A. Shinohara, “An efficient pattern-matching algorithm for strings with short descriptions,” *Nordic Journal of Computing*, vol. 4, no. 2, pp. 172–186, 1997.
[176] H. Khan and H. Fatmi, “Application of pattern recognition to text compression,” in *Proceedings, IEEE International Conference on Humans, Information and Technology*, vol. 2, pp. 1656–1659, October 1994.
[177] H. U. Khan and H. A. Fatmi, “A novel approach to data compression as a pattern recognition problem,” *Proceedings, IEEE Data Compression Conference*, 1993.
[178] V. R. Khapli and A. S. Bhalchandra, “Compressed domain image retrieval using thumbnails of images,” in *Proceedings, International Conference on Computational Intelligence*, (D. Al-Dabass, G. S. Tomar, R. Tripathi, and A. Abraham, eds.), pp. 392–396, 2009.
[179] V. R. Khapli and A. S. Bhalchandra, “Performance evaluation of image retrieval using VQ for compressed and uncompressed images,” in *Proceedings, International Conference on Emerging Trends in Engineering Technology*, pp. 885–888, 2009.
[180] T. Kida, T. Matsumoto, Y. Shibata, M. Takeda, A. Shinohara, and S. Arikawa, “Collage system: A unifying framework for compressed pattern matching,” *Theoretical Computer Science*, vol. 1, no. 298, pp. 253–272, 2003.
[181] T. Kida, M. Takeda, A. Shinohara, and S. Arikawa, “Shift-And approach to pattern matching in LZW compressed text,” *Proceedings, Combinatorial Pattern Matching, LNCS 1645*, pp. 1–13, 1999.
[182] T. Kida, M. Takeda, A. Shinohara, M. Miyazaki, and S. Arikawa, “Multiple pattern matching in LZW compressed text,” *Journal on Discrete Algorithms*, vol. 1, no. 1, pp. 133–158, 2000.
[183] J. C. Kieffer and E.-H. Yang, “Grammar-based codes: A new class of universal lossless source codes,” *IEEE Transactions on Information Theory*, vol. 46, no. 3, pp. 737–754, 2000.
[184] J. C. Kieffer, E.-H. Yang, G. J. Nelson, and P. C. Cosman, “Universal lossless compression via multilevel pattern matching,” *IEEE Transactions on Information Theory*, vol. 46, no. 4, pp. 1227–1245, 2000.
[185] S. T. Klein, “Improving static compression schemes by alphabet extension,” *Combinatorial Pattern Matching, LNCS 1848*, pp. 210–221, 2000.
[186] S. T. Klein and D. Shapira, “Pattern matching in Huffman encoded texts,” *Information Processing and Management*, vol. 41, no. 4, pp. 829–841, 2005.

[187] S. T. Klein and D. Shapira, "Compressed pattern matching in JPEG images," *International Journal of Foundations of Computer Science*, vol. 17, no. 6, pp. 1297–1306, 2006.
[188] S. T. Klein and D. Shapira, "Compressed matching in dictionaries," *Algorithms*, vol. 4, no. 1, pp. 61–74, 2011.
[189] S. T. Klein and D. Shapira, "On improving Tunstall codes," *Information Processing and Management*, vol. 47, no. 5, pp. 777–785, 2011.
[190] S. T. Klein and D. Shapira, "The string-to-dictionary matching problem," in *Proceedings, IEEE Data Compression Conference*, (J. A. Storer and M. W. Marcellin, eds.), pp. 143–152, 2011.
[191] S. T. Klein and Y. Wiseman, "Parallel Lempel Ziv coding," *Discrete Applied Mathematics*, vol. 146, no. 2, pp. 180–191, 2005.
[192] M. Klimesh, V. Stanton, and D. Watola, "Hardware implementation of a lossless image compression algorithm using a field programmable gate array," *TMO Progress Report*, vol. 42–144, February 15 2001.
[193] J. R. Knight and E. W. Myers, "Super-pattern matching," Technical Report TR-92-29, Department of Computer Science, University of Arizona, 1999.
[194] D. Knuth, J. Morris, and V. Pratt, "Fast pattern matching in strings," *SIAM Journal of Computing*, vol. 6, no. 2, pp. 323–350, June 1977.
[195] M. Kobayashi and K. Takeda, "Information retrieval on the web," *ACM Computing Surveys*, vol. 32, no. 2, pp. 144–173, 2000.
[196] S. Kosaraju, "Pattern matching in compressed texts," in *Proceedings, Foundations of Software Technology and Theoretical Computer Science*, (P. Thiagarajan, ed.), pp. 349–362, Springer-Verlag, 1995.
[197] J.-P. Kouma and H. Li, "Large-scale face image retrieval: A Wyner-Ziv coding approach," in *Proceedings, New Approaches to Characterization and Recognition of Faces*, (P. Corcoran, ed.), Intechopen.com, 2011.
[198] C. Kouzinopoulos and K. G. Margaritis, "Parallel implementation of exact two dimensional pattern matching algorithms using MPI and OpenMP," *Hellenic European Research on Mathematics and Informatics Science*, vol. 11, pp. 87–92, 2010.
[199] C. S. Kouzinopoulos and K. G. Margaritis, "String matching on a multicore GPU using CUDA," *Proceedings, Panhellenic Conference on Informatics*, vol. 14–18, no. 2009, 2009.
[200] S. Kreft and G. Navarro, "LZ77-Like compression with fast random access," in *Proceedings, IEEE Data Compression Conference*, (J. A. Storer and M. W. Marcellin, eds.), pp. 239–248, 2010.
[201] G. M. Landau and U. Vishkin, "Pattern matching in a digitized image," *Algorithmica*, vol. 12, no. 4/5, pp. 375–408, 1994.
[202] G. Langdon and J. Rissanen, "A simple general binary source code," *IEEE Transactions on Information Theory*, vol. IT-28, pp. 800–803, 1982.
[203] S. Lawrence and C. Giles, "Context and page analysis for improved web search," *IEEE Internet Computin*, vol. 2, no. 4, pp. 38–46, 1998.
[204] S. Lazebnik, C. Schmid, and J. Ponce, "A sparse texture representation using local affine regions," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 8, pp. 1265–1278, 2005.

[205] D.-S. Lee and J. J. Hull, "Duplicate detection in symbolically compressed documents," in *Proceedings, International Conference on Document Analysis and Recognition*, pp. 305–308, 1999.
[206] J. S. Lee, D. K. Kim, K. Park, and Y. Cho, "Efficient algorithms for approximate string matching with swaps," *Proceedings, Combinatorial Pattern Matching,* LNCS vol. 1264, pp. 28–39, 1997.
[207] D. LeGall, "MPEG: A video compression standard for multimedia applications," *Communications of the ACM*, vol. 34, no. 4, pp. 47–58, April 1991.
[208] A. Lempel and J. Ziv, "Compression of two-dimensional data," *IEEE Transactions on Information Theory*, vol. 32, no. 1, pp. 2–8, 1986.
[209] M. Lew and T. Huang, "Image compression and matching," in *Proceedings, IEEE International Conference on Image Processing*, vol. 2, pp. 720–724, November 1994.
[210] M. S. Lew, N. Sebe, C. Djeraba, and R. Jain, "Content-based multimedia information retrieval: State of the art and challenges," *TOMCCAP*, vol. 2, no. 1, pp. 1–19, 2006.
[211] C.-S. Li, J. Turek, and E. Feig, "Progressive template matching for content-based retrieval in Earth Observing Satellite image database," *SPIE Proceedings 2606*, pp. 134–144, 1995.
[212] S. Li and W. Li, "Shape-adaptive discrete wavelet transforms for arbitrarily shaped visual object coding," *IEEE Transactions on Circuits Systems for Video Technology*, vol. 10, no. 5, pp. 725–743, 2000.
[213] X. Li and M. T. Orchard, "Edge-directed prediction for lossless compression of natural images," *IEEE Transactions on Image Processing*, vol. 10, no. 6, pp. 813–817, 2001.
[214] Y. Lifshits, "Solving classical string problems an compressed texts," in *Proceedings, Combinatorial and Algorithmic Foundations of Pattern and Association Discovery*, vol. 06201 of *Dagsthul Seminar Proceedings,* Internationales Begegnungs- und Forschungszentrum fuer Informatik (IBFI), (R. Ahlswede, A. Apostolico, and V. I. Levenshtein, eds.), Germany: Schloss Dagstuhl, 2006.
[215] Y. Lifshits, "Processing compressed texts: A tractability border," in *Proceedings, Combinatorial Pattern Matching*, LNCS vol. 4580, (B. Ma and K. Zhang, eds.), pp. 228–240, Springer, 2007.
[216] A. D. Lillo, A. H. Daptardar, K. Thomas, J. A. Storer, and G. Motta, "Applications of compression to content based image retrieval and object recognition," in *Proceedings, Conference Computational Physics*, pp. 179–189, 2011.
[217] A. D. Lillo, A. H. Daptardar, K. Thomas, J. A. Storer, and G. Motta, "Compression-based tools for navigation with an image database," *Algorithms*, vol. 5, no. 1, pp. 1–17, 2012.
[218] A. D. Lillo, G. Motta, and J. A. Storer, "Shape recognition using vector quantization," in *Proceedings, IEEE Data Compression Conference*, (J. A. Storer and M. W. Marcellin, eds.), pp. 484–493, 2010.
[219] C.-Y. Lin and S.-F. Chang, "A robust image authentication method distinguishing JPEG compression from malicious manipulation," *IEEE Transactions on Circuits and Systtems for Video Technology*, vol. 11, no. 2, pp. 153–168, 2001.

[220] J. Lin and D. A. Adjeroh, "All-against-all circular pattern matching," *The Computer Journal*, vol. 55, no. 7, pp. 897–906, 2012.

[221] K.-J. Lin and C.-W. Wu, "A low power CAM design for LZ data compression," *IEEE Transactions on Computers*, vol. 49, pp. 1139–1145, October 2000.

[222] T. W. Lin, "Compressed quadtree representations for storing similar images," *Image and Vision Computing*, vol. 15, no. 11, pp. 833–843, 1997.

[223] O. Lipsky and E. Porat, "Approximate pattern matching with the L_1, L_2 and L_∞ Metrics," *Algorithmica*, vol. 60, no. 2, pp. 335–348, 2011.

[224] L.-Y. Liu, J.-F. Wang, R.-J. Wang, and J.-Y. Lee, "Design and hardware architectures for dynamic Huffman coding," *IEE Proceedings on Computers and Digital Techniques*, vol. 142, no. 6, pp. 411–418, November 1995.

[225] E. Loupias, N. Sebe, S. Bres, and J.-M. Jolion, "Wavelet-based salient points for image retrieval," in *Proceedings, IEEE International Conference on Image Processing*, 2000.

[226] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.

[227] C.-S. Lu and H.-Y. M. Liao, "Structural digital signature for image authentication: An incidental distortion resistant scheme," *IEEE Transactions on Multimedia*, vol. 5, no. 2, pp. 161–173, 2003.

[228] S. Lu, L. Li, and C. L. Tan, "Document image retrieval through word shape coding," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 11, pp. 1913–1918, 2008.

[229] Y. Lu and C. L. Tan, "Document retrieval from compressed images," *Pattern Recognition*, vol. 36, no. 4, pp. 987–996, 2003.

[230] T. Luczak and W. Szpankowski, "A suboptimal lossy data compression based on approximate pattern matching," *IEEE Transactions on Information Theory*, vol. 43, pp. 1439–1451, 1997.

[231] C.-Y. Maa, "Identifying the existence of bar codes in compressed images," in *Proceedings, IEEE Data Compression Conference*, (J. Storer and M. Cohn, eds.), p. 457, Los Alamitos, California, 1993.

[232] V. Mäkinen, E. Ukkonen, and G. Navarro, "Approximate matching of run-length compressed strings," *Algorithmica*, vol. 35, pp. 347–369, 2003.

[233] U. Manber, "A text compression scheme that allows fast searching directly in the compressed file," *ACM Transactions on Information Systems*, vol. 15, no. 2, pp. 124–136, April 1997.

[234] M. K. Mandal, F. Idris, and S. Panchanathan, "A critical evaluation of image and video indexing techniques in the compressed domain," *Journal of Image and Vision Computing*, vol. 17, no. 7, pp. 513–529, 1999.

[235] M. K. Mandal, S. Panchanathan, and T. Aboulnas, "Fast wavelet histogram techniques for image indexing," *Computer Vision and Image Understandin*, vol. 75, no. 1, pp. 99–110, 1999.

[236] B. S. Manjunath and W. Y. Ma, "Texture features for browsing and retrieval of image data," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 18, no. 8, pp. 837–841, 1996.

[237] W. Matsubara, S. Inenaga, A. Ishino, A. Shinohara, T. Nakamura, and K. Hashimoto, "Efficient algorithms to compute compressed longest common substrings and compressed palindromes," *Theoretical Computer Science*, vol. 410, no. 8–10, pp. 900–913, 2009.

[238] N. D. Memon, D. L. Neuhoff, and S. M. Shende, "An analysis of some common scanning techniques for lossless image coding," *IEEE Transactions on Image Processing*, vol. 9, no. 11, pp. 1837–1848, 2000.

[239] N. D. Memon and X. Wu, "Recent developments in context-based predictive techniques for lossless image compression," *The Computer Journal*, vol. 40, no. 2/3, pp. 127–136, 1997.

[240] B. Meyer and P. E. Tischer, "Extending TMW for near lossless compression of greyscale images," in *Proceedings, IEEE Data Compression Conference*, pp. 458–470, 1998.

[241] K. Mikolajczyk and C. Schmid, "A performance evaluation of local descriptors," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 10, pp. 1615–1630, 2005.

[242] V. Miller and M. Wegman, "Variations on a theme by Ziv and Lempel," in *Proceedings, Combinatorial Algorithms on Words*, vol. 12, (A. Apostolico and Z. Galil, eds.), Berlin: Springer-Verlag, 1985.

[243] S. Mitarai, M. Hirao, T. Matsumoto, A. Shinohara, M. Takeda, and S. Arikawa, "Compressed pattern matching for SEQUITUR," in *Proceedings, IEEE Data Compression Conference*, p. 469, 2001.

[244] M. Miyazaki, A. Shinohara, and M. Takeda, "An improved pattern matching algorithm for strings in terms of straight-line programs," in *Proceedings, Combinatorial Pattern Matching*, LNCS vol. 1264, (A. Apostolico and J. Hein, eds.), pp. 1–11, Springer, 1997.

[245] A. Moffat, "Two-level context based compression of binary images," in *Proceedings, IEEE Data Compression Conference*, (J. A. Storer and J. H. Reif, eds.), pp. 382–391, 1991.

[246] A. Moffat, R. M. Neal, and I. H. Witten, "Arithmetic coding revisited," *ACM Transactions on Information System*, vol. 16, no. 3, pp. 256–294, 1998.

[247] V. Monga and B. L. Evans, "Perceptual image hashing via feature points: Performance evaluation and tradeoffs," *IEEE Transactions on Image Processing*, vol. 15, no. 11, pp. 3452–3465, 2006.

[248] M. Mongeau and D. Sankoff, "Comparison of musical sequencs," *Computers and the Humanities*, vol. 24, pp. 161–175, 1990.

[249] J. Moreno and X. Otazu, "Image coder based on Hilbert scanning of embedded quadtrees," in *Proceedings, IEEE Data Compression Conference*, (J. A. Storer and M. W. Marcellin, eds.), p. 470, 2011.

[250] E. S. Moura, G. Navarro, and R. Baeza-Yates, "Fast and flexible word searching on compressed text," *ACM Transactions on Information Systems*, vol. 18, no. 2, pp. 113–139, 2000.

[251] A. Mukherjee and T. Acharya, "Compressed pattern-matching," in *Proceedings, IEEE Data Compression Conference*, p. 468, 1994.

[252] A. Mukherjee and T. Acharya, "VLSI algorithms for compressed pattern search using tree based codes," in *Proceedings, International Conference on Application Specific Array Processors*, pp. 133–136, 1995.

[253] A. Mukherjee, N. Ranganathan, and M. Bassiouni, "Efficient VLSI designs for data transformation of tree-based codes," *IEEE Transactions on Circuits and Systems*, vol. 38, no. 3, pp. 306–314, 1991.

[254] J. Mukhopadhyay, *Image and Video Processing in the Compressed Domain*. US: Taylor & Francis, 2011.

[255] A. Murugappan, B. Ramachandran, and P. Dhavachelvan, "A survey of keyword spotting techniques for printed document images," *Artificial Intelligence Review*, vol. 35, no. 2, pp. 119–136, 2011.

[256] E. W. Myers, "A sublinear algorithm for approximate keyword searching," *Algorithmica*, vol. 12, pp. 345–374, 1994.

[257] Ö. U. Nalbantoglu, D. J. Russell, and K. Sayood, "Data compression concepts and algorithms and their applications to bioinformatics," *Entropy*, vol. 12, no. 1, pp. 34–52, 2010.

[258] G. Navarro, "A guided tour to approximate string matching," *ACM Computing Surveys*, vol. 33, no. 1, pp. 31–88, 2001.

[259] G. Navarro and M. Raffinot, "A general practical approach to pattern matching over Ziv-Lempel compressed text," *Proceedings, Combinatorial Pattern Matching, LNCS 1645*, pp. 14–36, 1999.

[260] G. Navarro and M. Raffinot, "Fast and flexible string matching by combining bit-parallelism and suffix automata," *ACM Journal of Experimental Algorithms*, vol. 5, no. 4, 2000.

[261] G. Navarro and M. Raffinot, "Practical and flexible pattern matching over Ziv-Lempel compressed text," *Journal on Discrete Algorithms*, vol. 2, no. 3, pp. 347–371, 2004.

[262] G. Navarro and J. Tarhio, "Boyer-Moore string matching over Ziv-Lempel compressed text," *Proceedings, Combinatorial Pattern Matching, LNCS 1848*, pp. 166–180, 2000.

[263] G. Navarro and J. Tarhio, "LZgrep: A Boyer-Moore string matching tool for Ziv-Lempel compressed text," *Software, Practice and Experience*, vol. 35, no. 12, pp. 1107–1130, 2005.

[264] M. Nelson, *The Data Compression Book*. San Mateo, CA (USA): M&T Books, A Division of M&T Publishing, Inc., 1991.

[265] C. G. Nevill-Manning and I. H. Witten, "Compression and explanation using hierarchical grammars," *The Computer Journal*, vol. 40, no. 2/3, pp. 103–116, 1997.

[266] C. G. Nevill-Manning and I. H. Witten, "Identifying hierarchical structure in sequences: A linear-time algorithm," *Journal on Artificial Intelligence Research (JAIR)*, vol. 7, pp. 67–82, 1997.

[267] H. Ney and S. Ortmanns, "Progress in dynamic programming search for LVCSR," *Proceedings of the IEEE*, vol. 88, no. 8, pp. 1224–1240, 2000.

[268] C. W. Ngo, T. C. Pong, and R. T. Chin, "Exploiting image indexing techniques in DCT domain," *Proceedings, MINAR98: Multimedia Information Analysis and Retrieval, LNCS 1464*, pp. 195–206, August 1998.

[269] K. Oehler and R. Gray, "Combining image classification and image compression using vector quantization," in *Proceedings, IEEE Data Compression Conference*, pp. 2–11, 1993.

[270] H. H. Otu and K. Sayood, "A new sequence distance measure for phylogenetic tree construction," *Bioinformatics*, vol. 19, no. 16, pp. 2122–2130, 2003.

[271] O. Owolabi and D. R. McGregor, "Fast approximate string matching," *Software — Practice and Experience*, vol. 18, pp. 387–393, 1988.

[272] R. Pajarola and P. Widmayer, "Pattern matching in compressed raster images," in *Proceedings, South American Workshop on String Processing WSP 1996, International Informatics Series 4*, pp. 228–242, Carleton University Press, 1996.

[273] R. Pajarola and P. Widmayer, "An image compression method for spatial search," *IEEE Transactions on Image Processing*, vol. 9, no. 3, pp. 357–365, 2000.

[274] V. Pankratius, A. Jannesari, and W. F. Tichy, "Parallelizing Bzip2: A case study in multicore software engineering," *IEEE Software*, vol. 26, no. 6, pp. 70–77, 2009.

[275] W. A. Pearlman and A. Said, "Set partition coding: Part I of set partition coding and image wavelet coding systems," *Foundations and Trends in Signal Processing*, vol. 2, no. 2, pp. 95–180, 2008.

[276] W. A. Pearlman and A. Said, "Set partition coding: Part II of set partition coding and image wavelet coding systems," *Foundations and Trends in Signal Processing*, vol. 2, no. 3, pp. 181–246, 2008.

[277] W. B. Pennebaker and J. L. Mitchell, *JPEG: Still Image Data Compression Standard*. NY: Van Nostrand Reinhold, 1993.

[278] A. Pentland, R. W. Picard, and S. Scaroff, "Photobook: Content-based manipulation of image databases," *International Journal of Computer Vision*, vol. 18, no. 3, pp. 233–254, 1996.

[279] E. C. Posner and E. R. Rodemich, "Epsilon entropy and data compression," *The Annals of Mathematical Statistics*, vol. 42, no. 6, pp. 2079–2125, 1971.

[280] P. Poursistani, H. Nezamabadi-pour, R. A. Moghadam, and M. Saeed, "Image indexing and retrieval in JPEG compressed domain based on vector quantization," *Mathematical and Computer Modelling*, 2011.

[281] W. K. Pratt, *Digital Image Processing*. New York: John Wiley, 1991.

[282] S. J. Puglisi, W. F. Smyth, and A. Turpin, "A taxonomy of suffix array construction algorithms," *ACM Computing Surveys*, vol. 39, no. 2, 2007.

[283] S.-W. Ra and J.-K. Kim, "A fast mean-distance-ordered partial codebook search algorithm for image vector quantization," *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, vol. 40, no. 9, pp. 576 –579, September 1993.

[284] N. Ranganathan and S. Henriques, "High speed VLSI design for Lempel-Ziv-based data compression," *IEEE Transcations on Circuits and Systems*, vol. 40, pp. 96–106, February 1993.

[285] R. Reeves, K. Kubik, and W. Osberger, "Texture characterization of compressed aerial images using DCT coefficients," *Proceedings of SPIE: Storage and Retrieval for Image and Video Databases*, vol. 3022, pp. 398–407, 1997.

[286] I. Rigoutsos and A. Califano, "Searching in parallel for similar strings," *IEEE Computational Science and Engineering*, vol. 60–67, no. summer issue, 1994.

[287] F. Rizzo, J. A. Storer, and B. Carpentieri, "LZ-based image compression," *Information Science*, vol. 135, no. 1–2, pp. 107–122, 2001.

[288] R. Y. Rubinstein and D. P. Kroese, *The Cross-Entropy Method: A Unified Approach to Combinatorial Optimization, Monte-Carlo Simulation, and Machine Learning*. New York, USA: Springer-Verlag, 2004.

[289] L. M. S. Russo, G. Navarro, and A. L. Oliveira, "Parallel and distributed compressed indexes," in *Proceedings, Combinatorial Pattern Matching*, LNCS vol. 6129, (A. Amir and L. Parida, eds.), pp. 348–360, Springer, 2010.

[290] W. Rytter, "Compressed and fully compressed pattern matching in one and two dimensions," *Proceedings of the IEEE*, vol. 88, no. 11, pp. 1769–1778, 2000.

[291] W. Rytter, "Application of Lempel-Ziv factorization to the approximation of grammar-based compression," *Theoretical Computer Science*, vol. 302, no. 1-3, pp. 211–222, 2003.

[292] A. Said and W. A. Pearlman, "A new fast and efficient image codec based on set partitioning in hierarchical trees," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 6, no. 3, pp. 243–250, 1996.

[293] H. Sakoe and S. Chiba, "Dynamic programming algorithm optimization for spoken word recognition," *IEEE Transactions on Acoustic, Speech and Signial Processing*, vol. 26, no. 2, pp. 43–49, 1978.

[294] D. Salomon, *Data Compression: The Complete Reference*. New York: Springer, 3rd Edition, 2004.

[295] D. Salomon, "Special issue on lossless data compression," *Algorithms*, vol. 3, no. 2, 2010.

[296] H. Samet, "The quadtree and related hierarchical data structures," *ACM Computing Survey*, vol. 16, no. 2, pp. 187–260, 1984.

[297] S. A. Savari and R. G. Gallager, "Generalized Tunstall codes for sources with memory," *IEEE Transactions on Information Theory*, vol. 43, no. 2, pp. 658–668, 1997.

[298] K. Sayood, *Introduction to Data Compression*. San Francisco, CA, USA: Morgan Kaufmann Publishers, 4th Edition, 2012.

[299] G. Schaefer, "Compressed domain image retrieval by comparing vector quantisation codebooks," in *Proceedings of SPIE*, pp. 959–966, 2002.

[300] M. Schmalz, "General theory for the processing of compressed and encrypted imagery, with taxonomic analysis," *Proceedings, SPIE: Hybrid Image and Signal Processing III*, vol. 1702, pp. 250–263, 1992.

[301] M. Schmalz, "An introduction to the recognition of patterns in compressed data. 2. Optical processing of data transformed by block-, transform-, and runlength-encoding, as well as vector-quantization," *Proceedings of the SPIE*, vol. 2490, pp. 334–349, 1995.

[302] M. Schmalz, "An introduction to the recognition of patterns in compressed data. 1. Image-template operations over block-, transform-, runlength-encoded, and vector-quantized data," *Proceedings of the SPIE*, vol. 2484, pp. 256–269, 1995.

[303] C. E. Shannon, "A mathematical theory of communication," *Bell System Technical Journal*, vol. 27, pp. 379–423, 623–656, 1948.

[304] C. E. Shannon, "Coding theorems for a discrete source with a fidelity criterion," *IRE National Convention Record, Part 4*, pp. 142–163, 1959.

[305] J. M. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," *IEEE Transactions on Signal Processing*, vol. 41, no. 12, pp. 3445–3462, 1993.

[306] N. Sharman, T. Bell, and I. Witten, "Compression of pyramid coded images for progressive transmission," in *Proceedings of New Zealand Image Processing Workshop*, pp. 171–176, University of Canterbury, Christchurch, New Zealand, 1992.

[307] Y. Shibata, T. Kida, S. Fukamachi, T. Takeda, A. Shinohara, S. Shinohara, and S. Arikawa, "Byte-pair encoding: A text compression scheme that accelerates pattern matching," Technical Report, Department of Informatics, Kyushu University, Japan, 1999.

[308] Y. Shibata, T. Matsumoto, M. Takeda, A. Shinohara, and S. Arikawa, "A Boyer-More type algorithm for compressed pattern matching," *Proceedings, Combinatorial Pattern Matching,* LNCS vol. 1848, pp. 181–194–13, 2000.

[309] Y. Shibata, M. Takeda, A. Shinohara, and S. Arikawa, "Pattern matching in text compressed by using antidictionaries," *Proceedings, Combinatorial Pattern Matching, LNCS 1645*, pp. 37–49, 1999.

[310] M. Shneier and M. Abdel-Mottaleb, "Exploiting the JPEG compression scheme for image retrieval," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 18, no. 8, pp. 849–853, 1996.

[311] T. Sikora, "The MPEG-4 video standard verification model," *IEEE Transactions on Circuit and Systems for Video Technology*, vol. 7, no. 1, 1997.

[312] T. Sikora, S. Bauer, and B. Makai, "Efficiency of shape–adaptive 2D transforms for coding of arbitrarily shaped image segments," *IEEE Transactions on Circuits and Systems for Video Technolog*, vol. 5, pp. 254–258, 1995.

[313] E. Silberg and O. Etzion, "The metacrawler architecture for resource aggregation on the web," *IEEE Expert*, vol. 12, no. 1, pp. 11–14, 1997.

[314] D. Sleator and R. Tarjan, "Amortized efficiency of list update and paging rules," *Communications of the ACM*, vol. 28, pp. 202–208, 1985.

[315] A. W. M. Smeulders, M. Worring, S. Santini, A. Gupter, and R. Jain, "Content-based image retrieval at the end of the early years," *IEEE Transactions on Pattern Analysis and Machine Intellingence*, vol. 22, no. 12, pp. 1349–1380, 2000.

[316] J. R. Smith and S. F. Chang, "Transform features for texture classification and discrimination in large image databases," *Proceedings, IEEE International Conference on Image Processing*, vol. 3, pp. 407–411, 1994.

[317] W. Smyth, *Computing Patterns in Strings.* Addison-Wesley, 2003.

[318] C. Snoek and M. Worring, "Multimodal video indexing: A review of the state-of-the-art," *Multimedia Tools and Applications*, vol. 25, no. 1, pp. 5–35, 2005.

[319] C. G. M. Snoek and M. Worring, "Concept-based video retrieval," *Foundations and Trends in Information Retrieval*, vol. 2, no. 4, pp. 215–322, 2009.

[320] J. Song and B. L. Yeo, "Fast extraction of spatially reduced image sequences from MPEG-2 compressed video," *IEEE Transactions on Circuits and Systems for Video Technolog*, vol. 9, no. 7, pp. 1100–1114, 1999.

[321] Y. Steinberg and M. Gutman, "An algorithm for source coding subject to a fidelity criterion, based on string pattern matching," *IEEE Transactions on Information Theory*, vol. 39, pp. 877–886, 1993.

[322] J. Storer, *Data Compression: Methods and Theory.* Rockville, Maryland: Computer Science Press, 1988.

[323] J. Storer and M. Cohn, eds., *Proceedings of IEEE Data Compression Conference.* Los Alamitos, California, Snowbird, Utah, March 1993.

[324] J. Storer and T. Syzmanski, "Data compression via textual substitution," *Journal of the ACM*, vol. 29, pp. 928–951, 1982.

[325] J. A. Storer, "Special issue on lossless data compression," *Proceedings of the IEEE*, vol. 88, no. 11, pp. 1685–1809, 2000.

[326] J. A. Storer and H. Helfgott, "Lossless image compression by block matching," *The Computer Journal*, vol. 40, no. 2/3, pp. 137–145, 1997.

[327] J. A. Storer and M. W. Marcellin, eds., *2010 IEEE Data Compression Conference (DCC 2010).* 24-26 March 2010, Snowbird, UT, USA, 2010.

[328] J. A. Storer and M. W. Marcellin, eds., *2011 IEEE Data Compression Conference (DCC 2011).* 29-31 March 2011, Snowbird, UT, USA, 2011.

[329] M. D. Swanson, S. Hosur, and A. H. Tewfik, "Image coding for content-based retrieval," *Proceedings of SPIE: Visual Communications and Image Processing*, vol. 2727, pp. 4–15, 1996.

[330] W. Szpankowski, "Asympotic properties of data compression and suffix trees," *IEEE Transactions on Information Theory*, vol. 39, pp. 1647–1659, 1993.

[331] M. Tagliasacchi, G. Valenzise, and S. Tubaro, "Hash-based identification of sparse image tampering," *IEEE Transactions on Image Processing*, vol. 18, no. 11, pp. 2491–2504, 2009.

[332] T. Takaoka, "Approximate pattern matching with grey scale values," in *Proceedings, CATS 96 (Computing: the Australasian Theory Symposium)*, pp. 196–203, 1996.

[333] Y. Tamakoshi, T. I, S. Inenaga, H. Bannai, and M. Takeda, "From run length encoding to LZ78 and back again," in *Proceedings, IEEE Data Compression Conference*, pp. 143–152, 2013.

[334] T. Tanaka, T. I, S. Inenaga, H. Bannai, and M. Takeda, "Computing convolution on grammar-compressed text," *Proceedings, IEEE Data Compression Conference*, pp. 451–460, 2013.

[335] J. Tang, W. Zhang, and C. Li, "An approach to compressed image retrieval based on JPEG2000 framework," in *Advanced Data Mining and Applications*, vol. 3584, (X. Li, S. Wang, and Z. Y. Dong, eds.), pp. 391–399, Springer, 2005.

[336] J. W. H. Tangelder and R. C. Veltkamp, "A survey of content based 3D shape retrieval methods," *Multimedia Tools and Applications*, vol. 39, no. 3, pp. 441–471, 2008.

[337] T. Tao and A. Mukherjee, "Compressed pattern matching for predictive lossless image encoding," in *Proceedings, International Conference on Distributed Multimedia Systems*, p. 120, September 24-26, Miami, FL, 2003.

[338] T. Tao and A. Mukherjee, "Pattern matching in LZW compressed file," *IEEE Transactions on Computers*, vol. 54, no. 8, pp. 929–938, 2005.

[339] T. Tao, A. Mukherjee, and R. Satya, "A search-aware JPEG-LS variation for compressed image retrieval," in *Proceedings, International Symposium on Intelligent Multimedia. Video and Speech Processing*, pp. 20–222, October 20-22, 2004 Hong Kong, 2004.

[340] C. M. Taskiran, J.-Y. Chen, A. Albiol, L. Torres, C. A. Bouman, and E. J. Delp, "ViBE: A compressed video database structured for active browsing and search," *IEEE Transactions on Multimedia*, vol. 6, no. 1, pp. 103–118, 2004.

[341] D. Taubman, "High performance scalable image compression with EBCOT," *IEEE Transactions on Image Processing*, vol. 9, no. 7, pp. 1158–1170, July 2000.

[342] D. S. Taubman and M. W. Marcellin, *JPEG2000: Image Compression Fundamentals, Standards, and Practice*. Boston: Kluwer Academic Publishers, 2002.

[343] W. Thies, S. Hall, and S. Amarasinghe, "Manipulating lossless video in the compressed domain," in *Proceedings, ACM International Conference on Multimedia*, pp. 331–340, New York, NY, USA, 2009.

[344] Q. Tian, N. Sebe, M. S. Lew, E. Loupias, and T. S. Huang, "Image retrieval using wavelet-based salient points," *Journal on Electronic Imaging*, vol. 10, no. 4, pp. 835–849, 2001.

[345] TREC, "Official webpage for TREC — Text REtrieval Conference series," http://trec.nist.gov, 2000.

[346] TRECVID, "Official webpage for NIST TRECVID — TREC Video Retrieval Evaluation," http://trecvid.nist.gov, 2012.

[347] W. S. Tsai and S. S. Yu, "Attributed string matching with merging for shape recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 7, no. 4, pp. 453–463, 1985.

[348] P. Tseng, Y. Chang, Y. Huang, H. Fang, C. Huang, and L. Chen, "Advances in hardware architectures for image and video coding: A survey," *Proceedings of the IEEE*, vol. 93, no. 1, pp. 184–197, January 2005.

[349] N. Tuck, T. Sherwood, B. Calder, and G. Varghese, "Deterministic memory-efficient string matching algorithms for intrusion detection," in *Proceedings, IEEE International Conference on Computer Communications (IEEE INFOCOM)*, pp. 333–340, 2004.

[350] E. Tuncel, P. Koulgi, and K. Rose, "Rate-distortion approach to databases: Storage and content-based retrieval," *IEEE Transactions on Information Theory*, vol. 50, no. 6, pp. 953–967, 2004.

[351] B. Tunstall, "Synthesis of noiseless compression codes," Technical Report, PhD Dissertation, Georgia Institute of Technology, Atlanta, GA, 1967.

[352] A. Tversky, "Features of similarity," *Psychological Review*, vol. 84, no. 4, pp. 327–352, 1977.

[353] T. Uemura and H. Arimura, "Sparse and truncated suffix trees on variable-length codes," in *Proceedings, Combinatorial Pattern Matching*, LNCS vol. 6661, (R. Giancarlo and G. Manzini, eds.), pp. 246–260, Springer, 2011.

[354] T. Uemura, S. Yoshida, T. Kida, T. Asai, and S. Okamoto, "Training parse trees for efficient VF coding," in *Proceedings, String Processing and Information Retrieval*, LNCS vol. 6393, (E. Chávez and S. Lonardi, eds.), pp. 179–184, Springer, 2010.

[355] E. Ukkonen, "Finding approximate patterns in strings," *Journal of Algorithms*, vol. 6, pp. 132–137, 1985.

[356] J. Vaisey and A. Gersho, "Image compression with variable block size segmentation," *IEEE Transactions on Signal Processing*, vol. 40, no. 8, pp. 2040–2060, August 1992.

[357] N. Vasconcelos, "Minimum probability of error image retrieval," *IEEE Transactions on Signal Processing*, vol. 52, no. 8, pp. 2322–2336, 2004.

[358] N. Vasconcelos, "On the efficient evaluation of probabilistic similarity functions for image retrieval," *IEEE Transactions on Information Theory*, vol. 50, no. 7, pp. 1482–1496, 2004.

[359] N. Vasconcelos and A. Lippman, "Library-based coding: A representation for efficient video compression and retrieval," in *Proceedings, IEEE Data Compression Conference*, (J. A. Storer and M. Cohn, eds.), pp. 121–130, 1997.

[360] N. Vasconcelos and A. Lippman, "A multiresolution manifold distance for invariant image similarity," *IEEE Transactions on Multimedia*, vol. 7, no. 1, pp. 127–142, 2005.

[361] A. Vellaikal, C. C. J. Kuo, and S. Dao, "Content-based retrieval of remote-sensed images using vector quantization," *Proceedings of SPIE*, vol. 2488, pp. 178–189, 1995.

[362] R. C. Veltkamp, "Shape algorithmics," *Algorithmica*, vol. 38, no. 1, pp. 1–4, 2003.

[363] R. Venkatesan, S.-M. Koon, M. H. Jakubowski, and P. Moulin, "Robust Image Hashing," in *Proceedings, International Conference on Image Processing*, 2000.

[364] S. Verma and A. Pandit, "Quad tree based Image compression," in *Proceedings, TENCON 2008 — 2008 IEEE Region 10 Conference*, pp. 1–4, November 2008.

[365] G. Wallace, "The JPEG still picture compression standard," *Communications of the ACM*, vol. 34, no. 4, pp. 31–44, April 1991.

[366] O. Watanabe, T. Fukuhara, and H. Kiya, "Fast identification of JPEG 2000 images for digital cinema profiles," in *Proceedings, International Conference on Acoustics Speech and Signal Processing*, pp. 881–884, 2011.

[367] O. Watanabe, A. Kawana, and H. Kiya, "An identification method for JPEG2000 images using the signs of DWT coefficients," Technical Report of IEICE 106, IE2006-217, pp. 177–181, 2007.

[368] M. S. Waterman, *Mathematical Methods for DNA Sequences.* Boca Raton, FL: CRC Press, 1989.

[369] G. Wei, D. Li, and I. K. Sethi, "Web-WISE: Compressed image retrieval," *Proceedings, MINAR'98: Multimedia Information Analysis and Retrieval,* LNCS vol. 1464, pp. 33–46, August 1998.

[370] M. J. Weinberger, G. Seroussi, and G. Sapiro, "The LOCO-I lossless image compression algorithm: Principles and standardization into JPEG-LS," *IEEE Transactions on Image Processing*, vol. 9, no. 8, pp. 1309–1324, 2000.

[371] T. Welch, “A technique for high performance data compression,” *IEEE Computer*, vol. 17, pp. 8–20, June 1984.

[372] M. V. Wickerhauser, “Two fast approximate wavelet algorithms for image processing, classification, and recognition,” *Optical Engineering*, vol. 33, no. 7, pp. 2225–2235, (Special issue on Adapted Wavelet Analysis), July 1994.

[373] T. Wiegand, G. J. Sullivan, G. Bjntegaard, and A. Luthra, “Overview of the H.264/AVC video coding standard,” *IEEE Transactions on Circuits Systems and Video Technology*, vol. 13, no. 7, pp. 560–576, 2003.

[374] I. Witten and J. Cleary, “Picture coding and transmission using adaptive modeling of quad trees,” in *Proceedings, International Electrical, Electronics Conference*, pp. 222–225, Toronto, 1983.

[375] I. Witten, R. Neal, and J. Cleary, “Arithmetic coding for data compression,” *Communications of the ACM*, vol. 30, no. 6, pp. 520–541, June 1987.

[376] I. H. Witten, A. Moffat, and T. C. Bell, *Managing Gigabytes: Compressing and Indexing Documents and Images*. Morgan Kaufman, 2nd Edition, 1999.

[377] J. W. Woods, ed., *Subband Image Coding*. Amsterdam: Kluwer Academic Publishers, 1985.

[378] S. Wu and U. Manber, “Fast text searching allowing errors,” *Communications of the ACM*, vol. 35, no. 10, pp. 83–91, October 1992.

[379] M. Yadav, A. Venkatachaliah, and P. D. Franzon, “Hardware architecture of a parallel pattern matching engine,” in *Proceedings, International and Symposium on Circuits and Systems*, pp. 1369–1372, IEEE, 2007.

[380] E.-H. Yang, A. Kaltchenko, and J. C. Kieffer, “Universal lossless data compression with side information by using a conditional MPM grammar transform,” *IEEE Transactions on Information Theory*, vol. 47, no. 6, pp. 2130–2150, 2001.

[381] E. H. Yang and J. Kieffer, “On the performance of data compression algorithms based on string pattern matching,” *IEEE Transactions on Information Theory*, vol. 41, 1995.

[382] L. Yang, Q. Chen, J. Tian, and D. Wu, “Content based image hashing using companding and Gray code,” *Security and Communication Networks*, vol. 4, no. 12, pp. 1378–1386, 2011.

[383] S.-H. Yang and C.-F. Chen, “Robust image hashing based on SPIHT,” in *Proceedings, International Conference on Information Technology: Research and Education*, pp. 110–114, 2005.

[384] B. L. Yeo and B. Liu, “Rapid scene analysis in compressed video,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 5, no. 6, p. 533, 1996.

[385] X. Yin, I. Düntsch, and G. Gediga, “Quadtree representation and compression of spatial data,” *Transactions on Rough Sets*, vol. 13, pp. 207–239, 2011.

[386] S. Yoshida and T. Kida, “On performance of compressed pattern matching on VF codes,” in *Proceedings, IEEE Data Compression Conference*, (J. A. Storer and M. W. Marcellin, eds.), p. 486, 2011.

[387] S. Yoshida, T. Uemura, T. Kida, T. Asai, and S. Okamoto, “Improving parse trees for efficient variable-to-fixed length codes,” *Journal of Information Processing*, vol. 20, no. 1, pp. 238–249, 2012.

[388] T. Yoshida, T. Endoh, N. Kawamura, and H. Kato, “Image reduction system,” US Patent 5,159,468, October 1992.

[389] F. Zargari, A. Mosleh, and M. Ghanbari, "Compressed domain JPEG2000 image indexing method employing full packet header information," in *Proceedings, International Workshop on Content-Based Multimedia Indexing, 2008 (CBMI 2008)*, pp. 410–416, June 2008.

[390] K. Zeger and A. Gersho, "Pseudo-Gray coding," *IEEE Transactions on Communications*, vol. 38, no. 12, 1990.

[391] A. Zhang, B. Cheng, and R. Acharya, "Texture-based image retrieval using fractal codes," Technical Report 95-19, Department of Computer Science, SUNY Buffalo, Tue, 19 Sep 1995, August 16 1995.

[392] N. Zhang, A. Mukherjee, D. A. Adjeroh, and T. Bell, "Approximate pattern matching using the Burrows-Wheeler transform," in *Proceedings, IEEE Data Compression Conference*, p. 458, 2003.

[393] W. Zhang, Z. Nie, and Z. Zeng, "Image retrieval based on salient points from DCT domain," in *Proceedings, Mexican International Conference on Artificial Intelligence*, vol. 3789, (A. F. Gelbukh, A. de Albornoz, and H. Terashima-Marín, eds.), pp. 386–395, Springer, 2005.

[394] Y. Zhang and D. A. Adjeroh, "Prediction by partial approximate matching for lossless image compression," *IEEE Transactions on Image Processing*, vol. 17, no. 6, pp. 924–935, 2008.

[395] Y.-Q. Zhang, F. Pereira, T. Sikora, and C. Reader, "Special Issue on MPEG-4," *IEEE Transactions on Circuit and Systems for Video Technology*, vol. 7, no. 1, 1997.

[396] R. Zhu and T. Takaoka, "A technique for two-dimensional pattern matching," *Communications of the ACM*, vol. 32, no. 9, pp. 1110–1120, September 1989.

[397] J. Ziv and A. Lempel, "A universal algorithm for sequential data compression," *IEEE Transactions on Information Theory*, vol. IT-23, pp. 337–343, 1977.

[398] J. Ziv and A. Lempel, "Compression of individual sequences via variable rate coding," *IEEE Transactions on Information Theory*, vol. IT-24, pp. 530–536, 1978.

[399] N. Ziviani, E. S. Moura, G. Navarro, and R. Baeza-Yates, "Compression: A key for next generation text retrieval systems," *IEEE Computer*, vol. 33, no. 11, pp. 37–44, 2000.