Content-Based Retrieval of Compressed Videos

Jurandy Almeida, Neucimar J. Leite, and Ricardo da S. Torres
Institute of Computing, University of Campinas – UNICAMP
13083-852, Campinas, SP – Brazil
Email: {jurandy.almeida, neucimar, rtorres}@ic.unicamp.br

***Abstract*—The fast evolution of technology has led to a growing demand for video data, increasing the amount of research into efficient systems to manage video material. Most of existing systems for video retrieval rely on algorithms and methods which are computationally expensive, in terms of both time and space, limiting their application to the academic world or big companies. Contrary to this trend, the market has shown a growing demand for mobile and embedded devices. In this scenario, it is imperative the development of effective and efficient techniques in order to make such technologies available to a larger public. In this context, this PhD thesis introduces five novel approaches for the analysis, indexing, and retrieval of digital videos. All these contributions are combined to create a computationally fast system for content-based video retrieval, which is able to achieve a quality level superior to current solutions.**

***Keywords*-digital video; multimedia systems; information retrieval; databases; image processing;**

I. Introduction

Advances in data compression, data storage, and data transmission have facilitated the way videos are created, stored, and distributed. The increase in the amount of video data has enabled the creation of large digital video libraries. This has spurred great interest for systems that are able to efficiently manage video material. Making efficient use of video information requires the development of systems able to extract high-level semantics from low-level features of the video content, known as content-based video retrieval systems [1].

Figure 1 shows a basic architecture of those systems. Two main functionalities are supported [2]: data insertion and query processing. The data insertion is responsible for extracting appropriate features from videos and storing them into the video database (see dashed modules and arrows). This process is usually performed offline.

The query processing, in turn, is organized as follows: the interface allows a user to specify a query pattern and to visualize the similar videos. The query-processing module extracts a feature vector from the query pattern and applies a distance function to evaluate its similarity regarding to videos in the database. Next, it ranks those videos in a decreasing order of similarity to the query pattern and forwards the most similar ones to the interface module.

Figure 2 shows a flowchart of systems based on this architecture. Due to the complexity of the video material, there are five main challenges in designing such systems [1]: (1) to divide the video stream into manageable segments according to its organization structure; (2) to implement algorithms for encoding the low-level features of each video segment into feature vectors; (3) to develop similarity measures for comparing these segments by using their feature vectors; (4) to quickly answer similarity queries over a huge amount of video sequences; and (5) to present the list of results in a user-friendly way.

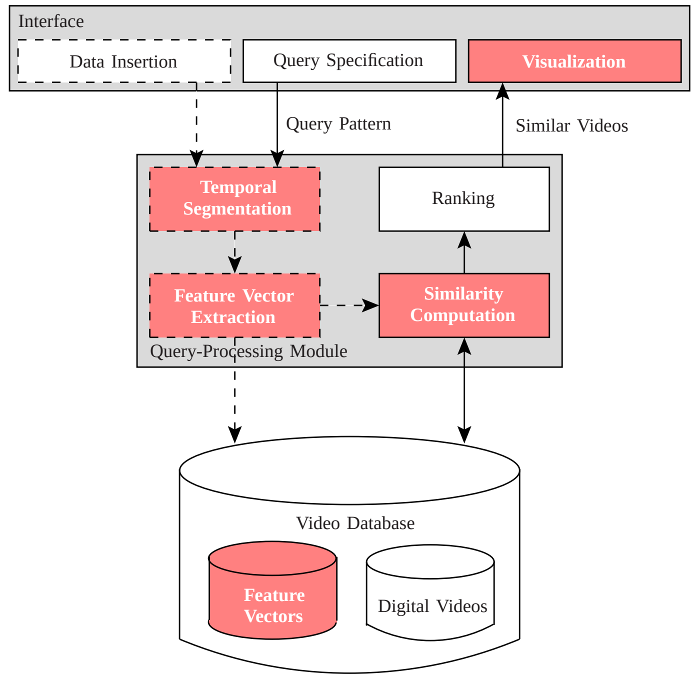


Fig. 1. Typical architecture of a content-based video retrieval system.

Numerous techniques have been proposed to support such requirements. Most of existing works involve algorithms and methods which are computationally expensive, in terms of both time and space, limiting their application to the academic world or big companies. Contrary to this trend, the market has shown a growing demand for mobile and embedded devices. In this scenario, it is imperative the development of effective and efficient techniques in order to make such technologies available to a larger public.

The work developed in this PhD thesis has contributed to address such a problem. In this context, its goal is to offer solutions for several open problems in the literature of video processing, more specifically, focused on the analysis, indexing, and retrieval of digital videos; and to contribute for overcoming the research challenges involved in specifying and implementing content-based video retrieval systems that can

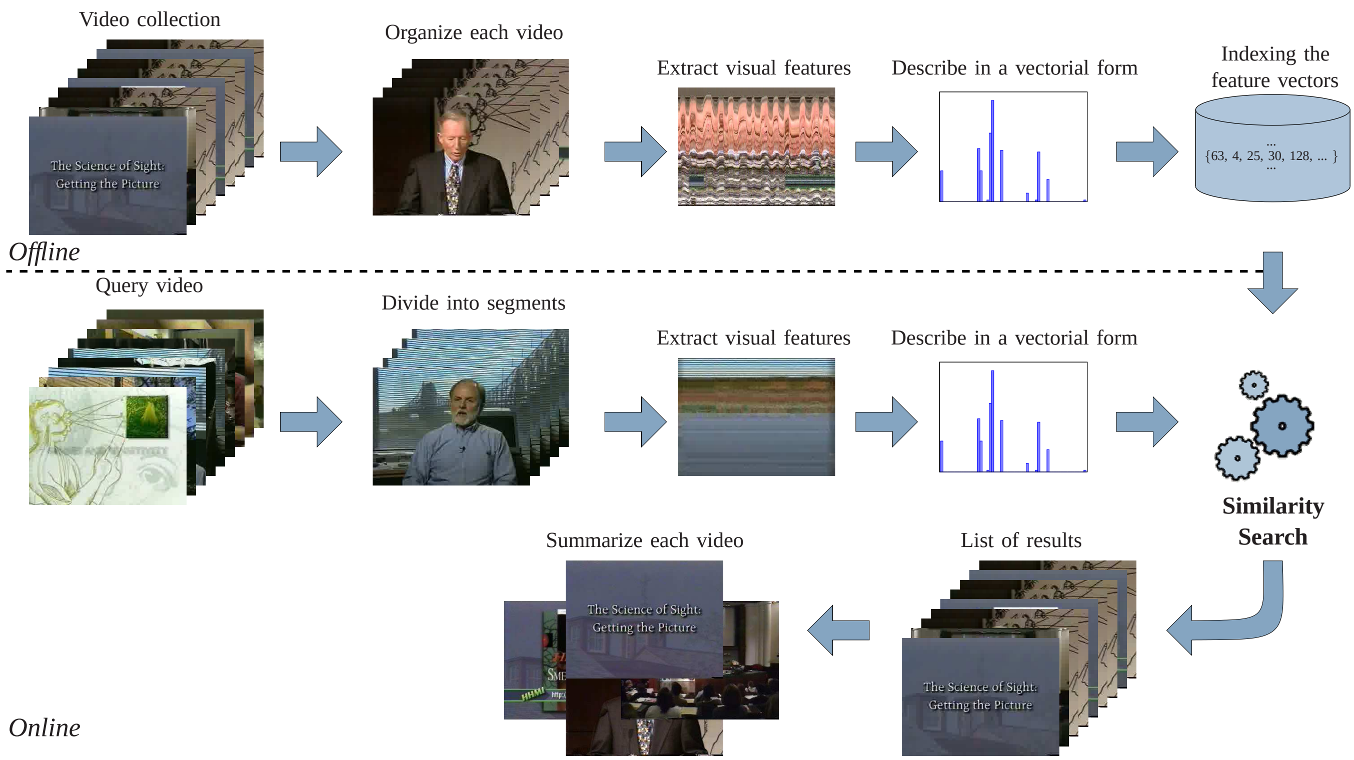


Fig. 2. Flowchart of a content-based video retrieval system.

be applied to devices with a limited hardware capacity and/or environments which require a quick response, maintaining a quality level comparable to the state-of-the-art solutions.

The main contribution of the thesis is the introduction of five novel approaches, one for each module of the basic architecture of a content-based video retrieval system, these which are highlighted in the Figure 1. All of those techniques were designed to be, at the same time, efficient and effective, in order to make them scalable and, hence, suitable for large video collections. In this way, they can be applied successfully to devices with a limited hardware capacity. They are [3]:

1) A novel approach for temporal segmentation of video sequences (Temporal Segmentation). The computational efficiency of this technique makes it suitable for online tasks [4].
2) A novel approach for estimating camera motion in video sequences (Feature Vector Extraction). In this technique, motion parameters are directly related to physical operations of the camera [5].
3) A novel approach for comparing video sequences (Similarity Computation). The computational efficiency of this technique makes it suitable for large video collections [6].
4) A novel indexing structure for performing similarity search in metric spaces (Feature Vectors). This technique is scalable, which makes it suitable for large volume of data [7].
5) A novel approach for summarizing video sequences that allows for user interaction (Visualization). This technique was designed for producing both static and dynamic video summaries in online tasks [8]–[10].

Another contribution is an empirical evaluation of those techniques against classical methods for the analysis, indexing, and retrieval of digital videos. For that, lots of experiments were conducted on several video collections, using unbiased, controlled, and reproducible experimental protocols. All of those experiments were carefully designed to ensure statistical significance, allowing to discern between genuine and casual differences in performance.

Finally, all of those contributions are combined to create a computationally fast system for content-based video retrieval, which is able to achieve a quality level superior to current solutions. Besides the aforementioned contributions, many others derived from this PhD work were published in [11]–[16]. In the following sections, we detail each of the contributions obtained from the developed research.

II. Temporal Segmentation

The first step to manage video data is to divide them into a set of meaningful and manageable units, so that the video content remains consistent in terms of camera operations and visual events. This has been the goal of a well-known research area, called video segmentation.

Different techniques have been proposed in the literature to address the temporal segmentation of video sequences [17], [18]. Many of those research works have focused on the uncompressed domain. Although existing methods provide a high quality, decoding and analyzing of a video sequence are two extremely time-consuming tasks and require a huge amount of space.

The contribution published in [4] introduces a novel approach for temporal segmentation of video sequences that operates directly in the compressed domain. It relies on exploiting visual features extracted from the video stream and on a simple and fast algorithm to detect the video transitions. The improvement of the computational efficiency makes this technique suitable for online tasks.

Figure 3 shows a block diagram of the proposed algorithm. For each type of video frame (intra-coded, predicted, or bidirectionally predicted), it is defined an independent detection module, which implements appropriate detection criteria to such a type. The partial decoding module simply parses an input video and dispatches data to the three detection modules. Finally, an information fusion module collects the results and determines the video transitions.

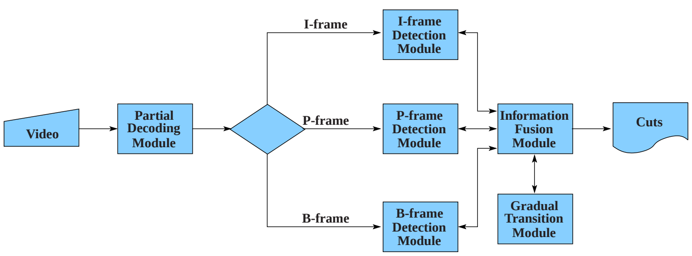


Fig. 3. Block diagram of the proposed algorithm for cut detection.

The proposed algorithm was evaluated on a real-world video dataset with different video genres and compared with the most popular approaches for temporal video segmentation. Results from an experimental evaluation over several types of video transitions show that such method presents high accuracy and computational speed. For a more detailed discussion, refer to [3], [4].

III. Representation of Visual Content

Making efficient use of video information requires that the data be stored in a compact way. For this, it must be associated with appropriate features in order to allow any future retrieval. An important feature in video sequences is the temporal intensity change between successive video frames, which is generally attributed to the motion caused by object movement or introduced by camera operation.

Numerous algorithms have been proposed in the literature to estimate camera motion from video sequences [11], [19]–[21]. These solutions are typically based on a two step approach: first, identifying correspondences (motion) between consecutive frames and, then, associating to a parametric form (model) describing the displacement of the video content. The most popular and widely used model is the affine model, which is not directly related to the physical operations of the camera.

The contribution published in [5] introduces a novel approach to estimate camera motion in video sequences, based on optical flow models. The proposed method generates the camera model using linear combinations of prototypes of optical flow produced by each camera operation.

Figure 4 illustrates the prototypes of optical flow generated by panning, tilting, zooming, and rolling, respectively. By considering that the visual field of the camera is small, an ideal model of the optical flow produced by a camera operation can be established from a numerical expression for the relationship between motion vectors, creating a prototype of optical flow. Thus, a real model of the optical flow produced by a camera motion can be approximated by a weighted combination of prototypes for camera operations.

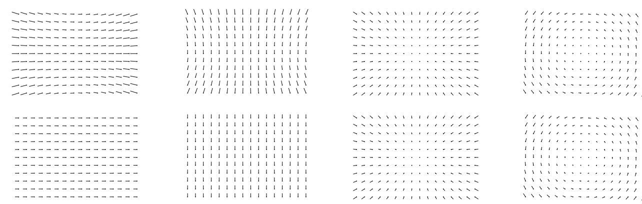


Fig. 4. The real (top) and ideal (bottom) models of optical flow generated by panning, tilting, zooming, and rolling, respectively (left to right).

For the validation of such a technique, several experiments were conducted on a synthetic test set and real-world video sequences, including all kinds of camera motion and many of their possible combinations. Results show that the proposed method is more effective than approaches based on the affine model. For a more detailed discussion, refer to [3], [5].

IV. Similarity Measure

After obtaining compact representations, it is still necessary to define a similarity measure for comparing video sequences from their signatures. There are two issues concerning this task: robustness and discriminability. Robustness is the amount of data inconsistency tolerated by the system before the occurrence of a false positive. Discriminability is the ability of the system to reject irrelevant data and reduce false positives [22].

Different techniques have been proposed in the literature to address the problem of comparing video sequences [22], [23]. Although existing methods provide a high quality in terms of robustness and discriminability, the main drawback of them is that the generated signature is prohibitive in terms of storage space, and their comparison using a similarity measure based on a frame-by-frame approach is impractical for very large databases.

The contribution published in [6] introduces a novel approach for comparing video sequences that operates directly in the compressed domain. It relies on recognizing motion patterns extracted from the video stream, which are accumulated to form a normalized histogram. This computationally simple approach is robust to several distortions and transformations. The improvement of the computational efficiency makes this technique suitable for huge collections of video data.

Figure 5 presents a flowchart of the proposed method. Initially, each macroblock is represented by the average intensity value of its four luminance blocks. Next, an ordinal matrix is obtained by ranking the intensity values of the macroblocks. This strategy is employed for computing the spatial ordinal measure of the 4-blocks of a macroblock and the temporal ordinal measure of the corresponding blocks in three frames (previous, current, and next). Finally, the spatio-temporal pattern of all the macroblocks of the video sequence are accumulated to form a normalized histogram.

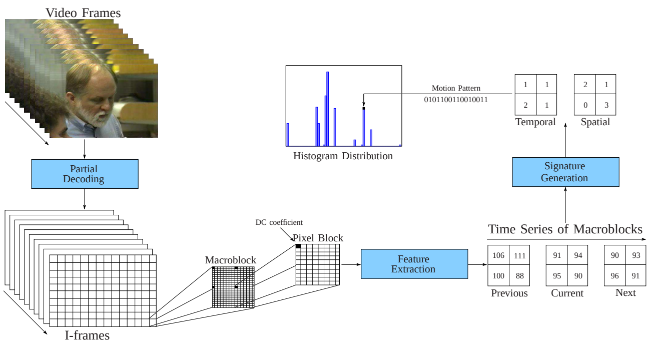


Fig. 5. An overview of the proposed method for similarity computation.

The proposed algorithm was evaluated on about 11,500 videos (400 hours) of the TRECVID 2010 dataset (IACC.1) and compared with recent approaches in the literature of video similarity detection. Results from an experimental evaluation over several types of video transformations show that such method presents high accuracy and computational speed in identifying similar videos. Moreover, this approach was successfully applied for video geocoding. Considering methods based only on visual information, it ranked 1st out of 4 groups in the MediaEval 2011 Placing Task [15]. For a more detailed discussion, refer to [3], [6].

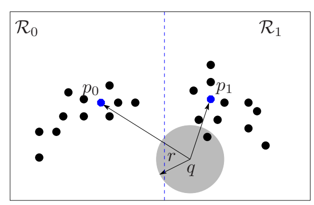
V. Data Indexing

When a user specifies a query pattern to a system, a video signature is extracted from the input pattern and a similarity measure is applied to identify all of the similar signatures in a database. To ensure a quick response, it is imperative the development of algorithms to speed up that process. Elaborated indexing structures have been proposed in order to organize video signatures and facilitate the search for similar videos [24].

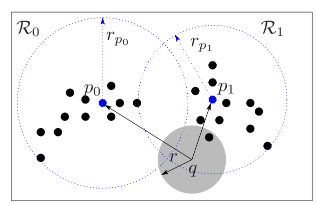
Most of existing indexes employed to accelerate data retrieval are constructed by partitioning a set of objects using only information of distances between them. In order to keep the balance of the structure, the dataset is divided into even-sized parts, ignoring the inherent grouping of their objects. In general, those techniques can be divided into two different categories. One type of methods produces partitions and, hence, they are disjoint, which may separate nearby objects, seriously affecting the search performance. The other type of methods produces groups and, hence, they may overlap, which may considerably degrade the query time [25].

The contribution published in [7] analyzes the performance of a new indexing-structure, called Ball-and-Plane tree (BP-tree), which is constructed by dividing a set of objects into compact clusters. It combines the advantages of both disjoint (partitions) and non-disjoint (groups) paradigms in order to achieve a structure of tight and low overlapping clusters, yielding significantly improved performance on performing similarity search.

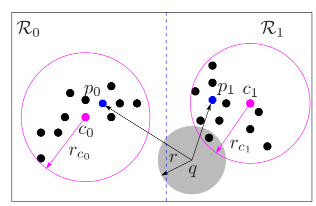
Figure 6 illustrates how BP-tree and the two types of indexing paradigms handle the dataset and the query. The disjoint approaches partition the dataset by defining a cut (dashed line) between representatives (Figure 6(a)). In the example, the query region intersects two partitions, thus two nodes must be accessed in order to answer the query. The non-disjoint approaches cluster the dataset around representatives and use a bounding region (dotted circle) to represent each group (Figure 6(b)). Usually, those bounding regions do overlap. In the figure, the region of the query response is covered by two bounding regions and, hence, two nodes must be accessed for the query resolution. BP-tree partitions the dataset by defining a cut (dashed line) between representatives and, for each partition, it selects the reference object which establishes a bounding region (solid circle) with the minimum covering radius (Figure 6(c)). In this case, triangle inequality is used to prune one subtree, thus only one node must be accessed.



(a) Partitions (disjoint).



(b) Groups (non-disjoint).



(c) BP-tree.

Fig. 6. Example of how BP-tree benefits and combines the advantages of both disjoint and non-disjoint approaches.

Those properties of BP-tree are supported by an extensive experimental evaluation performed on several real-world datasets. Results demonstrate that such method consistently outperforms the traditional solutions. Moreover, BP-tree is scalable, exhibiting a sublinear behavior regarding the number of indexed objects, which makes it well-suited for very large datasets. For a more detailed discussion, refer to [3], [7].

VI. Visualization of Results

In the end, users are presented with a list of videos related to a given query-pattern. It is impractical to assume that a user will watch all the content of those videos or a reasonable part of them in order to know what topic they are really discussing. Therefore, it is important to provide users with a concise video representation to give an idea of a video content, without having to watch it entirely, so that a user can decide whether watch the entire video or not.

Different techniques have been proposed in the literature to address the problem of summarizing a video sequence [26], [27]. Many of those research works have focused on the uncompressed domain. Due to the long time spent for decoding and analyzing a video sequence, the summaries are often produced completely offline, stored, and delivered to a user when requested. The drawback of this scheme is the complete lack of user customization.

The contribution published in [8]–[10] introduces VISON[1] (VIdeo Summarization for ONline applications), a novel approach for video summarization that operates directly in the compressed domain. It offers customization: users can control the quality of video summaries and also specify the time they are willing to wait. Such a user interaction is becoming more and more important in the current scenario, as users often have different demands and resources.

VISON was designed to be simple and efficient in order to produce video summaries in a reasonable time and with an acceptable quality, so as to allow online usage. The online production of a video summary implies a number of challenges with respect to the offline approaches. The small delay, progressive generation, and lack of complete information about the video sequence (e.g., content and length) potentially affect the quality of a video summary.

A flowchart of VISON is shown in Figure 7. For each frame of an input sequence, visual features are extracted from the video stream for describing its visual content. After that, a simple and fast algorithm is used to detect groups of video frames with a similar content and for selecting a user-defined number of representative frames per each group. Finally, the selected frames are filtered in order to avoid possible redundant or meaningless frames in the video summary.

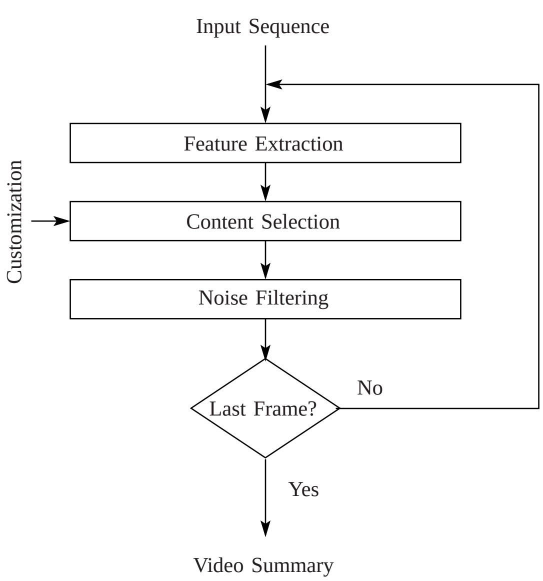


Fig. 7. Flowchart of VISON.

[1]http://www.recod.ic.unicamp.br/~jurandy/vison/ As of May 2012.

The proposed algorithm was evaluated both on videos from the Open Video Project and also on videos from the YouTube and compared with recent approaches in the literature of video summarization. The experiments were diligently designed to ensure statistical significance. Results from a subjective evaluation with users show that such method produces video summaries with high quality and computational speed. For a more detailed discussion, refer to [3], [8]–[10].

VII. Conclusion

This PhD work has performed a comprehensive study on content-based video retrieval systems, covering topics that range from video representation using visual properties, such as color, motion, shape, and texture; until aspects of indexing and storing those information. It was completed in four years and has resulted in five international journal papers [9], [10], [12], [13], [16] and eight international conference papers [4]–[8], [11], [14], [15].

More specifically, this PhD thesis has introduced five novel approaches for analyzing, indexing, and retrieving digital videos. The result of combining the proposed methods is the specification and implementation of a computational framework for video retrieval that is, at the same time, efficient and effective, and can be applied successfully to devices with a limited hardware capacity and/or environments which require a quick response, achieving a quality level superior to current solutions.

In summary, the main contributions of the thesis are:

1) A comprehensive study on algorithms and methods for content-based video retrieval.
2) Proposal of five novel approaches for the analysis, indexing, and retrieval of digital videos [4]–[10].
3) A statistically well-founded experimental evaluation of the proposed methods compared to current solutions.
4) Specification and implementation of a computationally fast system for devices with a limited hardware capacity.

There are several planned extensions, both from a theoretical and implementation point of view, for the different modules of the system architecture. Some of those extensions, by the way, are being analyzed. They include:

- **Extension to other video formats.** To gain computational efficiency, the proposed system exploits video compression standards in order to process videos directly in the compression domain. Due to its stability and simplicity, the MPEG video compression standard and, more specifically, the MPEG-1/2 video format, was chosen as the basis on which the proposed methods were founded. In general, different video compression standards follow a same basic paradigm. Therefore, those solutions can be extended to different compression standards.
- **Extension to other visual features.** This thesis has focused on motion information. This thesis has focused on motion information. However, other visual properties can be explored in the analysis, indexing, and retrieval of digital videos. A potential feature is the temporal

consistency of video data. Temporal consistency refers to the observation that temporally adjacent video shots have similar visual and semantic content. This implies that relevant shots matching a specific semantic concept or a query topic tend to gather in temporal neighborhoods or even appear next to each other consecutively. In this way, it is possible to make a more detailed analysis of the relevance of a video shot by considering the relevance of its neighboring shots.

- **Integration of different data types.** The query processing involving video content may include different data types, such as text, sound, and image. In this sense, an extension of the proposed system consists in investigating the use of fusion strategies for combining results from methods to analyze different data types. One way of integrating those data is to extend the current system as a composition of smaller subsystems, one for each data type. This can be done in a hierarchical (i.e., the output of a subsystem is used as the input of other ones, improving results at each iteration) or parallel (i.e., each subsystem is processed in an independent manner and the final result is given by combining individual responses) fashion.

Acknowledgment

This research was partially supported by Brazilian agencies FAPESP (Grant 07/52015-0, 08/50837-6, and 09/18438-7), CNPq (Grant 311309/2006-2, 472402/2007-2, and 306587/2009-2), and CAPES (Grant 01P-05866/2007).

References

[1] A. F. Smeaton, "Techniques used and open challenges to the analysis, indexing and retrieval of digital video," *Information Systems*, vol. 32, no. 4, pp. 545–559, Jun. 2007.

[2] R. S. Torres, J. A. M. Zegarra, C. D. Ferreira, O. A. B. Penatti, F. A. Andaló, and J. Almeida, "Recuperação de imagens: Desafios e novos rumos," in *Seminário Integrado de Software e Hardware (SEMISH'08)*. Belem, PA, Brazil: SBC, Jul. 12–18 2008, pp. 223–237.

[3] J. Almeida, "Recuperação de vídeos comprimidos por conteúdo," Ph.D. dissertation, Instituto de Computação, Unicamp, Campinas, SP, Brazil, Nov. 2011.

[4] J. Almeida, N. J. Leite, and R. S. Torres, "Rapid cut detection on compressed video," in *Iberoamerican Congress on Pattern Recognition (CIARP'11)*, ser. Lecture Notes in Computer Science, C. S. Martín and S.-W. Kim, Eds., vol. 7042. Pucón, Chile: Springer, Nov. 15–18 2011, pp. 71–78.

[5] J. Almeida, R. Minetto, T. A. Almeida, R. S. Torres, and N. J. Leite, "Robust estimation of camera motion using optical flow models," in *International Symposium on Advances in Visual Computing (ISVC'09)*, ser. Lecture Notes in Computer Science, G. Bebis, R. D. Boyle, B. Parvin, D. Koracin, Y. Kuno, J. Wang, R. Pajarola, P. Lindstrom, A. Hinkenjann, M. L. Encarnacao, C. Silva, and D. Coming, Eds., vol. 5875. Las Vegas, NV, USA: Springer, Nov. 30 – Dec. 2 2009, pp. 435–446.

[6] J. Almeida, N. J. Leite, and R. S. Torres, "Comparison of video sequences with histograms of motion patterns," in *IEEE International Conference on Image Processing (ICIP'11)*, B. Macq and P. Schelkens, Eds. Brussels, Belgium: IEEE Computer Society, Sep. 11–14 2011, pp. 3673–3676.

[7] J. Almeida, R. S. Torres, and N. J. Leite, "BP-tree: An efficient index for similarity search in high-dimensional metric spaces," in *ACM International Conference on Information and Knowledge Management (CIKM'10)*. Toronto, ON, Canada: ACM Press, Oct. 26–30 2010, pp. 1365–1368.

[8] ——, "Rapid video summarization on compressed video," in *IEEE International Symposium on Multimedia (ISM'10)*. Taichung, Taiwan: IEEE Computer Society, Dec. 13–15 2010, pp. 113–120.

[9] J. Almeida, N. J. Leite, and R. S. Torres, "VISON: VIdeo Summarization for ONline applications," *Pattern Recognition Letters*, vol. 33, no. 4, pp. 397–409, Mar. 2012.

[10] ——, "Online video summarization on compressed domain," *Journal of Visual Communication and Image Representation*, 2012, DOI: 10.1016/j.jvcir.2012.01.009.

[11] J. Almeida, R. Minetto, T. A. Almeida, R. S. Torres, and N. J. Leite, "Estimation of camera parameters in video sequences with a large amount of scene motion," in *International Conference on Systems, Signals and Image Processing (IWSSIP'10)*. Rio de Janeiro, RJ, Brazil: IEEE Computer Society, Jun. 17–19 2010, pp. 348–351.

[12] J. Almeida, E. Valle, R. S. Torres, and N. J. Leite, "DAHC-tree: An effective index for approximate search in high-dimensional metric spaces," *Journal of Information and Data Management*, vol. 1, no. 3, pp. 375–390, Oct. 2010.

[13] S. M. Pinto-Cáceres, J. Almeida, V. P. A. Neris, M. C. C. Baranauskas, N. J. Leite, and R. S. Torres, "Navigating through video stories using clustering sets," *International Journal of Multimedia Data Engineering and Management*, vol. 2, no. 3, pp. 1–20, Sep. 2011.

[14] N. P. Kozievitch, J. Almeida, R. S. Torres, A. Santanchè, and N. J. Leite, "Reusing a compound-based infrastructure for searching video stories," in *IEEE International Conference on Information Reuse and Integration (IEEE-IRI'11)*. Las Vegas, NV, USA: IEEE Computer Society, Aug. 3–5 2011, pp. 222–227.

[15] L. T. Li, J. Almeida, and R. S. Torres, "RECOD working notes for placing task mediaeval 2011," in *Working Notes Proceedings of the MediaEval 2011 Workshop (MEDIAEVAL'11)*. Pisa, Italy: CEUR-WS.org, Sep. 11–14 2011.

[16] N. P. Kozievitch, J. Almeida, R. S. Torres, N. J. Leite, M. A. Gonçalves, U. Murthy, and E. A. Fox, "Towards a formal theory for complex objects and content-based image retrieval," *Journal of Information and Data Management*, vol. 2, no. 3, pp. 321–336, Oct. 2011.

[17] I. Koprinska and S. Carrato, "Temporal video segmentation: A survey," *Signal Processing: Image Communication*, vol. 16, no. 5, pp. 477–500, Jan. 2001.

[18] R. Lienhart, "Reliable transition detection in videos: A survey and practitioner's guide," *International Journal of Image and Graphics*, vol. 1, no. 3, pp. 469–486, Jul. 2001.

[19] M. V. Srinivasan, S. Venkatesh, and R. Hosie, "Qualitative estimation of camera motion parameters from video sequences," *Pattern Recognition*, vol. 30, no. 4, pp. 593–606, Apr. 1997.

[20] S.-C. Park, H.-S. Lee, and S.-W. Lee, "Qualitative estimation of camera motion parameters from the linear composition of optical flow," *Pattern Recognition*, vol. 37, no. 4, pp. 767–779, Apr. 2004.

[21] R. Minetto, N. J. Leite, and J. Stolfi, "Reliable detection of camera motion based on weighted optical flow fitting," in *International Conference on Computer Vision Theory and Applications (VISAPP'07)*, A. Ranchordas, H. Araújo, and J. Vitrià, Eds. Barcelona, Spain: Institute for Systems and Technologies of Information, Control and Communication, Mar. 8–11 2007, pp. 435–440.

[22] C. Kim and B. Vasudev, "Spatiotemporal sequence matching for efficient video copy detection," *IEEE Transactions on Circuits Systems and Video Technology*, vol. 15, no. 1, pp. 127–132, Jan. 2005.

[23] J. Law-To, L. Chen, A. Joly, I. Laptev, O. Buisson, V. Gouet-Brunet, N. Boujemaa, and F. Stentiford, "Video copy detection: A comparative study," in *ACM International Conference on Image and Video Retrieval (CIVR'07)*, N. Sebe and M. Worring, Eds. Amsterdam, The Netherlands: ACM Press, Jul. 9–11 2007, pp. 371–378.

[24] E. Chávez, G. Navarro, R. A. Baeza-Yates, and J. L. Marroquín, "Searching in metric spaces," *ACM Computing Surveys*, vol. 33, no. 3, pp. 273–321, Sep. 2001.

[25] P. Zezula, G. Amato, V. Dohnal, and M. Batko, *Similarity Search: The Metric Space Approach*. Secaucus, NJ, USA: Springer-Verlag, Inc., 2005.

[26] A. G. Money and H. W. Agius, "Video summarization: A conceptual framework and survey of the state of the art," *Journal of Visual Communication and Image Representation*, vol. 19, no. 2, pp. 121–143, Feb. 2008.

[27] B. T. Truong and S. Venkatesh, "Video abstraction: A systematic review and classification," *ACM Transactions on Multimedia Computing, Communications, and Applications*, vol. 3, no. 1, pp. 1–37, Feb. 2007.