

Ten Research Questions for Scalable Multimedia Analytics

Björn Þór Jónsson, Marcel Worring, Jan Zahálka, Stevan Rudinac, Laurent Amsaleg

► **To cite this version:**

Björn Þór Jónsson, Marcel Worring, Jan Zahálka, Stevan Rudinac, Laurent Amsaleg. Ten Research Questions for Scalable Multimedia Analytics. The 22nd International Conference On Multimedia Modelling, Special Session Perspectives on Multimedia Analytics, Jan 2016, Miami, United States. The 22nd International Conference On Multimedia Modelling, Special Session Perspectives on Multimedia Analytics The 22nd International Conference On Multimedia Modelling, Special Session Perspectives on Multimedia Analytics. <hal-01214847>

HAL Id: hal-01214847

<https://hal.inria.fr/hal-01214847>

Submitted on 13 Oct 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Ten Research Questions for Scalable Multimedia Analytics

Björn Þór Jónsson^{1,2}, Marcel Worring², Jan Zahálka²,
Stevan Rudinac², and Laurent Amsaleg³

¹ School of Computer Science, Reykjavik University, Iceland, bjorn@ru.is

² Informatics Institute, University of Amsterdam, Netherlands,

{m.worring|j.zahalka|s.rudinac}@uva.nl

³ IRISA, Rennes, France, laurent.amsaleg@irisa.fr

Abstract. The scale and complexity of multimedia collections is ever increasing, as is the desire to harvest useful insight from the collections. To optimally support the complex quest for insight, multimedia analytics has emerged as a new research area that combines concepts and techniques from multimedia analysis and visual analytics into a single framework. State of the art multimedia analytics solutions are highly interactive and give users freedom in how they perform their analytics task, but they do not scale well. State of the art scalable database management solutions, on the other hand, are not yet designed for multimedia analytics workloads. In this position paper we therefore argue the need for research on *scalable multimedia analytics*, a new research area built on the three pillars of visual analytics, multimedia analysis and database management. We propose a specific goal for scalable multimedia analytics and present several important research questions that we believe must be addressed in order to achieve that goal.

1 Introduction

In the last decade, we have witnessed a revolution in all aspects of computing technology as the human ability to produce, store and share data has truly exploded. This data contains very valuable information for individuals, enterprises and society, and as a result we have seen a sharp rise in interest in big data analytics and related fields. Big data is typically characterised using “the three Vs”—Volume, Velocity, and Variety—which indicate respectively that the data is bountiful, is produced continuously, and contains a large variety of information.

While big data analytics has focused on relatively structured data, such as business data and transaction logs, much of the information explosion has taken the form of multimedia, in particular images and videos, along with user-generated annotations and automatically generated metadata, for example from the capturing device or a social media service. Current and potential applications involving large multimedia collections are numerous, ranging from personal applications (e.g., life-logging) through societal (e.g., digital heritage) and scientific (e.g., biology, astronomy, and medicine) to business applications (e.g., marketing

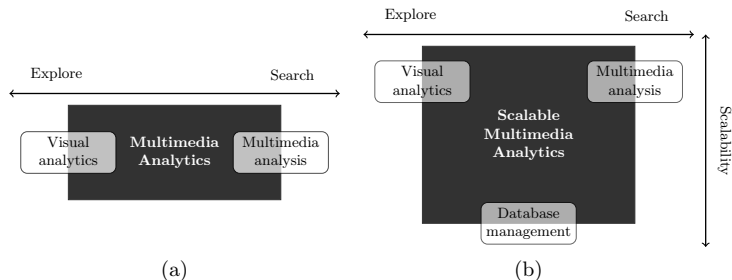


Fig. 1. Transition from (a) multimedia analytics to (b) scalable multimedia analytics

and profiling). What many of these media collections have in common is that they have the potential to significantly change the world in some way if we can manage to extract the knowledge and insight that they encode.

Unfortunately, existing big data analytics methods are not directly applicable to the multimedia domain. First, since the data is multimedia, automatically understanding the content and context of the data must be done at various levels of abstraction and, because it is very difficult, it is best done in collaboration between man and machine. This collaboration requires systems to learn in real-time the intention of users, the patterns they indicate, and support their interactions with the data. Developing general methods and tools for harvesting information from multimedia collections therefore represents a significant unsolved challenge.

1.1 Multimedia Analytics

Enter the new field of *multimedia analytics*. This field, which combines visual analytics with multimedia analysis as depicted in Figure 1(a), has been developing over the last half decade [2]. The goal of multimedia analytics is to produce the processes, techniques and tools to allow users to efficiently and effectively analyze multimedia collections in order to gain insight and knowledge [2]. While multimedia analysis should be stressed to its limits to extract information automatically from the media files, visual analytics proposes an interactive process that must involve the user, through data selection and visualization techniques.

Understanding semantics of multimedia content brings many difficult challenges. Machines cannot match humans, both in the richness of semantics extracted and the speed of the extraction, while humans have great difficulty processing large multimedia collections. Combining the strengths of human and machine while alleviating the weaknesses, and providing interactive experience for a variety of analytics tasks, is the major challenge of multimedia analytics.

Addressing the analytic challenge is already daunting when collections are of moderate size. As the collections grow in size and scope, supporting the analytics process efficiently and effectively through data management becomes increasingly important. State of the art multimedia analytics solutions are highly interactive and give users freedom in how they perform their task, but they do

not scale well. State of the art scalable data management solutions, on the other hand, are not designed for interactive analysis.

1.2 Scalability Challenges

In big data analytics, the requirements for data management are often described using the three Vs—volume, velocity and variety.¹ For multimedia analytics, since the nature of multimedia collections and applications leads to some specific requirements, the following major axes of scalability must be addressed:

Volume: The size of the collection is obviously the first scalability axis. Due to the large file sizes of multimedia items, some collections are enormous in their sheer size, making any manipulation of such collections very difficult. While storage capacity and cost is generally not an issue anymore, user time—and therefore processing time—is. The main interest of volume is therefore the impact on remaining axes.

Variety: Current multimedia analytics research projects generally focus on solving a particular problem. Data may have complex structure and arise from many sources, but significant effort is spent on reducing both data quantity and complexity to make the analytics process more manageable. We predict, on the other hand, that the aim will become to understand and analyze whole domains, with a variety of multimedia data coming from sources that may not yet be fully understood, such as the “Internet of Things.” This requires gathering data for future use and retaining much more of the potentially useful data. Such data will be even more voluminous, but in particular more complex, requiring more effort to manage and analyze.

Velocity: As these large-scale collections grow fast and are long-lived, data is added incrementally and continuously, and data sources may come and go, so we must necessarily support incremental analysis. Furthermore, multimedia represents a world that is changing fast, both literally and in terms of its representation in the multimedia collections. Many different users will work with each collection over a long period of time and they must contribute their knowledge to the collection for the benefit of concurrent and future users, who will then add to this knowledge or even negate it.

Visual Interaction: Users are not considered a scalability axis for big data analytics, as big data tends to be used by few expert users. For multimedia analytics, however, users and user interactions play perhaps the most important role. In some domains, the number and diversity of users may be very significant, as aspects of the collection they work on may be separated depending on their objectives, roles, location and time. Because visual interfaces are mandatory, we call this axis visual interaction.

Note that the first two axes stem from the multimedia collection itself, while the latter two stem from the user interaction with the multimedia collection.

¹ The requirements are sometimes given as four, six or even seven Vs. The additional Vs—veracity, validity, viability and value—are not important for this discussion.

1.3 Scalable Multimedia Analytics

Based on these scalability axes, we propose the following definition of the main goal of *Scalable Multimedia Analytics*:

to produce the processes, techniques and tools to allow *many diverse* users to efficiently and effectively analyze *large* and *dynamic* multimedia collections over a *long period of time* to gain insight and knowledge.

We argue that scalable multimedia analytics must rest on the three pillars shown in Figure 1(b). Visual Analytics must still contribute advanced methods for presenting information and interacting with users, while Multimedia Analysis must contribute increasingly accurate and rich methods for analysing multimedia to add semantic information about its content. However, in order to scale gracefully to very large collections, Database Management must contribute advanced methods for managing storage of and access to the large multimedia collections.

Large collections obviously require advanced storage management due to their size. Their dynamic nature—rapid growth, rapid development of file formats and capturing technology, and rapid evolution of sharing models and analysis of this information—also requires techniques for supporting multimedia analytics on dynamic collections, especially when analyzing recently added material.

Furthermore, the multimedia analytics process may span a long period of time, possibly requiring the cooperation of several analysts, which must then share the insight and knowledge extracted from the collection. For this purpose a data model is required that can seamlessly integrate the insight and knowledge into the information extracted from the existing multimedia collection. Such a data model must persistently keep track of and structure the relationships between data, knowledge and context.

Finally, query processing must be supported in the analytics process, with potentially different requirements depending on the context of the analytics application, such as whether the scope of the analysis is wide or narrow. Maintaining real-time performance in this environment will require managing resources very effectively, using a range of techniques for knowledge representation, database management, and computation management.

Database techniques to address the above challenges have been proposed to support either visual analytics or multimedia analysis, but the techniques used in each case are very different. All of the work to date has thus only addressed the combination of two of the pillars, leaving a wide gap in the middle of Figure 1(b). It is clearly necessary to focus on all three pillars at the same time, if we are to make progress towards truly scalable multimedia analytics.

1.4 Contributions of this Paper

The contributions of this position paper are to i) propose and elaborate on the above definition of the goal of scalable multimedia analytics, ii) briefly review the work in the fields of multimedia analytics and database management which could help in reaching this goal, and iii) present some important research directions that we believe must be addressed in order to achieve progress in this field.

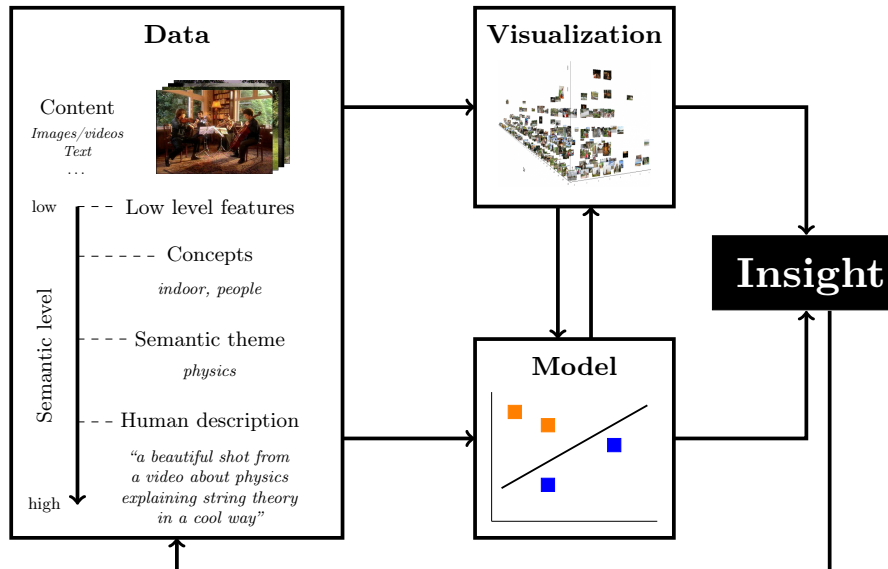


Fig. 2. Multimedia analytics process [13], adapted from the visual analytics diagram by Keim et al. [5]

2 Multimedia Analytics

In this section we discuss the current state of the art in multimedia analytics. We first consider the specific requirements of multimedia, before describing the multimedia analytics process. We conclude by considering the current approach in the field to achieving scalability.

2.1 From Multimedia Analysis: Multimedia Representation

For structured and numeric data the interpretation of the data items themselves is always at the same level of abstraction. Analytics comes about when aggregating the data and studying patterns through statistics. In contrast, an individual image or video can be interpreted in many different ways. Depending on the nature of analytics task, as illustrated in Figure 2, the multimedia content may have to be interpreted at different semantic levels, associated with: (a) low-level visual features (e.g., colour histograms), (b) semantic concepts (e.g., objects, settings and events), (c) semantic theme (e.g., physics, immigrants or cultural identity) and (d) complex human interpretation, including factors such as sentiment and aesthetic appeal.

Recent advances in multimedia analysis have opened the door to enabling search and exploration at all semantic levels. However, while the features extracted from the content are getting increasingly descriptive, the size of the resulting feature collection is still prohibitively large for real-time user interactions, a key aspect of multimedia analytics. Furthermore, there is no “universal”

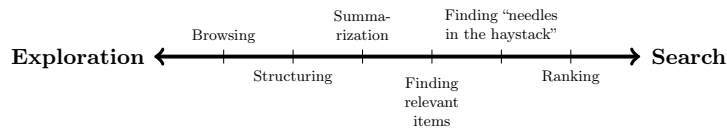


Fig. 3. The exploration-search axis with example multimedia analytics tasks [13]

feature representation satisfying relevance criteria in a wider range of analytics tasks. For example, a particular video search query may require a simple query-by-example matching based on low-level visual features, while utilizing concept detectors or automatic speech recognition may yield better results in other cases [10]. In the past several attempts have been made to standardising multimedia content descriptions, as well as multimedia items and user interactions with them (e.g., the MPEG standards). Some of the main reasons for their limited adoption were the insufficient effectiveness of early content analysis approaches and the inflexibility of description schemes with regard to accommodating emerging information-rich content sources (e.g., social multimedia portals) and complex user interaction modes.

2.2 From Visual Analytics: Multimedia Analytics Process

A large body of work related to the constituent parts of multimedia analytics has been surveyed by Zahálka and Worring [13]. The objective of multimedia analytics is *user insight*, a complex understanding of the analyzed data accumulated over time using all or most of the relevant data at hand [8]. The concept of insight is quite familiar in the field of visual analytics: the conceptual diagram by Keim et al., instantiated for multimedia analytics in Figure 2, is one of the cornerstones of the field. In contrast with visual analytics, however, a media item is more complex than a data point and the analyst cannot fully understand it before seeing it; there is thus a trade-off between giving an overview of the collection and showing the individual media items in detail.

The palette of tasks leading towards insight is quite colourful. Nevertheless, all multimedia analytics tasks have a key common characteristic: they consist of a certain proportion of exploration and search. Hence, an *exploration-search axis*, shown in Figure 3, has been proposed as the task model for multimedia analytics [13]. The analysts tilt back and forth on this axis during their quest for insight, and hence a multimedia analytics system should support the entire axis. *Analytic categorization*, i.e., maintaining a set of analyst-defined categories based on the semantic and metadata content of the multimedia collection and updating it as the analyst progresses towards insight, has been proposed as an umbrella task for the exploration-search axis [13].

Many challenges for multimedia analytics arise due to the need for overcoming two gaps. The *semantic gap* is defined as the disproportion between the semantics extractable by the humans on the one side and the machines on the

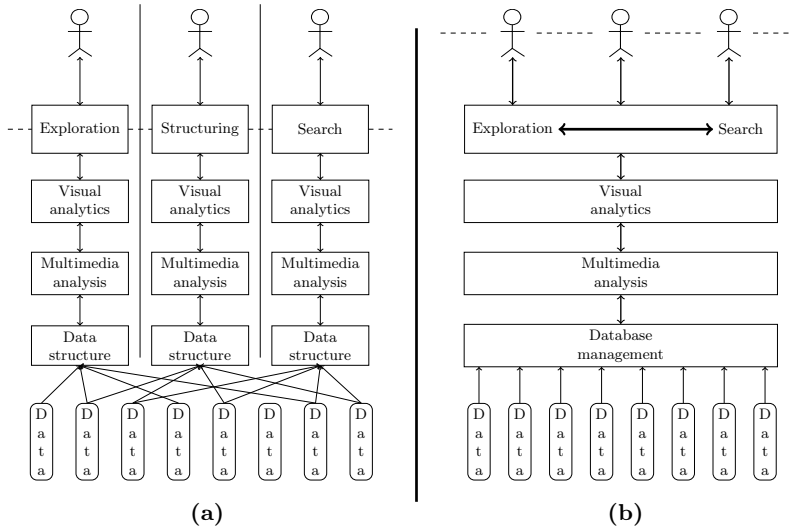


Fig. 4. Transition from (a) ad hoc to (b) scalable multimedia analytics systems

other. This longstanding challenge aiming at understanding the content of a single multimedia item is increasingly being addressed using deep learning based on huge amounts of training data. Yet the semantic gap is as prevalent at the (sub-) collection level at which it has hardly been addressed. The *pragmatic gap*, defined as the gap between the highly adaptable mental categorical model of the user on the one side and the strict, bounded definition of categorization in the machine world on the other, comes into play when exploring a multimedia collection in its context [13]. In order to close the pragmatic gap, the data model of the analytics system must fully adapt to user intent and understanding, as it varies over the duration of the analytics process, so that it accurately represents the view of the user at each time.

2.3 Scalability Considerations

Multimedia analytics state of the art, however, has up to now to the best of our knowledge not explicitly considered the issue of data management, despite aiming for large-scale analytics. This applies both to the model described above, as well as to the pioneer systems conceived so far (e.g., [1,11]). As Figure 4(a) indicates, data is only present in main memory, which limits the scale of the systems both in data volume and duration and makes the data processing ad hoc for each respective multimedia analytics system.

As illustrated in Figure 4(b), a modern multimedia database should support search and exploration through optimal utilization of available query analysis and retrieval algorithms and ideally eliminate the need for constructing a separate framework for each analytic task from features to user interaction models.

Tightly integrating a suitable data model and query processing techniques with multimedia analytics has the possibility to increase the scale of multimedia analytics and truly utilize multimedia collections as knowledge bases, rather than individual datasets. Effective data models and query languages must be able to support exploration and search based on relevance criteria defined at various semantic levels [9]. More particularly, the query model should facilitate user intent analysis and the definition of complex relevance criteria.

As the size and heterogeneity of multimedia collections increase, analysing them in their entirety becomes infeasible. The data model should thus facilitate efficient filtering of parts of the collection needed in a given analytics session. It should further enable translation of features into representations allowing for seamless interactions with the system. Finally, both data model and query language should facilitate the dynamic choice of optimal retrieval algorithms.

3 Database Support

In this section we review the existing support for multimedia search on one hand and analytics workloads on the other. We conclude that existing work is indeed insufficient to cope with scalable multimedia analytics, as defined above, and discuss some techniques that can pave the way forward.

3.1 Multimedia Search

High-dimensional indexing has been studied for decades, but it is only during the last decade that some breakthroughs have been made: it is now possible to run efficient similarity searches at a truly large scale. Some of those approaches are main-memory based [3], while others adapt gracefully to disk-based collections [6], but they all employ some sort of approximation—either in the descriptor generation or during query processing—to trade efficiency for accuracy.

Recently, researchers have started applying big data analytics tools, such as Hadoop and Spark with their map-reduce programming model, to high-dimensional indexing [7]. The conclusion of this work so far is that while these big data tools can support very large collections and provide excellent throughput, they come nowhere close to providing the interactive query response times that are required for multimedia analytics workloads. Furthermore, none of these tools provide a data model that can represent any of the complexities of multimedia analytics collections adequately.

3.2 Analytics Workloads

Analytics workloads have mostly been considered in two domains: business analytics and the more general big data analytics. As discussed above, current big data analytics tools are not suitable for multimedia analytics due to their response time. In the business analytics domain, however, data warehouses are used to extract data from their sources and stored in a database schema that keeps

sufficient information to facilitate obtaining useful and meaningful insights, yet in a format that is sufficiently simple to do so interactively. Of course, business analytics is concerned with numerical data which is much simpler than multimedia data and, unlike multimedia data, supports easy aggregation of information.

A data model for multimedia analytics must provide sufficient semantics to facilitate long-term accumulation of data and insights, yet be simple enough to allow relatively simple query formulation and efficient query processing and optimization. The relational data model, for example excels at the latter, but handles complex relationships poorly. Ad hoc data structures can handle any relationships, but query formulation then amounts to low-level programming. The multi-dimensional model of OLAP applications seems to represent a good middle ground for those applications; similarly a good middle ground must be developed for multimedia analytics workloads. This work seems to provide a direction for going forward towards multimedia analytics and indeed recent proposals for data models for multimedia analytics have been based on this work, including Multimedia PivotTables [12] and the O³ data model [4].

3.3 Database Management

Other aspects of database management are also relevant in the multimedia analytics domain. While space constraints prevent both complete coverage and full citations, the following list indicates the range of techniques that could be used: Transaction support ensures data integrity by enforcing the ACID properties of atomicity, consistency, isolation and durability; Query optimization dynamically chooses the best query processing algorithms and access paths, depending on query, data and hardware characteristics; Caching is used to dynamically retain as much data as possible in memory and process this data while fetching the remainder, to hide the latency of the underlying collection; Parallel and distributed processing are used to divide the workload to as many computing cores or computers as possible; Approximation and sampling is used to reduce the work needed to produce a first answer, which may then be incrementally updated if more time becomes available. A complete database management solution for multimedia analytics must undoubtedly draw on all of these aspects. In some cases, tried and tested techniques will be directly applicable (e.g., for transaction management) while in other cases entirely new methods must be developed based on the data model and associated query model.

4 Research Questions

In this section, we discuss several research directions arising from the discussion above that we believe must be addressed in order to make progress towards scalable multimedia analytics. This list, summarized in Table 1, expands on the multimedia analytics research agenda of [13], focusing on issues related to scalability. For clarity, we divide the research questions into the four axes—or Vs—of scalability described in Section 1.2, but of course a complete solution must address all of these issues.

Table 1. Ten Research Questions for Scalable Multimedia Analytics

- RQ1** How can database management techniques facilitate multimedia analytics for increasingly large-scale collections of multimedia items and metadata?
- RQ2** Is a novel multimedia query language needed for the database to fully support multimedia analytics, or is an extension of classic query languages sufficient?
- RQ3** Can the database management system support the different workloads that arise along the exploration-search axis?
- RQ4** Can database management techniques be utilized to improve the quality of modality fusion?
- RQ5** How can the multimedia analytics system facilitate up-to-date interactive analysis of collections that are dynamically and rapidly evolving?
- RQ6** Can persistence of the machine learning model be used to improve the quality and efficiency of the analysis?
- RQ7** How can insight history and the context of insight be represented as the insight develops through time?
- RQ8** How could long-term learning be leveraged to improve the extraction of high-level semantics?
- RQ9** How can database management techniques be leveraged to improve the user’s interactive experience?
- RQ10** How can the multimedia analytics system best support dynamic discovery of new analytic categories that lead to insight and knowledge?

4.1 Volume

With the ever increasing volume of multimedia collections in virtually all application domains, all components of multimedia analytics must handle large-scale data more efficiently. Managing increasingly large-scale data is not specific to multimedia analytics. With the advances in camera and smartphone technology, however, individual images and videos have increasingly higher resolution and detail, thus increasing not only the number of multimedia items, but also the size per item. These challenges are reflected in **RQ1**.

4.2 Variety

The variety of both data and tasks within the multimedia domain presents an interesting potential for information gain, but also brings many processing challenges. As mentioned in Section 2.1, multimedia data makes common database aggregations and data operations difficult or impossible. The limitations of current query languages with respect to semantics inspire **RQ2**.

Analysts conduct a variety of tasks in the multimedia domain. These tasks, modelled by the exploration-search axis, require different data to be presented to the user, which could be possibly handled by database management techniques, leading to **RQ3**.

A second aspect of variety is the variation in the multimedia data itself. Fusion of individual modalities is required to truly utilize the heterogeneous nature of multimedia. Efficient fusion at the database level could positively impact semantic quality, inspiring **RQ4**.

4.3 Velocity

The main challenge of velocity is that of handling collections that are growing rapidly, allowing the analysis of up-to-date information and merging these with existing information; this is represented by **RQ5**.

Furthermore, database persistence has a tremendous potential to improve incremental analysis not only by persisting data, but also by persisting elements of the analysis itself: the machine learning model used by the systems and the history of insight as the user develops it over time. Persisting and reusing the elements of the analysis reduces the start-up time of the analysis. Moreover, this enables the use of multimedia data as a true knowledge base: instead of starting the analysis every time anew, analysts are able to continue where they or their colleagues left off in previous sessions. Longitudinal analysis of the stored data might thus very well improve the accuracy of high-level semantic concepts and their boundaries, improving semantic quality in general. These considerations lead to research questions **RQ6** through **RQ8**.

4.4 Visual Interaction

The interactivity requirement of multimedia analytics places a rather stringent requirement on performance throughout the entire pipeline. Maintaining and improving interactivity with database management is a research challenge of its own, as witnessed by **RQ9**.

As mentioned in Section 2.2, analytic categorization was introduced as an umbrella model for user tasks [13]. Whether it is a sufficient model, however, remains an open question. The degree of categorization support from the database management component of the system is an open question as well. Moreover, analytic categorization involves enabling the user to discover new categories as she progresses towards the insight. These categorization-related concerns inspire the final research question **RQ10**.

5 Conclusions

In this paper we have argued that research is needed at the boundary of multimedia analytics and database management, and in fact that database management should be integrated as the third pillar of *scalable multimedia analytics*. We have presented a list of important research challenges that relate to the scalability of the multimedia analytics process. This list is no doubt incomplete and as these issues are addressed new will arise. What is important, however, is that the research community immediately starts tackling these research questions so that we can start harvesting the information encoded in today’s and tomorrow’s multimedia collections.

Acknowledgments. This work was in part supported by the CNRS PICS grant “MMAnalytics” and by sabbatical support from Reykjavik University. Thanks go to the anonymous reviewers, for comments that helped improve the paper.

References

1. Burtner, R., Bohn, S., Payne, D.: Interactive visual comparison of multimedia data through type-specific views. In: Wong, P.C., Kao, D.L., Hao, M.C., Chen, C., Healey, C.G. (eds.) *SPIE Conference on Visualization and Data Analysis (VDA)*. pp. 86540M–86540M–15. Burlingame, CA, USA (2013)
2. Chinchor, N.A., Thomas, J.J., Wong, P.C., Christel, M.G., Ribarsky, W.: Multi-media analysis + visual analytics = multimedia analytics. *IEEE Comput. Graph. Appl. Mag.* 30(5), 52–60 (2010)
3. Jégou, H., Douze, M., Schmid, C.: Product quantization for nearest neighbor search. *IEEE Trans. Pattern Anal. Mach. Intell.* 33(1), 117–128 (2011)
4. Jónsson, B.P., Tómasson, G., Sigurbórsson, H., Eiríksdóttir, Á., Amsaleg, L., Lárusdóttir, M.K.: A multi-dimensional data model for personal photo browsing. In: He, X., Luo, S., Tao, D., Xu, C., Yang, J., Hasan, M.A. (eds.) *MMM 2015*. LNCS, vol. 8936, pp. 345–356. Springer, Heidelberg (2015)
5. Keim, D.A., Kohlhammer, J., Ellis, G., Mansmann, F. (eds.): *Mastering The Information Age—Solving Problems with Visual Analytics*. Eurographics (2010)
6. Lejsek, H., Jónsson, B.P., Amsaleg, L.: NV-Tree: Nearest neighbours at the billion scale. In: De Natale, F.G.B., Del Bimbo, A., Hanjalic, A., Manjunath, B.S., Satoh, S. (eds.) *ACM International Conference on Multimedia Retrieval (ICMR)*. pp. 54:1–54:8. Trento, Italy (2011)
7. Moise, D., Shestakov, D., Guðmundsson, G.P., Amsaleg, L.: Indexing and searching 100M images with Map-Reduce. In: Jain, R., Prabhakaran, B., Worring, M., Smith, J.R., Chua, T.S. (eds.) *ACM International Conference on Multimedia Retrieval (ICMR)*. pp. 17–24. Dallas, TX, USA (2013)
8. North, C.: Towards measuring visualization insight. *IEEE Comput. Graph. Appl. Mag.* 26(3), 6–9 (2006)
9. Rudinac, S., Larson, M., Hanjalic, A.: Learning crowdsourced user preferences for visual summarization of image collections. *IEEE Trans. Multimedia* 15(6), 1231–1243 (2013)
10. Snoek, C.G.M., van de Sande, K.E.A., de Rooij, O., Huurnink, B., van Gemert, J.C., Uijlings, J.R.R., He, J., Li, X., Everts, I., Nedovic, V., van Liempt, M., van Balen, R., Yan, F., Tahir, M.A., Mikolajczyk, K., Kittler, J., de Rijke, M., Geusebroek, J.M., Gevers, T., Worring, M., Smeulders, A.W.M., Koelma, D.C.: The MediaMill TRECVID 2008 semantic video search engine. In: Over, P., Awad, G., Rose, R.T., Fiscus, J.G., Kraaij, W., Smeaton, A.F. (eds.) *TRECVID Workshop*. Gaithersburg, MD, USA (2008)
11. Viaud, M.L., Thièvre, J., Goëau, H., Saulnier, A., Buisson, O.: Interactive components for visual exploration of multimedia archives. In: Luo, J., Guan, L., Hanjalic, A., Kankanhalli, M.S., Lee, I. (eds.) *ACM International Conference on Image and Video Retrieval (CIVR)*. pp. 609–616. Niagara Falls, Canada (2008)
12. Worring, M., Koelma, D.C.: Insight in image collections by multimedia pivot tables. In: Hauptmann, A.G., Ngo, C.W., Xue, X., Jiang, Y.G., Snoek, C., Vasconcelos, N. (eds.) *ACM International Conference on Multimedia Retrieval (ICMR)*. pp. 291–298. Shanghai, China (2015)
13. Zahálka, J., Worring, M.: Towards interactive, intelligent, and integrated multimedia analytics. In: Chen, M., Ebert, D.S., North, C. (eds.) *IEEE Conference on Visual Analytics Science and Technology (VAST)*. pp. 3–12. Paris, France (2014)