

## The Big Data methodology in computer vision systems

Popov S.B.

Samara State Aerospace University,  
Image Processing Systems Institute, Russian Academy of Sciences

**Abstract.** I consider the advantages of using the big data methodology in the computer vision systems. It is noted that this solution provides transparent increasing the functionality of CVS and improvement its quality, the formation of new intellectual properties of the system based on the possibilities of a posteriori multivariate iterative processing of stored video in the background. The basic principles of such intelligent vision systems have been successfully used to create a distributed vision system of railway tanks registration.

**Keywords:** big data, information technologies, computer vision system, image processing, multimodal processing

**Citation:** Popov S.B. The Big Data methodology in computer vision systems. Proceedings of Information Technology and Nanotechnology (ITNT-2015), CEUR Workshop Proceedings, 2015; 1490: 420-425. DOI: 10.18287/1613-0073-2015-1490-420-425

### 1. Introduction

The emergence and development of new approaches and technologies in processing and analysis of big data has led to a shift in the methodology of formation of new knowledge. Some researchers polemically declare the end of science, since the scientific discoveries and insights can come from an exhaustive search of all possible models of any scientific phenomenon with subsequent clustering. It certainly is hyperbole, but there is a definite impact on modern information technologies. Transfer of big data technology to another subject domain does not involve simple copying a selected set of methods. It is important to understand how the basic principles of a new methodology allow to move to a qualitatively new level of target technology. Concrete solutions depend strongly on the specific application domain and its level of hardware and algorithmic advances.

In this work we are primarily interested in the impact of data science approaches to design solutions in the domain of computer vision.

### 2. The Big Data methodology

What is the main difference between the new approaches from the traditional data processing technologies?

New data, computational capabilities, and methods create opportunities and challenges [1]:

- Integrate statistics/machine learning to assess many models and calibrate them against “all” relevant data.
- Integrate data movement, management, workflow, and computation to accelerate data-driven applications.
- New computer facilities enable on-demand computing and high-speed analysis of large quantities of data.

Many models. Traditional research approaches are based on the initial formation of a certain model, then the accumulation and processing of data with subsequent estimation of the parameters of such pre-formed model. In the process of discovery of new knowledge big data technologies integrate simulation, experiment, and informatics. This integrated research solves the problem of finding the model or a set of models most appropriate to some experimental data. Moreover, this methodology is aimed at finding the most interesting (i.e., unexpected and effective) and robust models.

Data-driven applications. The relatively simple principles are at the core of big data processing technology: “Divide and Conquer” – distributed architecture of data storage system and total parallel processing at the lowest level; “Treat where Store” – data are not moved during processing, the data processing tasks are delivered and run on computing resources of distributed storage systems; “Data Are Forever” – data are not deleted or modified during processing, the results simply are saved in situ.

New computer facilities. The widespread introduction of parallelism even in commodity computers discovers new ability for developers of modern algorithms allowing to implement a multi variance and competitive methods of processing. Easy deployment of distributed systems allows transparently to increase intellectual capabilities of target systems.

Seeing computer vision tasks, the most interesting is real-time big data processing. Real-time big data analytics is an iterative process involving multiple tools and systems. It’s helpful to divide the process into five phases [2]: data distillation, model development, validation and deployment, real-time scoring, and model refresh. At each phase, the terms “real time” and “big data” are context dependent.

1. Data distillation – This phase includes extracting data features, combining different data sources, filtering for domains of interest, selecting relevant features and outcomes for modeling, and exporting sets of distilled data to a local data store.

2. Model development – Processes in this phase include feature selection, sampling and aggregation; variable transformation; model estimation; model refinement; and model benchmarking. The goal at this phase is creating a predictive model that is powerful, robust, comprehensible and implementable.

3. Validation and deployment – The goal at this phase is testing the model to make sure that it works in the real world. The validation process involves re-extracting fresh data, running it against the model, and comparing results with outcomes run on data that’s been withheld as a validation set. If the model works, it can be deployed into a production environment.

4. Real-time scoring – At this phase the generated decisions are scoring by consumers or by external control system.

5. Model refresh – Data is always changing, so there needs to be a way to refresh the data and refresh the model built on the original data. The existing programs are used to refresh the models in accordance with the newly accumulated data. It is also recommended simple exploratory data analysis, along with periodic (weekly, daily, or hourly) model refreshes.

Big Data methodology is the iterative process. Models should not be created once, then deployed and left in place unchanged. Instead, through feedback, refinement and redeployment, a model should continually adapt to conditions, allowing both the model and the work behind it to provide value to the organization for as long as the solution is needed [2].

### **3. Revised data processing technology in the computer vision**

Turning to the problems of developing computer vision systems (CVS), it should be noted that the traditional approach of using sequential processing steps performed by the separate frames of video data largely exhausted the resources to increase processing accuracy and the ability to adapt as conditions of surveillance change.

It is necessary to use some principles of big data technology. With regard to the field of computer vision the most suitable solution are the following:

- Total parallelization of the processing and distribution of data.
- Multivariate and competitive data models and methods of their processing.
- Continuous scoring of competitive solutions, both in the process of forming solutions and to adapt to changing surveillance conditions.

The implementation of these solutions leads to the multimodal approach proposed in this work. In this case, this term is broadly interpreted. First of all, the principle of multimodal processing means the simultaneous use of a family of algorithms at the critical stages, each of which is built on a fundamentally different approach. At the same time, multimodality is a coordinated use of video data from different points of view, with different spatial resolutions, a simultaneous use of multiple consecutive frames and additional a priori information about objects of interest.

The results of such versatile processing are analyzed together. This achieves the variability of approaches to solving complex problems, a significant increase in the stability of technology in general with significant changes in observation conditions and parameters of controlled or analyzed objects in video streams.

At a time when for a particular purpose computer vision requires a complex multi-step process, to take a final decision on choosing the best one in a separate step is not possible.

The effectiveness of a multimodal approach is best manifested in computer vision applications in the presence of well-defined criteria for achieving the goal of processing. Since in this case the final decision stage can be implemented as a procedure of joint analysis of the whole set of variants generated by the multivariate multistage process.

A good example of the effective use of multimodal processing is the problem of recognizing the numbers of cars trains. It is noted that similar systems for recognition of train car number should be adapted for work under the following challenges: presence of various digit outlines, variety of color combinations of digits and background, distortions of numbers due to line of sight being non perpendicular to the surface where the number is located (this is especially common for tank cars), various kinds of contamination of the object's surface, and the necessity to operate in both artificial and natural lighting, with the latter accompanied by significant changes in illumination conditions during the day.

The first design decision is the choice of distributed architecture for computer vision system. Distributed modular software allows transparent mapping onto the distributed computation system with a variable number of computers and makes it possible to scale the system towards both a higher quantity of video streams being processed and time reduce of video data processing.

Multimodal processing uses as the base the multistage recognition technology developed for vision system of the railway train registration [3]. Revised technology is as follows: at each stage of processing it is used the previously formed set of algorithms that implement a completely different approaches; each algorithm from the working set is applied to each frame of a video stream of the current wagon. More than one result is obtained at all stages. The results lists are generated. All results are ordered on the basis of relevant metrics. Thus a decision tree, corresponding to one wagon video stream processing, is obtained. When forming the recognition result, code protection of numbering system of train wagons and containers is used and it takes into account the characteristics of validity of recognition formed for each digit. These characteristics are accounted with weights that correspond to algorithms used in previous steps of the process of obtaining a fragment of number. The final scoring is provided by the operator who is responsible for confirmation or adjustment of the final results within the system. It provides additional opportunities for improving the operational quality of recognition technology. With correct result available, it makes sense the posterior multivariate processing of archived video data in the background iterative mode. In this way the quality of generated results may be estimated and parameters of processing algorithms may be selected.

The using structured lighting [4] makes it possible to expand the spectrum of algorithms of fast symbol localization by using the alternative method of delineation of areas with no symbols, which uses the analysis of luminous line geometry of structured lighting for recognition of structural elements of train wagons and containers. Acquired data on the positions of the retaining bands on tank cars and stiffening plates on train cars and containers may be used for the identification of the type of the currently processed train car, tank, or container and therefore obtaining additional information on typical car number locations for them. It is a clear manifestation of the proposed multimodal processing principle.

Allocation of special computational capacity for data storage makes it possible to extend the set of goals of posterior video data processing. This procedure makes it possible to analyze and reveal additional information on the data located in the system, retrieve specific frames of interest for further viewing, interpretation, etc., and

generate an annotated list of detected abnormalities of the observation process [5], ensuring fast access to them for the operator. The indicated opportunities make it possible to use the system for security control. Further analysis of this information allows CVS developers to improve processing and recognition algorithms.

The implementation of this approach involves the development of modifications of the basic algorithms of fast localization of the alphanumeric information in multiple video streams, enabling to form an ordered list of results; the formation of a representative set of recognition algorithms; the implementation of tasks of analysis of the obtained list and selection of the most reliable results.

The distributed architecture of the system ensures further increase in real time video stream processing rate due to scalability of computations [6] within hardware–software platforms of distributed multiprocessor computation systems, in particular using CUDA technology within NVIDIA data processing acceleration hardware.

Distributed video data archive storage on several servers is based on the cascade principle [7]: the working storage with which the video server interacts and the set of archive storages, each one containing video data for a specific period of time. As the working storage is filled, the data are transferred to a first level archive, where space is cleared beforehand by transferring existing data to a second level archive. The same procedure is performed throughout the whole cascade of archives. The most outdated information is deleted at the bottom level.

The presence of the distributed archive storage needed for the implementation of the first three phases of the process of real-time big data analytics and subsequent model refresh.

## 5. Conclusion

The use of new technologies has a cumulative effect. Their implementation opens new opportunities and challenges. The use of distributed software architecture in CVS design considerably facilitates the transfer to advanced technologies based on cloud computing. When complimented by an extended stock list of high resolution IP cameras with possibility for wireless connection to the communication network, it makes it possible to efficiently use three most promising technologies, which are now showing an explosive growth of interest, namely computer vision, cloud computing, and wireless data transfer. Integration of these three technologies significantly expands the CVS application market due to a higher quality of services, a significant reduction in time and complexity and, therefore, implementation costs, and minimization of cost of ownership.

## Acknowledgements

This work was partially supported by the Ministry of education and science of the Russian Federation in the framework of the implementation of the Program of increasing the competitiveness of SSAU among the world's leading scientific and educational centers for 2013-2020 years; by the Russian Foundation for Basic Research grants (# 13-07-00997, # 13-07-12181, # 13-07-97002, # 15-29-07077); by

the Presidium of the RAS program # 6 “Problems of creation of high-performance, distributed and cloud systems and technologies” 2015.

### References

1. **Foster I.** Taming Big Data: Accelerating Discovery via Outsourcing and Automation. Keynote Lecture, the International Winter School on Big Data Tarragona. Spain, January 26-30, 2015.
2. **Barlow M.** Real-Time Big Data Analytics: Emerging Architecture. O'Reilly Media, 2013; 30 p.
3. **Kazanskiy NL, Popov SB.** The distributed vision system of the registration of the railway train. *Computer Optics*, 2012; 36(3): 419-428. [in Russian]
4. **Popov SB.** The intellectual lighting for optical information-measuring systems. *Proc. SPIE 9533, Optical Technologies for Telecommunications 2014, 2015*; 95330 p.
5. **Kazanskiy NL, Popov SB.** Machine Vision System for Singularity Detection in Monitoring the Long Process. *Optical Memory and Neural Networks (Information Optics)*, 2010; 19(1): 23-30.
6. **Kazanskiy NL, Popov SB.** Distributed storage and parallel processing for large-size optical images. *Proc. SPIE 8410, Optical Technologies for Telecommunications 2011, 2012*; 84100I.
7. **Kazanskiy NL, Popov SB.** Integrated Design Technology for Computer Vision Systems in Railway Transportation. *Pattern Recognition and Image Analysis*, 2015; 25(2): 215-219.