Big Data Analytics in Secure Network of Video Surveillance Systems Based on Images Indexation

J.Kavitha[1], Mrs K.Bala[2], Mr V.Pandarinathan[3], Dr D.Rajinigirinath[4]
[1]Dept of CSE
[2,3] Assit.Prof. Dept of CSE,
[4]HOD, Dept of CSE
[1, 2, 3, 4] Sri Muthukumaran Institute Of Technology,Chikkarayapuram, Chennai, Tamil Nadu-600069, India

Abstract- *In recent years, the video data has very massive volume. Digitals cameras are capturing innumerable amounts of video data, the majority of which is unused and analyzed. However, image indexation is necessary to retrieve in time the most significant information. Our proposed framework is based on images indexation to advance a big data analysis in very secure video surveillance system. This framework is based on images indexation. In our work, we have studies the most important methods of indexing in the literature and we have classified them in a taxonomical scenario. Then, we have explains the function of each step in the proposed framework. The strength of this work lies in the fact that it gathers between Big Data analytics, Network security, intelligent video surveillance protection and image and video indexing. All this constitutes a diversified field of research area.*

Keywords- Big Data Analysis, Image indexing, Information retrieval, Video surveillance Sys-tem, Cloud Computing,Public Security.

I. INTRODUCTION

Actually, we live in time of unprecedented amounts of more accessible and varied data than ever before. Especially, video analytics science has scaled with advances in machine vision, and the image analytics remains like as a potential solution to social, economic, industry and political issues. For example, a network of camera IP in video surveillance system capture a large quantity of information, which proves the need for intelligent storage politics (cloud) and very fast and efficient processing algorithms. All studies in literature estimate that big data grows four times faster than the world economy. So, big data storage, treatment and use are becomes the key of any better change of the everyday life. In the form, big data can comprise unstructured and structured data. It can be founded in different high rates of change and in high volumes. In this case, big data storage is the management of data in a scalable way that should satisfies the requirements of applications that involve rapid access to the information. All data acquisition must suffer analysis, and then can be curate in order to be stored in a special way. Generally, the big data storage can be done in different ways, like memory data bases, NoSQL data bases, New SQL data bases, Cloud storage and Query interfaces and by other forms.

II. BIG DATA INDEXINGTEHNIQUES USESD IN VIDEOSURVEILLANCE SYSTEMS

In recent years, real time systems for acquisition, analysis, storage and any other type of access to big data have received a growing attention. Having a massive data pouring into your company is inevitable, store it, analyze it and visualize it in real time are intact different ball game. Actually, companies are moving to real time responses for massive data and various source of information. Responses in real time needs to acquire or to develop decision making tools systems to store, manage and build decision in real time. In recent time, the trend of big data processing is to attack real time systems. In literature, big data is known by tree characteristics, volume, variety and velocity(3Vs). But there is a very important trend of big data treatment, which is defined enclosing the fourth characteristic (4Vs), however big data is ened from many devices such as, PC, GPS, smart phones

2.1 Big Data image and video indexing

Index is the data structure used to organize data records on media storage in order to optimize retrieval access. Indexing big images and videos captured by a video surveillance system are crucial and reel trends in multimedia application. The important goal of video indexing is to develop techniques that can provide the possibility to store and retrieve video in order to access to their contents. Recently, the most important known trends of video surveillance images indexing are: Digital libraries, video on demand and multimedia information systems access etc. Especially, for real time systems, visual database systems require efficient indexing algorithms to simplify fast access to the frames in the database. In the literature, several image and video indexing techniques have been proposed, and we have two most known approaches: a structural method that uses images as hierarchy of objects, regions and portions. The other is based on a

content indexing, in what the content is indexed to encode the frames of the captured video. In the video surveillance process such as detection, tracking, profile analysis or re-identification stage, image database indexing is interest in searching through a very big number of images to find the similar image to the input image, which is known as the query image. Some techniques of image indexation are based on color, texture, sketch, shape, spatial relationship and indexing images in the compressed domain. A video sequence is a series of frames (sequentially ordered in time). The first important operation in video treatment is the storage in a database. The video streaming is segmented to elementary shots in order to be identified and indexed. This figure shows the three

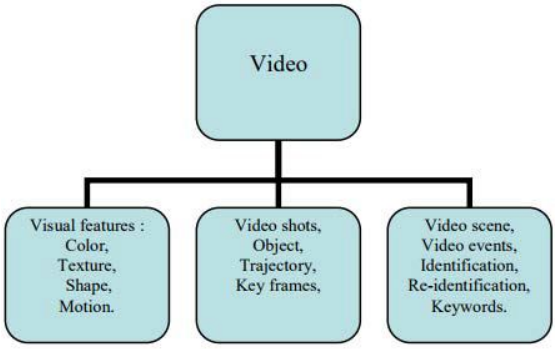

Figure 1: The three boxes of video indexing structure

boxes of structural video indexing. Firstly, we extract the low and mid-level features, then a semantic understanding of the video and feature analyzing is done(B. Meshram and all.).

2.2 Big Data indexation techniques

Big data is a hot topic in the world today and no doubts that querying and indexing with big data is now a great challenge. The most important operation in the big data techniques is the analysis process. In that analysis a statistical models are used and computing technologies [1]. During the analytics phases, the access of particular data and where we operate in real time require a secure and fast access in term of ability to retrieve the information. For doing it, indexing data in big volume data is obligatory. These indexes algorithms allow access to the source in a multitude of ways in the literature; we have many techniques to choose to index the data. This indices or indexes are a list of tags, characters or subjects of a dataset which directed where we can found data.

.

2.2.1 Hidden Markov Model

The HMM representation [2] apply the machine learning to big data that is represented as a sequence of observations over the time. HMM is used in all speech recognition systems, in data compression, in pattern recognition and in some others areas of artificial intelligence. HMM [9] is a tool for representing probability distributions over sequences of observations. For elucidation we consider this example: in a room that is not visible to an observer there is a genie. Urns in the room are: X1, X2, X3 .., in which one of them contains a mix of balls, which are labeledy1, y2, y3 etc. Each genie chooses an urn in that room and arbitrary draw a ball from that urn. Then the ball is puts against to a conveyor belt, where the observer agent watch only the sequence of the balls and not the sequence of the urns from which they were down. With: X : states, y : possible observations, a : state transition probabilities and b : output probabilities.

2.2.2 LSI strategy

The Latent Semantic Indexing adds a very important step to the process of document indexing. LSI [10]transforms the original data in a different space, in which each two words in same concept should have a higher cosine similarity. The LSI strategy uses statistically derived conceptual indices instead of individual words for retrieval. The key idea of LSI strategy is to instead of representing much known document and queries as vectors in a tdim space of terms in two ways : firstly, is by representing them as vectors in a low space dimension whose axes are concepts who works together as a similar words. And secondly, the

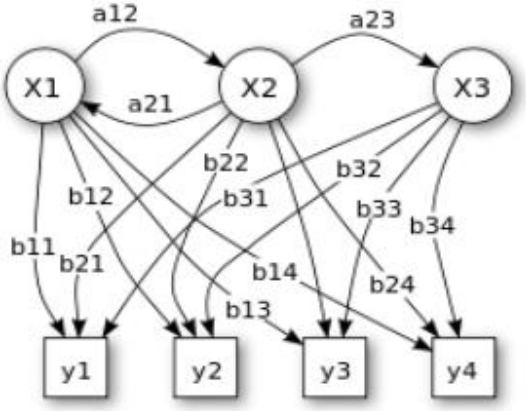

Figure 2: Example of the probabilistic parameter of a hidden

Markov model [3]. axes of the space are the ACP. Practically, the fundamental principle of the LSI strategy takes the vector representation in the original term space and transforms it to new space as shown in this figure [4]: In

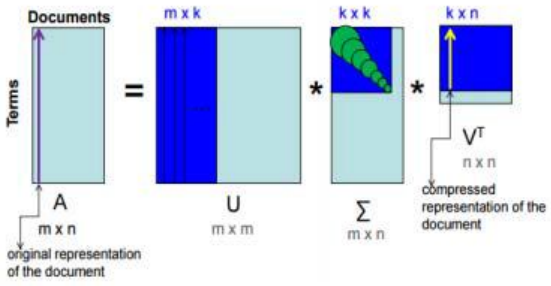

Figure 3: The functional principal of indexing document with LSI key [5]

this procedure, for an m x n matrix A of rank r, we can have the following singular factorization (singular value decomposition):

$$A = U\Sigma VT \quad (1)$$

Where, U is m x m, Σ is m x n and and *VT* is an n x n matrix. Then, the columns of U are orthogonal eigenvectors of *AAT* . in other hand, V columns are orthogonal eigenvectors of *ATA* and eigen values λ1, λ2 λ*r* of *AAT* are the eigen values of *ATA*.

Then:

$$\sigma i = x = \sqrt{\lambda i} \quad (2)$$

σ = diag (λ1.....λi) singular values, Generally, the main idea of the LSI is to map the documents and the queries into a lower dimensional space. Then, the key is that retrieval in this space might be superior to retrieval in the space on the index terms. Studies all indexing techniques demands to write a big book. In our paper we limits to dress the two precedent techniques which are mostly used and we propose in the last subsection a taxonomical organization of the very important algorithms used in the literature.

2.2.3 Hashing indexation

In real time video surveillance systems, indexing massive multimedia data, such as video, image and for any other type of document are named the natural works in the learning by hashing. The main objective of the hashing application is to represent high dimensional data with compact binary codes to obtain simple and faster search results in the time. Many supervised and semi-supervised hashing techniques have been studies for retrieval and searching the relevant information in a real time video surveillance In our classification of different techniques of data indexing, we have underlined that the hashing is very fast and efficient in retrieving. This property candidate this technique to be recommended to real time smart surveillance.

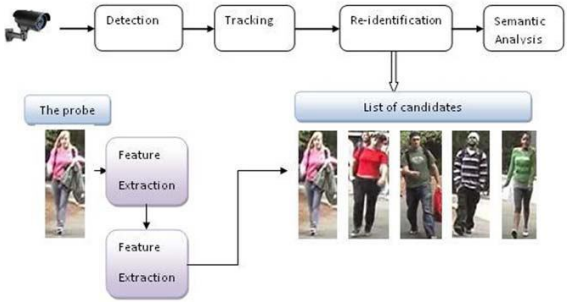

Figure 4: The process of indexing in smart surveillance system.

In our classification of different techniques of data indexing, we have underlined that the hashing is very fast and efficient in retrieving. This property candidate this technique to be recommended to real time smart surveillance.

2.2.4 Taxonomy of the popular indexing strategies used in Big Data

The following table organizes the most known strategies if big data indexing algorithms: This table shows some of the very important indexing techniques used in big data. We can classify them into three categories: NAI Non-Artificial Intelligence, AI Artificial Intelligence and CAI collaborative Artificial Intelligence. In the first class NAI

Indexing Strategies	Properties	Challenges
B-tree	- One dimensional access method -Tree structure with nodes and pointers - Scales linearly	- Waste storage space - Not suitable for multidimensional access -Consumes huge computing resources
R-tree	- More scalable than the B-tree - 2 to 3 dimensional access method	- Index consumes more memory space
X-tree	- Multidimensional access method	- Consumes memory space
Hash	- Presents the exact answer (uses '=' operator) - Quick information retrieval	- Computational overhead
GiST	- Arbitrary indexes - Based on the tree structures	- Slower query response
GIN	- Arbitrary indexes - Based on the tree structures	- Longer processing time
Inverted	- Index consumes less space - Full text search (keyword search)	- Longer data processing time - Limits the search space, not necessarily producing the exact answer - Can present wrong answers due to synonyms and polysemy
LSI	- Uses data and meaning of data for indexing - Presents accurate query results (since it uses more information)	- Demands high computational performance - Consumes more memory space
HMM	- Based on the Markov model - Recognizes relationships between data	- Demands high computational performance

Table 1: Taxonomy of the most known indexing techniques [6].

The establishment of indexes doesnt depend on the meaning of the data. B-tree [7], R-tree, X-tree and hashing index are NAI classified NAI techniques. LSI and HMM belongs to AI category. In real time video surveillance systems, all captured videos should be processed for doing investigation quickly. For doing this, all streams are divided to frames which are treated with looping for each pixel or for ROI. In images indexing process, each image in a database, a feature matrix capturing certain properties is stored and processed in a feature base.

The following scheme:

This figure demonstrates the different type of indexing techniques used for skill access to the content of all types of documents. Here we proposed this conception for a great secure smart video protection system. This architecture uses a VPDN [13] tunneling channel based on Layer Two Forwarding or Layer Two Transport Protocol [8].

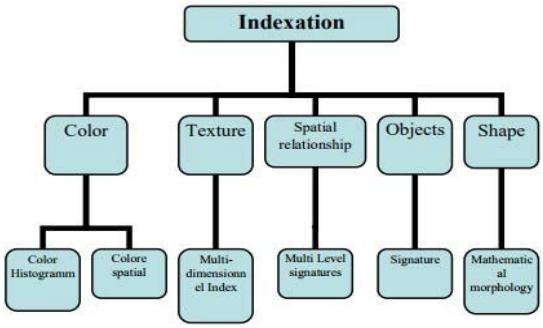

Figure 5: Different content based indexes

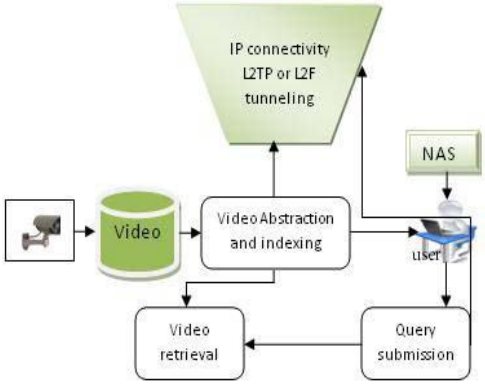

Figure 6: Architecture of secure Video surveillance system indexing and retrieval framework.

III. OVERVIEW OF THE PROPOSED METHODOLOGY

In the last section, Fig. 6 illustrates the structure and the functionality of our proposed framework for Big Data analytics [12] in Secure Video surveillance Systems Based on Indexation Images. Its based on four modules: Pre-processing which is elementary, query processing, retrieval processing and the secure phase. Furthermore, its proposed to works as an efficient and secure and convenient framework to store, process and analyze a massive data captured from the camera. An indexing suitable indexing technique is used to extract only the retrieval information. In the first loop of our framework, pre-processing module is done, in which the surveillance camera capture video streams. Then compressed and stored in a special Medias, in database based on their time and location.

The abstraction process is done in order to decrease computational cost. In the second step, an abstraction is

necessary. A dynamic video skimming and static video summary are done for abstracting all stored streams. The skimming is done from a collection of frames with their related audio. This operation helps to extract the key frames (The representative video frames). To do this extraction, we can use for example the clustering method for performing video abstraction. After abstraction, a video indexing is imposed. This processed can be done into three steps: video analysis, features extraction and then applying the loop of indexation to the Big Data. Object detecting is performed by an adaptive background modeling. The moving objects are classified into persons, cars and others and a feature vector is concluded. The two informations are used in the data indexing. And then a query module is used for retrieving. Finally, the retrieval step is coming and in information transmission process between user, gateway center and network access server in order to address all treatments result via a secure IP voice. This security is based on tunnelingL2F, which itself use the MD5 to hash the connectivity when the access of relevant information is requested by an investigator. And an alarm can be sets off if an abnormal gesture is introduced.

IV. CONCLUSIONS

This research presented a novel and perfect framework for automatic video surveillance indexing and retrieval based on the indexing of the big data which was stored along the time. However, successful implementation of our proposed framework will let the scientific community with a great and leadership scheme of real time Big Data storage analytic system. This system adopts a security strategy based on Virtual Private Dual Network secured by tunneling protocols like L2F or L2TP in IP connectivity for retrieving video when doing query transmission between Gateway, user and NAS. It Indexation Color Texture Objects Signature Shape Color Histogramm Colore spatial Spatial relationship Multidimensionnel Index Multi Level signatures Mathematic al morphology remains to test the feasibility and the credibility of our proposed framework, will be shown by performing a variety of experiments using different Voluminous datasets and an efficient and adaptive indexing techniques to a real time video surveillance system.

REFERENCES

[1] B. B. Meshram and G. P. Gaikwad, Different IndexingTechniques,International Journal of Engineering Research andApplications (IJERA), Vol. 3, Issue 2, March -April 2013,pp.1230-1235.

[2] R.Kune, P. K. Konugurthi , A. Agarwal , R. Rao Chillarigeand R.Buyya, The anatomy of big data computing, Softw.Pract. Exper. 2016; 46:79105.

[3] Z. Ghahramani, In introduction to Hidden Markov Modelsand Bayesian Networks, Journal of Pattern Recognition andArtificial Intelligence, 15(1) : 9:42, 2001.

[4] E. Xing, Lecture : Latent Semantic Indexing (LSI),Unsupervised learning 2011

[5] F. Li, K. Yi, W. Le(2010) Top-k queries on temporaldata. VLDB 19(5):715733

[6] D. Wu ,G. Cong , CS Jensen (2012) A framework forefficient spatial web object retrieval. VLDB J 21(6):797822.

[7] W. Hsu, Ml Lee , Ooi Bc, Pk Mohanty , Kl Teo, Xia C(2002) Advanced database technologies in a diabetic healthcaresystem. Paper presented at the proceedings of the 28th internationalconference on very large data bases, Hong Kong, China

[8] Paul ,B-W Chen , K. Bharanitharan , J-F Wang (2003)Video search and indexing with reinforcement agent for interactivemultimedia services. ACM Trans Embed Comput Syst (2):116.doi:10.1145

[9] J. Wang and W. Liu, Learning to Hash for IndexingBig Data - A Survey , Proceding of IEEE, 17 sep.2015.

[10] H. Xu, N. Yao, W. Hu, H. Pan, and X. Gao, The designand implementation of image information retrieval, inComputer Science Service System (CSSS), 2012 InternationalConference on, pp. 15471550, Aug 2012.

[11] A.Ez-zahout, R. Oulad Haj Thami, and Y. Hadi, Quantitativeand qualitative evaluation based on misclassified pixels forpeoples detection, ICCS12, Novenmber 5-6, Agadir 2012.

[12] E-zahout, Scurit rseaux sous le protocole L2F pourles VPDN, Edition 2016. N du dpt lgal : 2016MO4773, ISBN : .978-9954- 38-997- 5, Maro