

Article

Movement-Oriented Objectified Organization and Retrieval Approach for Heterogeneous GeoVideo Data

Chen Wu ^{1,2}, Qing Zhu ^{2,3,*}, Yeting Zhang ^{2,*}, Xiao Xie ^{4,5,6}, Han Qin ⁷, Yan Zhou ⁸,
Pengcheng Zhang ⁹ and Weijun Yang ⁹

¹ National Engineering Research Center for E-Learning, Central China Normal University, Wuhan 430079, China; wuchen@whu.edu.cn

² State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430072, China

³ Faculty of Geosciences and Environmental Engineering, Southwest Jiaotong University, Chengdu 610031, China

⁴ Laboratory for Environment Computation & Sustainability of Liaoning Province, Institute of Applied Ecology, Chinese Academy of Sciences, Shenyang 110016, China; xiexiao@iae.ac.cn

⁵ School of Geodesy and Geomatics, Wuhan University, Wuhan 430072, China

⁶ Nuclear Industry Huzhou Engineering Investigation Institute, Huzhou 313000, China

⁷ Department of Geography and Geoinformation Science, George Mason University, Fairfax, VA 22030, USA; hqin@gmu.edu

⁸ School of Resource and Environment, University of Electric Science and Technology, Chengdu 611731, China; zhouyan_gis@uestc.edu.cn

⁹ Guangzhou Urban Planning & Design Survey Research Institute, Guangzhou 510060, China; guangzhou2000@126.com (P.Z.); fazeyang@gmail.com (W.Y.)

* Correspondence: zhuq66@263.net (Q.Z.); zhangyeting@whu.edu.cn (Y.Z.)

Received: 20 April 2018; Accepted: 25 June 2018; Published: 28 June 2018



Abstract: With the wide deployment of the video sensor network and the rapid development of video spatialization technology, the large volume of complex GeoVideo data necessitates improvements in the application efficiency of the GeoVideo database and GeoVideo surveillance system. Traditional storage management approaches focus on the optimization of access to the GeoVideo stream. However, they suffer from poor management of the diverse movement processes contained within it; for example, they cannot support associative queries or comprehensive analysis of multi-type GeoVideo data in complex geographic environments. This paper takes physical movement process in GeoVideo as a new type of object and carries out an objectified organization of the heterogeneous GeoVideo data around it (including the video stream, spatial references, interpretations of the video data, and the overall scene) in a Not only SQL-Structured Query Language (NoSQL-SQL) hybrid GeoVideo database. This paper systematically explores the hybrid spatiotemporal indexes and multi-modal retrieval methods around movement processes, which enrich the query modes of the GeoVideo data. A prototype implementation and experimental analysis are presented to prove the feasibility and effectiveness of this organization and retrieval approach.

Keywords: GeoVideo data; movement process; database organization; multi-modal retrieval

1. Introduction

GeoVideo is a type of geographic semantics enhanced video data that makes spatiotemporal semantic associations between the local spatial scene inside the video and the global geographic scene outside the video. It has become an important data source in geographic information systems (GIS)—namely Video GIS [1–6]—providing an associative mechanism for the integrated analysis of

mass video data and large-scale geographic scenes, and playing an important role in smart cities and urban security. With the widespread deployment of the GeoVideo sensor network and continuous technical advancements in video spatialization and video spatial information extraction (such as tracking the position and trajectory of moving objects, reconstructing 3D models, and so on) [7–11], the complexity and volume of GeoVideo data has increased dramatically. This has intensified pressure on integrated data storage and scheduling of GeoVideo databases, challenging the application efficiency of GeoVideo surveillance systems. Therefore, the efficient management of complex GeoVideo data has become a challenging issue in spatial science research.

Traditional GeoVideo databases focus on the management of GeoVideo segments. The storage management approaches of it are mainly divided into four categories: the archived management of GeoVideo segments on a dedicated video server, GeoVideo metadata management on relational databases (RDB), GeoVideo services based on cloud storage, and the GeoVideo pipeline based on main memory databases (MMDB). The storage environment of the archived management of GeoVideo includes magnetic tape, a digital video recorder (DVR), and a network video recorder (NVR); these require a high financial cost for the dedicated video server, storage device, and maintenance, but have limited scalability [12]. GeoVideo metadata includes video text annotation, low-level image features, high-level semantics, and spatial references (such as the camera position, attitude, and parameters, etc.) [13–16], which are usually stored in RDB, while related GeoVideo segments are stored in a file system (FS). The “RDB–FS” hybrid storage architecture is not easy to expand due to the reduced access efficiency of the RDB when a large amount of GeoVideo data is stored. The robust performance of cloud storage in big data environments, such as its high scalability, high performance, low cost, and ease of deployment, has driven the advent of “video surveillance as a service” [17]. However, at present cloud storage is weak in supporting the spatial information of GeoVideo data and mainly uses a light-weight spatial database engine (SDE) or spatial plug-in to implement limited spatial data management functions, or makes a compromise by carrying out dimension reduction processes [18,19]. The GeoVideo pipeline makes full use of the performance advantage of MMDB for real-time access, which enhances real-time analytical performance of online GeoVideo streams [20,21]. These methods improve access efficiency and management capability of GeoVideo stream. However, they suffer from poor management of the diverse movement processes contained within it. It can still be difficult to directly track movements in GeoVideo data.

To overcome the shortcomings of traditional methods, the management approaches of GeoVideo data around movements has become a new research hotspot in recent years. Greater understanding of the validity of massive GeoVideo data is reflected in the observation of movement processes within complex geographical environments [22]. In current research, the storage management approaches of it are classified into three main types: GeoVideo storage management on movement snapshots, GeoVideo storage management on movement carriers, and the GeoVideo storage management on movement processes. The GeoVideo storage management approach on movement snapshots stores GeoVideo stream as a set of key GeoVideo frames describing movement processes [23]. It uses the image processing algorithms (such as morphological filtering, median filtering, and vertical projection) to trigger the storage operation of movement processes by detecting the presence of foreground information, so as to filter out still video frames within a long period of time. By superimposing the temporal information of the movement snapshot set as a label layer on the GeoVideo stream, it can support simple temporal queries of movement processes. GeoVideo storage management approach on movement carriers focus on managing moving states of spatiotemporal objects (such as pedestrian, vehicle, suspect, etc.) in movement processes. According to spatiotemporal cognitive theory, the movement process can be simplified as the sequence of the moving state of a spatiotemporal object. For example, Yattaw et al. summarized 12 types of moving states of spatiotemporal objects by combining three kinds of temporal characteristics (continuous, cyclic, and discrete) with four kinds of geometric structures (point, line, surface, and volume) [24]. Pelekis et al. summarized 8 types of moving states based on a change combination of geometry, topology, and the thematic

attributes of spatiotemporal objects [25]. Xue et al. summarized 7 types of moving states based on the comprehensive analysis of the change characteristics of attributes, functions, and relationships of spatiotemporal objects [26]. The GeoVideo storage management approach on movement processes stores the intrinsic mechanisms of the movement processes. It contains the process semantics (such as the semantics of generation, development, decline, demise, etc.) of each development stage in the life cycle of movement, as well as the temporal relationships and causal relationships between movement processes [27,28]. These methods improve query efficiency and reduce data redundancy in GeoVideo queries and analyses. However, it can still be difficult to support associative queries and comprehensive analyses of multi-period and cross-regional GeoVideo data in a complex geographic environment. The difficulties arise from the implicit storage of the diverse pieces of spatiotemporal correlation information in large-scale movements, the deficiency in rich association relationships between movement processes and multidimensional spatiotemporal data as well as the discrete storage management of heterogeneous GeoVideo data. Therefore, traditional management approaches of GeoVideo data cannot support spatiotemporal associative analyses around movement processes as they evolve over time.

To address above issues, the basic organization of a GeoVideo stream should be changed from being based on a GeoVideo segment to using a complete movement with associated spatiotemporal data. The rich movement information and diverse spatiotemporal data in GeoVideos help people to quickly focus on thematic video content and intuitively understand movements in GeoVideo scenes. It is therefore necessary to focus on the unified organization and storage of valuable GeoVideo streams that track movements as heterogeneous spatiotemporal data. In this paper, we present a novel organization and storage management approach for a massive GeoVideo data orienting spatiotemporal changing process, which takes movements as a new geographic object in the GeoVideo data as a basis for an objectified organization of the heterogeneous data, including the video stream, spatial references, interpretations of the video data, and the geographical scene, etc. The approach summarizes the semantic framework of evolving physical processes in GeoVideo data (movement), abstracts the key elements to dynamically express the formation and development process of the movement, and proposes an integrated management framework for the GeoVideo database. In summary, the GeoVideo database realizes the unified expression, organization, storage, and retrieval of massive and heterogeneous GeoVideo data around movements.

The remainder of the paper is organized as follows. Section 2 describes the objectified organization of the heterogeneous GeoVideo data around multiple movement processes. Section 3 presents the hybrid spatiotemporal index of the GeoVideo data. Section 4 elaborates on the multi-model retrieval method of GeoVideo data. Section 5 shows the implementation of the proposed approach and the experimental results. Finally, Section 6 presents the conclusion of the study.

2. Objectified Organization of Movements in Heterogeneous GeoVideo Data

2.1. Semantic Framework of Movement

GeoVideo data records offer a wealth of resources when it comes to processes evolving over time, including tracking changes in trajectory, structure, and morphology. Considering the diversity of the processes covered, it is necessary to abstract the semantic elements linked specially to movement from GeoVideo segments. The data should encapsulate the relevant heterogeneous data, including the video stream, spatial references, interpretations of the video data, and geographical scenes. The additional data can enrich retrieved GeoVideo data and allow users to focus on analysis of the valuable information captured by video monitoring [6,29].

Semantic data related to the causes of movements, including changes in scope and content, classify a variety of the GeoVideo content. We distinguish the relationship between the spatiotemporal characteristics and the changes in GeoVideo data. Changes in scope, termed the external change factor, stems from the change of the camera's spatial reference (such as the camera's position, attitude, field of

view, etc.) or the camera imaging parameter (such as the resolution, pixel size, frame rate, spectral characteristics, etc.). Changes in content, termed the internal change factor, stems from the dynamic evolution of the surrounding environment (such as the movement of an individual, the replacement of the geographical scene, etc.).

From the perspective of a user's progressive understanding of the GeoVideo change content, movements observed in GeoVideo data can be defined as consisting of three key elements: participant, driving force, and running mode, as illustrated in Figure 1. A participant is any subject that participates in the movement. A participant that changes states throughout the movement is defined as the target participant, such as the active geographic entity. A participant that is only present during the movement (without changing) is defined as the source participant, such as the static geographic scene. The driving force is the impetus that propels the change, including any externally driven forces that bring discrete change states, and internally driven forces that bring about continuous spatiotemporal processes, in the forms of a sequence of changes of state or a natural/humanistic process with thematic attributes as independent variables. The running mode is an abstraction of the execution pattern of the movement, which includes the constraint condition and the operation boundary that determine how the driving force acts on participants in the spatiotemporal process—different movements with the same participants and driving forces can present different results ascribed to different running mode initial conditions and constraints.

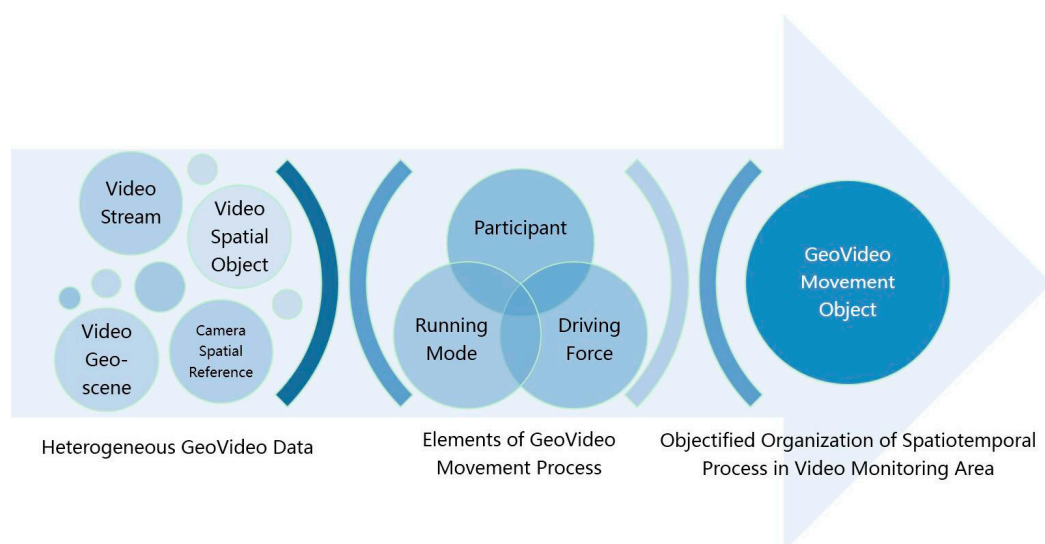


Figure 1. The semantic elements of spatiotemporal GeoVideo processes.

We use three categories of correlations to extract possible relationships in GeoVideo data: temporal correlations, scenario correlations and change-driven correlations. Temporal correlation refers to the chronological order of GeoVideo changes in time. Scenario correlation expresses the spatio-temporal distribution pattern of the GeoVideo changes, including the relationships of direction, scale, and topology. Change-driven correlation means that one change in the GeoVideo has an effect on another change. Change-driven correlation includes four sub-types: evolutionary correlation, causal correlation, trend correlation, and scale correlation. The evolutionary association relationship means that any output of change C1 can be the input of change C2; the causal association relationship means that the change C1 constrains change C2; the trend association relationship signifies that change C1 affects the original process of change C2; the scale association relationship stands for the inclusive relationship between change C1 and C2 in scale, which means that C1 and C2 are expressions of the same movements at different spatiotemporal scales.

The definition of the change semantic abstracts movements, which can then explicitly describe dynamic process of the observed scenario.

2.2. Objectified Organization of Heterogeneous GeoVideo Data

The semantic framework of movements provides the organizational structure for the objectified management of heterogeneous GeoVideo data. By objectifying the heterogeneous data around the spatiotemporal processes in GeoVideos, users can browse the movements they find most significant, but not isolated multi-type data in the form of the video stream, 3D scene model, and trajectory, etc. By abstracting a GeoVideo movement class with features (class attributes) and behaviors (class functions), the GeoVideo movement object (GeoVideo MVO) offers a more intuitive representation. In this paper, we take a single movement as a geographic object to achieve the objectified organization, storage, and management of the multi-type and heterogeneous GeoVideo data.

To achieve a unified expression of heterogeneous GeoVideo data in a spatiotemporal process, the GeoVideo data model describes the common characteristics of GeoVideo data in five dimensions: space, time, theme, scale, and semantics, as illustrated in Figure 2.

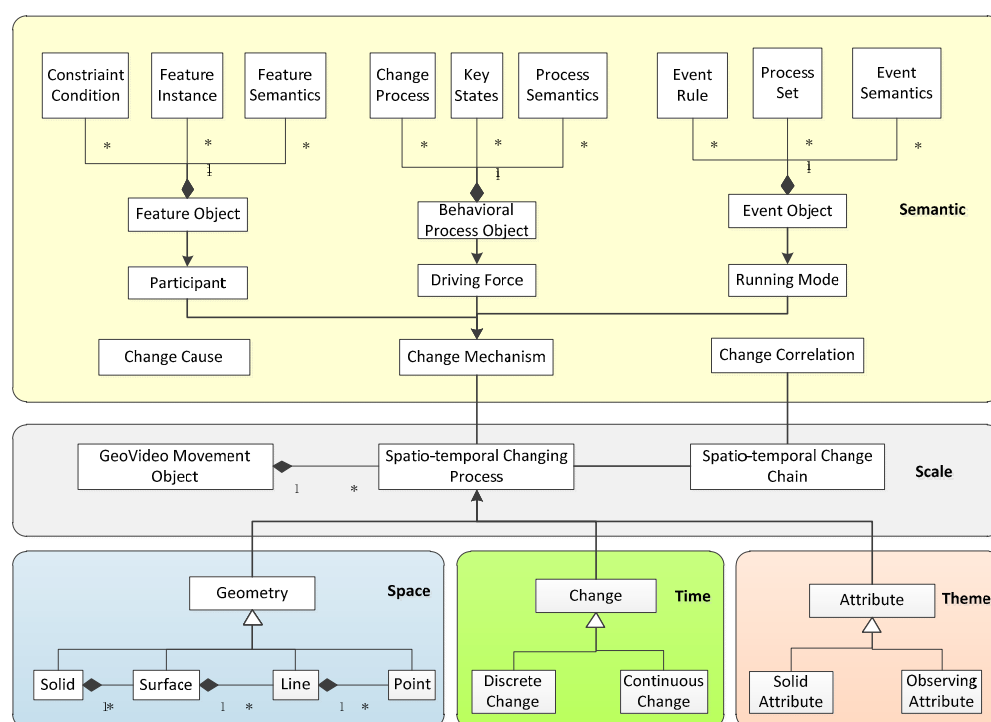


Figure 2. The unified expression of the heterogeneous GeoVideo data.

The spatial dimension expresses the geometric information of the GeoVideo MVO in the video image space and the video geographic scene space, including the video spatial entities in the forms of the camera viewpoint model, the trajectory of the moving object, the video annotation object, and so forth, and video geographical scenes such as complex 3D models in the GeoVideo monitoring area. The uniform geometric representation of the video spatial entity and the video geographical scene is composed of four classes of basic geometric elements—point, line, surface, and volume—which meet the Open Geospatial Consortium (OGC) international standard and support the geometric derivation for interactive visualized analyses in specific video applications.

The temporal dimension expresses the temporal features of GeoVideo MVO from the perspective of the accumulative effect of change during the time interval of the movement, including discrete and continuous change. Discrete change corresponds to the change driven by external force, in the forms of discrete time-sliced objects such as the frames of the initial state, mutation state, and the final state of the spatiotemporal process. Continuous change corresponds to change driven by the internal energy of the entity. The movement is therefore a time-based function.

The theme dimension abstracts the multi-dimensional attributes involved in the GeoVideo MVO, which is divided into the solid attribute and the observed attribute. The solid attribute is a static property independent of the dynamic movement, such as GeoVideo management metadata, and fixed camera location information, etc. The observed attribute is directly obtained from the GeoVideo, including the image feature attributes obtained from a single GeoVideo frame, such as the position of a monitoring object, object type, etc., and the video feature attribute obtained from the video sequence, such as the trajectory of a moving object.

The scale dimension refers to the multi-scale expression of the GeoVideo MVO. In human vision, different observation scales produce different cognition results. The chain of the GeoVideo MVO can comprehensively reflect the evolution process of the GeoVideo monitoring scene at different spatiotemporal scales, which can be described on two levels: horizontal change chain and vertical change chain. The horizontal change chain achieves an orderly collection of different evolution stages of the movement in the same spatiotemporal dimension. The vertical change chain is a multi-detail description of the same process.

The semantic dimension defines the semantic structure of the GeoVideo MVO and is composed of a feature object, a behavioral process object, and an event object [6]. The feature object is the concrete expression of the participant in GeoVideo movement, involving three parts: a constraint condition, a feature instance, and feature semantics. The constraint condition defines the structure type (such as the image, graph, 3D model, or annotated object) and spatiotemporal scale (such as video resolution, or level of detail of 3D model of the feature object). The feature instance is the mapping state of the feature object in the GeoVideo frame under certain constraint conditions. Feature semantics express the intrinsic characteristics and external attributes of the feature object, including the feature parameter semantics that illustrate a unique change property, change condition, and change capability of the feature object; and location semantics that express where the feature object is (based on the absolute coordinate or relative coordinate) within the GeoVideo monitoring scene. The behavioral process object is the concrete expression of the driving force of the movement, involving three parts: the change process, key states, and process semantics. The change process is a description of the procedure by which the movement occurs, including the change runtime environment, life cycle, and process model, in the form of a function or discrete state sequence. The key states record special states of movement, including the input, intermediate result, and output, as well mutant state instances in the discrete spatiotemporal process. The process semantics describe changes in the behavioral process object, including the action semantics, which describe the action types and action characteristics, as well as the trajectory semantics, which describe the process stages and results. The event object is the concrete expression of the running mode of the movement, involving three parts: event rules, process group, and event semantics. Event rules define the rules of judgement and reasons for movement, as well as conditions that trigger the start and end of events. The process group defines the ordered set of behavioral process objects that constitutes the event object—the change chain. The event semantics describe the meaning and characteristics of an event in a specific GeoVideo application and includes the content semantics that describe the event type, the higher-level understanding of the process group, etc., as well as the scene semantics, which describe the geographic scene and time period of the entire spatiotemporal process from the perspective of macro-understanding of the event.

2.3. The NoSQL–SQL Hybrid GeoVideo Database

Based on the objectified organization of heterogeneous GeoVideo data, this section describes the structure of the GeoVideo database, along with the characteristics of the NoSQL–SQL hybrid databases. Compared with SQL (Structured Query Language) databases, NoSQL (Not only SQL) databases have better read/write performance and expansibility and are therefore suitable for dealing with ‘big’ GeoVideo datasets. However, GeoVideo MVO is a typical example of relational data, so a combined utilization of the SQL and NoSQL databases can make use of the performance benefits of both types [30].

The GeoVideo database is organized in a hierarchical structure, from highest to lowest: project, movement class, movement object set, movement element set-multi-granularity GeoVideo stream data. The logical structure of the GeoVideo database is shown in Figure 3. The project table GVDB_ALL_PROJ stores the name and scope of the management area. The GeoVideo monitoring network can be divided into several projects according to the administrative division. For example, GeoVideo data from Beijing, Nanjing, and Wuhan can be managed in different projects. The regional table GVDB_ALL_REGION supports the sub-management units of the city according to the jurisdiction division; for example, Wuhan can be divided into Wuchang district, Hongshan district, and Jiangnan district, etc.

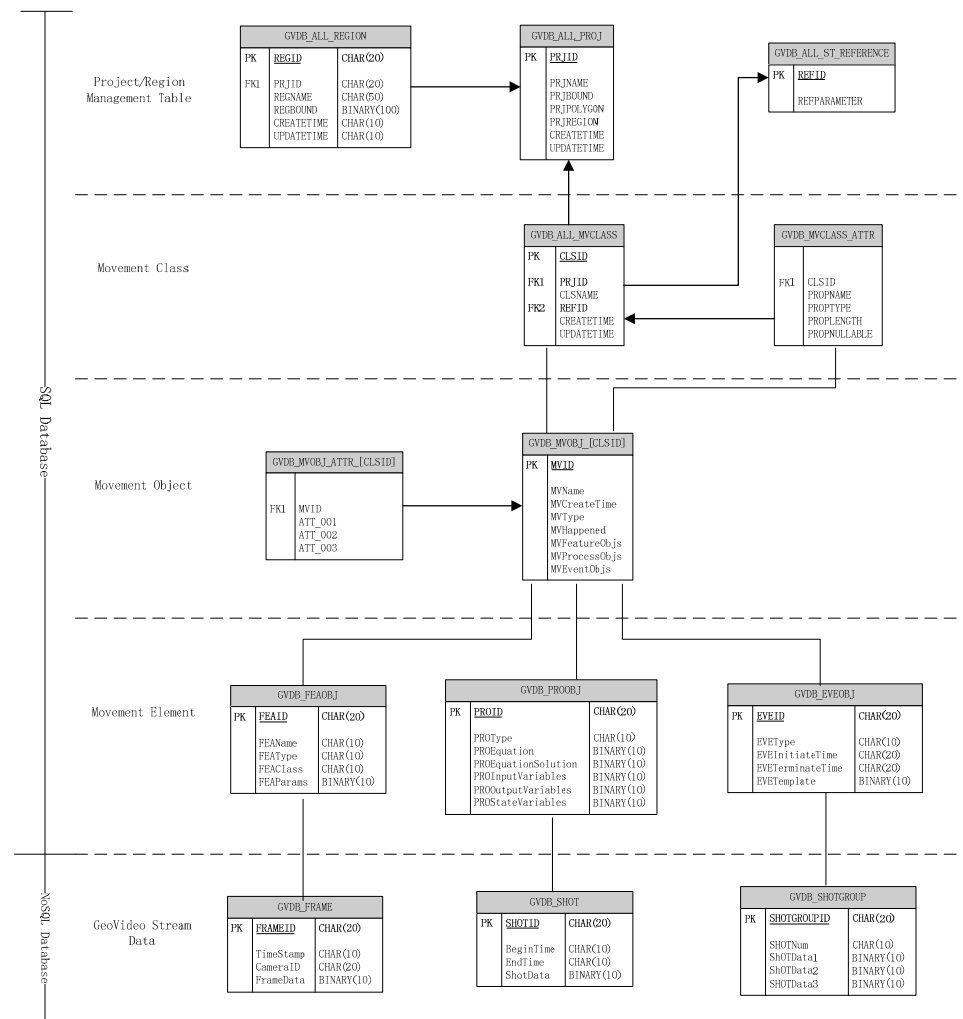


Figure 3. The hierarchical structure of the GeoVideo database.

The GeoVideo database realizes the unified management of the movement objects via movement classes. The movement class is a collection of objects with the same change characteristics and attribute structures. The tables GVDB_ALL_MVCLASS and GVDB_MVCLASS_ATTR store all the management and attribute information for a movement classes. The tables GVDB_MVOBJ[CLSID] and GVDB_MVOBJ_ATTR[CLSID] store certain classes of movement objects and attribute information, while [CLSID] is the unique identifier of the movement class. The feature object, the behavioral process object, and the event object are the key elements for constructing the GeoVideo MVO, and have the corresponding tables GVDB_FEAOBJ, GVDB_PROOBJ, and GVDB_EVEOBJ.

The GeoVideo stream data is the video stream carrier of the movement, with three levels of granularity: the GeoVideo frame, the GeoVideo shot, and the GeoVideo shot group. Multi-granularity GeoVideo stream data objects are the mapping and decimation of the three change elements of a spatiotemporal process on the original video stream, which is beneficial for reducing the video information dimension and improving video comprehension efficiency. The GeoVideo frame object records the discrete key state of the feature object and has a many-to-one aggregation relationship, which means that one feature object may have multiple GeoVideo frames that record the different change states of the participant. The GeoVideo shot object records the spatiotemporal process of the behavioral process object and has a many-to-one aggregation relationship, so one behavior process object may have multi-resolution or multi-view GeoVideo shots for the interpretation of the movement. The GeoVideo shot group object records the whole movement event and has a many-to-one aggregation relationship, which means one event object may have a cross-region and multi-period GeoVideo shot group or a multiple spatiotemporal scale GeoVideo shot group. The GeoVideo frame table GVDB_FRAME, the GeoVideo shot table GVDB_SHOT, and the GeoVideo shot group table GVDB_SHOTGROUP fulfill the storage management of the multi-granularity video data in the collections of NoSQL databases.

3. Hybrid Spatiotemporal Indexes of GeoVideo Data

3.1. A Spatiotemporal Index for GeoVideo MVO

An efficient GeoVideo spatiotemporal index is an important foundation for rapid queries of GeoVideo database. GeoVideo MVO have the features of a discrete distribution, an uneven life cycle of the movement, and different geometric shapes, which are suitable for applying the three dimensional R-tree (3DR-tree) indexing method [31]. The data structure of the non-leaf node and the leaf node in the 3DR-tree are $\langle oid, MBR, t_s, t_e, children[1, \dots, n] \rangle$ and $\langle oid, MBR, t_s, t_e, tuple \rangle$ respectively, where *oid* is the unique identification of the GeoVideo MVO, *MBR* is the minimum bounding rectangle of the spatiotemporal range of the change object, t_s and t_e indicate the effective time interval of the change (start and end times), *children* is the child node pointed to by the non-leaf node, and the *tuple* is the pointer of the entity data associated with the leaf node. The 3DR-tree is intended for offline data whose time interval is fixed, and cannot manage real-time GeoVideo MVO, which reduces the immediacy of the GeoVideo database. In response to this problem, active GeoVideo MVO are updated in real-time in a hash table HASH(AOid, RL) in the main memory for real-time queries, as a complement of the 3DR-tree.

The large-scale GeoVideo MVO may lead to problems of a large, invalid MBR space and node overlapping, which reduce the update/query efficiency of the spatiotemporal index. Large-scale movement can be decomposed into several sub-processes segmented by the field of view (FOV) of the camera, according to the organization of the graph structure. Nodes of the graph express successive sub-processes of the movement object within the scope of a certain camera; paths of the graph express the spatiotemporal semantic topological relationships between sub-processes. The index node of the large-scale GeoVideo MVO can be built on each graph node combined with path information so that it can be easy to find the connected sub-process to constitute the complete the spatiotemporal process.

3.2. Spatiotemporal Index Concerned with Change Semantics

The change semantics (cause, correlation, feature parameter semantics, location semantics, action semantics, trajectory semantics, content semantics, scene semantics etc.) described in the GeoVideo data model in Section 2.2 are important search conditions for querying GeoVideo MVO. For example, the query statement “querying the surveillance video of a pedestrian wearing a red jacket running into the south gate of Wuhan University on 1 August 2017 from 11:00 a.m. to 12:00 a.m.” is a GeoVideo MVO search condition composed of a semantic description of a spatiotemporal range and a movement. The spatiotemporal index of the 3DR-tree mentioned above focuses on the query of the spatiotemporal range of the movement and takes the change semantics as video attributes, which can be inefficient for the query task listed above. The text index can quickly find the GeoVideo data set containing the independent change semantic

words [32], but do not have a good support for ubiquitous spatial information as well as the structural relationship between semantic words associated with the change. Therefore, it is necessary to develop a spatiotemporal index concerned with change semantics to support the intelligent search of GeoVideo data.

The semantic description of movement can be abstracted to the triple structure of subject, predicate, and object. Subject describes the individual characteristics of the participants. Predicate describes the behavioral and trajectory characteristics in the spatiotemporal process. Object describes the geographic location of the relative reference object (geographic entity or static geographic scenario) during the movement. Taking the query statement listed above as an example, the subject is “a pedestrian wearing a red coat”, the predicate is “run into” and the object is “south gate of Wuhan University”. In addition, according to the theory of spatiotemporal cognition, the location information is the spatial carrier of a movement and is the focus of people’s attention. For this reason, the semantic index of movements in GeoVideo data should center on the location semantics. The key steps of the index construction includes the extraction of the location semantics in the GeoVideo MVO; the construction of an R-tree for the location semantics; and the construction of an inverted index for the change semantics of the GeoVideo MVO in the leaf node of the R-tree, which uses the location semantics as the spatial unit to construct the inverted index of the semantic text of the GeoVideo MVO in the spatial range. Different from traditional inverted indexes based on text, the retrieving object in this paper is the GeoVideo MVO and, thus, the vocabulary records all the semantic words of the GeoVideo MVO in the corresponding spatial unit, while the inverted list consists of a series of index items including the unique identification of the GeoVideo MVO, the term frequency of the change semantic word, and the location of the change semantic word (composed of the movement class, the tuple ID of GeoVideo MVO, attribute ID). A B-tree index combined with the location semantic identifier LSID and the start time t_s of the life cycle of the GeoVideo MVO is built to support the temporal query under the location semantic constraints. The index structure is shown in Figure 4.

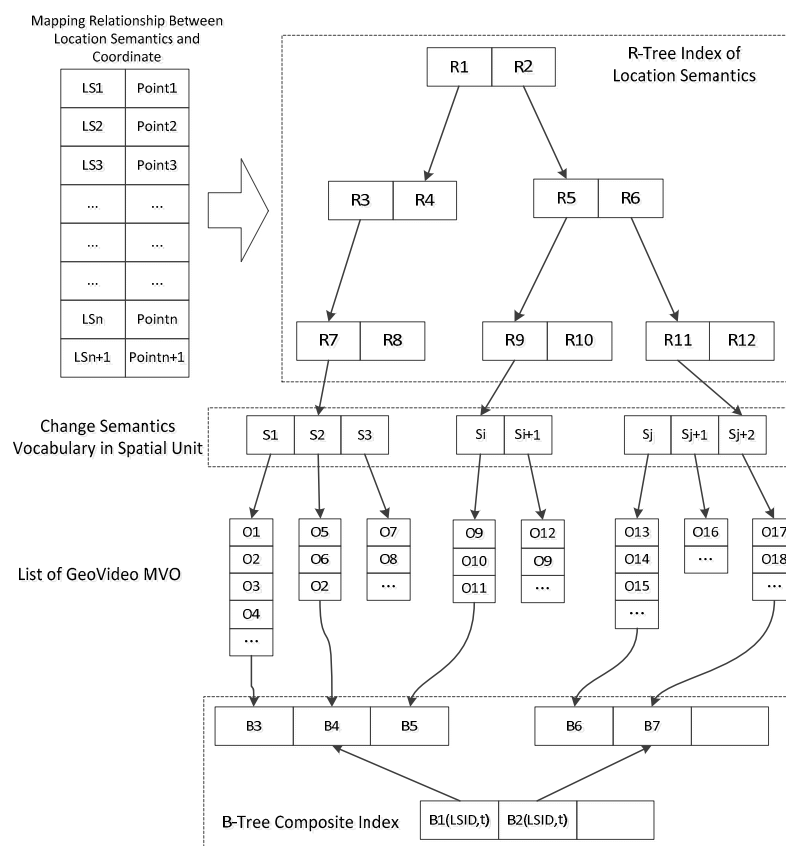


Figure 4. The spatiotemporal index concerned with change semantics.

4. Multi-Model Retrieval Methods of GeoVideo Data

The diversified retrieval methods of GeoVideo data are an important technical means to achieve the efficient application of a GeoVideo database. This section analyzes the retrieval methods of GeoVideo data from the point of view of GeoVideo MVO and movements, which can enrich the retrieval modes of GeoVideo data.

4.1. Retrieval Method of GeoVideo MVO

GeoVideo MVO realizes the objectified organization of the heterogeneous GeoVideo data around movement and frees users from the miscellaneous information of the fragmented GeoVideo stream, the three-dimensional scene, the trajectory of the moving object, etc., by focusing on the behavioral characteristics of the spatiotemporal process in GeoVideo with intuitive change semantics. Being new geographic data, GeoVideo MVO inherits the data characteristics of traditional spatiotemporal data and supports spatiotemporal relationship queries (such as spatial query, topological query, and directional query), but has the particularity of the movement processes. Therefore, we particularly analyzed the retrieval method of GeoVideo MVO from the perspectives of the similarity of change semantics in movements.

Semantic query of movements is an incremental process, with the query condition evolving from fuzzy to clear. We need to perform a semantics similarity measurement between the query condition and the GeoVideo MVO to retrieve the most relevant movement object. The vector space model (VSM) based similarity measurement is the principle method for fuzzy query currently [33] and satisfies the retrieval requirements of GeoVideo MVO change semantics. The VSM converts the semantic description of the GeoVideo MVO and the query condition into a vector representation and compares the vector distance between them. It then judges the degree of correlation. The more similar semantics the two objects contain, the closer the vector distance is and the higher similarity they have. The utilization of the VSM on the change semantics of GeoVideo MVO mainly involves three implementation procedures: establishing a change semantics vocabulary (CSV), calculating the weight of the change semantics term (CST), and calculating the metric space of the change semantics vector between the GeoVideo MVO and the query condition. CSV can be built from the CST of the semantic dimension in the GeoVideo data model, which can be extended with thematic attributes for specific application scenarios. The weight of CST w_{ij} composed of the term frequency (TF) and inverse document frequency (IDF) represents the importance of a CST i in GeoVideo MVO j , which increases proportionally with the occurrences of the CST in the GeoVideo MVO (the parametric expression is tf_{ij}) and decreases inversely to the occurrences of the CST in the GeoVideo database (parametric expression is df_i). idf maps df to a smaller value range through normalization and n is the total number of GeoVideo MVO in the database. The formula for the weight of the CST is as follows.

$$idf_i = \begin{cases} \log\left(\frac{n}{df_i}\right) + 1, & df_i > 0 \\ 1, & df_i = 0 \end{cases}$$

$$w_{ij} = tf_{ij} \times idf_i = \frac{f_{ij}}{m_j} \times \left(\log\left(\frac{n}{df_i}\right) + 1 \right)$$

Through weight calculation, the GeoVideo MVO can be represented as a vector $v_i = \langle w_{i1}, w_{i2}, \dots, w_{im} \rangle$. All the GeoVideo MVO in the thematic database form a matrix: $(w_{ij})_{m \times n}$. A row in the matrix represents one GeoVideo MVO and a column of a row represents one CST. When the user inputs a query condition according to a specific semantic template, the input condition will be converted to a vector $v_q = \langle w_{q1}, w_{q2}, \dots, w_{qm} \rangle$, then the Euclidean distance will be calculated with the matrix and the most relevant calculation result will be chosen as the output data. The specific formula of the similarity calculation is as follows.

$$\text{sim}(v_q, v_i) = \frac{v_q \times v_i}{|v_q| |v_i|}$$

4.2. Retrieval Method of Movement Associated with Geospatial Elements

A movement is a geospatial-context-constrained spatiotemporal process of a physical entity in a GeoVideo monitoring scenario. Integrating the key elements of a movement's "feature object-behavioral process object-event object" with spatial elements such as the "point-line-polygon" in the GeoVideo monitoring scenario allows spatiotemporal correlated semantics to enrich the query mode of complex processes.

A feature object records the key status and position information of the participant at a certain moment in the changing process. The association relationship between the feature object and the geospatial elements of the GeoVideo monitoring scenario can be defined as "stay at the point", "stay by the line", and "stay in the polygon". The behavioral process object records the development process and trajectory information of the movement. The association relationship between the behavioral process object and the geospatial elements can be defined as "point element-pass by", "line element-enter/in/leave/cross/intersect", and "polygon element-enter/in/leave/cross" [34]. The event object is the expression of the running mode during the movement, such as the constraints, operating boundaries, etc., and does not have explicit spatial characteristics. Therefore, this section mainly analyzes the geospatial element correlated model and the query pattern of the first two types of movement elements.

In the spatial dimension, the GeoVideo MVO, the feature object, and the behavioral process object are represented as entity tables in the geospatial element correlated model, namely *MVObject* (*mvid*, *mbr*, *life*), *FEAObject* (*feaid*, *mvid*), and *PROObject* (*proid*, *feaid*, *mvid*), as illustrated in Figure 5. The feature object in the movement may have multiple key state objects whose relational structure is defined as *FEAKeyState* (*ksid*, *feaid*, *time*, *point*). *ksid* is the unique time incremental identifier of the key state object, while *feaid* is the unique identifier of the associative feature object. *time* is the record time of the key state and *point* is the position of the key state object at that time. The behavioral process object can be regarded as a series of movement behavior objects with specific action characteristics whose relational structure is defined as *PROMoveBehavior* (*mbid*, *proid*, *ksid1*, *ksid2*, *mbr*). *mbid* is the unique identifier of the movement behavior object, while *proid* is the identifier of the corresponding behavioral process object. *ksid1* and *ksid2* identify the key state object at the start point and end point of the movement behavior. The movement behavior object is associated with the key state object through *ksid1* and *ksid2*, while both the movement behavior object and the key state object are associated with the GeoVideo MVO via *proid* and *feaid*, respectively.

Based on the above analysis, the association relationship between the elements of movement and the geospatial elements of the GeoVideo monitoring scenario can be mapped as the relationship between the key state object/movement behavior object and the geospatial elements, as illustrated in Table 1. The relational structure of the key state object and the three types of geospatial element are universally expressed as *STAY*(*ksid*, *oid*, *type*), in which *oid* is the identifier of the geospatial elements and the *type* represents the corresponding spatial relationship of "stayAt/stayBy/stayIn". The relational structure of the movement behavior object and the three types of geospatial element are respectively expressed as *MVPT*(*mbid*, *oid*, *time*), *MVLN*(*mbid*, *oid*, *time1*, *time2*, *type*), and *MVPG*(*mbid*, *oid*, *time1*, *time2*, *type*), in which *time1/time2* represent the beginning/ending moment of the movement behavior, while *type* refers to the corresponding spatial relationship discussed above. The model of correlation between the elements of the movement and the geospatial elements is illustrated in Figure 5.

Table 1. The classified correlation mode.

Sub-Element of Movement	Geospatial Element	Correlation Mode
Key state object	point/line/polygon	STAY(<i>ksid</i> , <i>oid</i> , <i>type</i>)
Movement behavior object	point	MVPT(<i>mbid</i> , <i>oid</i> , <i>time</i>)
	line	MVLN(<i>mbid</i> , <i>oid</i> , <i>time1</i> , <i>time2</i> , <i>type</i>)
	polygon	MVPG(<i>mbid</i> , <i>oid</i> , <i>time1</i> , <i>time2</i> , <i>type</i>)

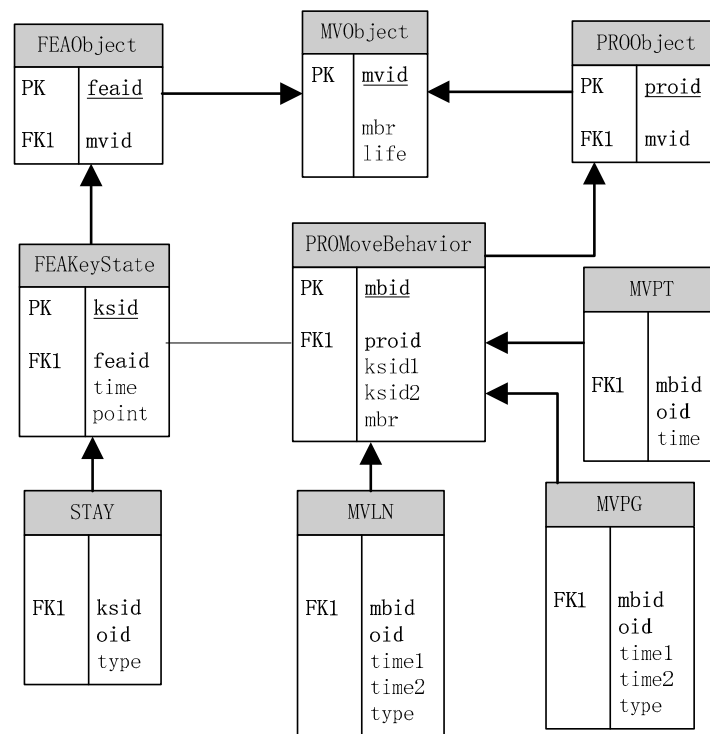


Figure 5. The model of correlation between the movement and geospatial elements.

Based on the relational model, three typical GeoVideo movement queries, including location query, sequential query, and relationship query are presented. These are inefficient or unavailable in a traditional GeoVideo database.

Q1—location query: find all the GeoVideo MVO that enter the camera A at T1, cross the bus station B during T2–T3, and then leaves the camera A at T4.

```
SELECT fea.mvid FROM FEAObject fea, PROObject pro, FEASKeyState ks1, FEASKeyState ks2, PROMoveBehavior mb, STAY st1, STAY st2, MVPT mt WHERE
st1.oid = A AND st1.type = "StayIn" AND ks1.time = T1 AND ks1.ksid = st1.ksid—L1
AND mt.oid = B AND mt.time BETWEEN T2 AND T3 AND mb.mbid = mt.mbid—L2
st2.oid = A AND st2.type = "StayIn" AND ks2.time = T4 AND ks2.ksid = st2.ksid—L3
AND ks1.feaid = ks2.feaid AND ks1.feaid = fea.feaid AND mb.proid = pro.proid AND pro.mvid = fea.mvid—L4
```

L1 returns all the key state objects that come into the FOV of camera A at T1; L2 returns all the movement behavior objects that pass through bus station B in the period of T2–T3; L3 returns all the key state objects that leaves the FOV of camera A at T4; L4 guarantees that the GeoVideo MVO simultaneously satisfies the query condition of L1, L2 and L3.

Q2—sequential query: find all the GeoVideo MVO that appear in the FOV of camera A firstly, then cross square B, and finally pass by office building C.

```
SELECT fea.mvid FROM FEAObject fea, PROObject pro, FEASKeyState ks, PROMoveBehavior mb1, PROMoveBehavior mb2, STAY st, MVPT mt, MVPG mg WHERE
st.oid = A AND st.type = "StayAt" AND ks.ksid = st.ksid AND ks.feaid = fea.feaid—L1
AND mg.oid = B AND mg.type = "Cross" AND mb1.mbid = mg.mbid—L2
AND mt.oid = C AND mb2.mbid = mt.mbid—L3
AND mb1.proid = mb2.proid AND mb1.proid = pro.proid AND pro.mvid=fea.mvid AND ks.ksid < mb1.ksid1
AND mb1.ksid1 < mb2.ksid2—L4
```

L1 returns all the feature objects that appear in the FOV of camera A; L2 returns all the movement behavior objects that pass through square B; L3 returns all the movement behavior objects that pass

by the office building C; L4 guarantees that the GeoVideo MVO satisfies the query condition of L1, L2 and L3 simultaneously, and making sure that L1 occurs before L2 and L2 occurs before L3 by *ksid* comparison.

Q3—relationship query: find all the GeoVideo MVO that pass the FOV of camera A and cross warning line B during the occurrence time of the GeoVideo MVO C1 when passing through camera A.

```
SELECT fea.mvid FROM FEAObject fea, PROObject pro1, PROObject pro2, FEAKeyState ks,
PROMoveBehavior mb1, PROMoveBehavior mb2, STAY st, MVPG mg, MVLN ml WHERE
mg.oid = A AND mg.type = "Cross" AND pro1.mvid = C1 AND mg.mbid = mb1.mbid AND mb1.proid =
pro1.proid—L1
AND ml.oid = B AND ml.type = "Cross" AND ml.mbid = mb2.mbid AND mb2.proid=pro2.proid—L2
AND st.oid = A AND st.type = "StayIn" AND st.time BETWEEN mg.time1 AND mg.time2—L3
AND ks.ksid = st.ksid AND ks.feaid = fea.feaid AND fea.mvid = pro2.mvid—L4
```

L1 returns the movement behavior object and time range of C1 when passing through camera A; L2 returns the movement behavior objects that cross the warning line B; L3 returns the key state objects that cross the FOV of camera A and have a time intersection with the time period of C1 when passing through camera A; L4 guarantees that the GeoVideo MVO satisfies the query condition of L2 and L3 simultaneously.

The retrieval method of the movements associated with geospatial elements takes advantages of the conjunctive query of relational tables in RDB and makes the query of the complex spatiotemporal process more efficient.

5. System Implementation and Experimental Analysis

Being a geographical awareness platform that integrates a video analysis system and a GIS system under a unified geographic reference, Video GIS realizes the management, analysis, and visualization of ‘big’ GeoVideo data and realistic/virtual monitoring scenarios. In this section, we report a prototype of the GeoVideo database based on the mixed storage environment of Redis/MySQL/MongoDB, and present the simulation experiments to verify the effectiveness of the organization and retrieval approach described above.

5.1. Implementation of the Prototype System

The prototype system of the GeoVideo database is composed of three parts: the data repository in hybrid storage environment, the Video GIS database engine, and the GeoVideo database management tool. The data repository implements the data structure of the movement-oriented data model in the hybrid storage environment of Redis/MySQL/MongoDB and realizes the efficient access of the heterogeneous GeoVideo data by applying the application programming interface (API) of these databases in the C++ programming language. The database engine adopting the client/server (C/S) system architecture is composed of a unified access interface module (UAIM), a communication module (CM), a spatiotemporal index module (STIM), a retrieval module (RM), and a database underlying management module (DMM). The client-side implements the database operations of data loading, query, management, and so forth by calling the unified interfaces in UAIM. The server-side utilizes the asynchronous communication mechanism in CM to achieve concurrent accesses of the multiple users and converts the access requests to multi-threaded scheduling tasks of STIM, RM, and DMM to execute efficient concurrent operations in the GeoVideo database. The GeoVideo database management tool provides a visual management interface for the creation, update, deletion operations of the data repository, and spatiotemporal index, etc. Figure 6 shows the visual effect of Video GIS platform based on GeoVideo database, which realizes the unified management of heterogeneous GeoVideo data.

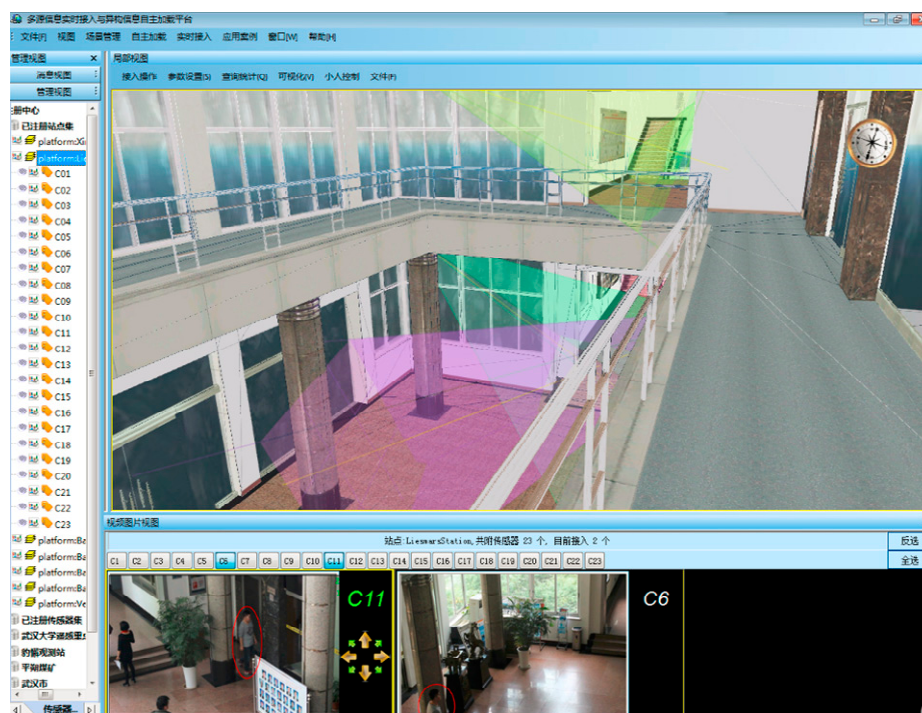


Figure 6. The visual effect of Video graphic information systems (GIS) platform based on GeoVideo database.

5.2. Experimental Analysis

The experimental area of the video surveillance system is 1.3 km², which contains the complex 3D scene (including the 3D model elements of building, road, vegetation, street light, etc.) and 58 distributed cameras (the sampling rate is 25 frames per second and each frame size is about 100 KB). The data volume of a 3D scene is 1.5 GB while the data volume of a GeoVideo stream is about 6 TB. The trajectories extracted from the GeoVideo by utilizing the open-source feature extraction algorithm library OpenCV are more than 5000 and have been mapped to a 3D scene. All experiments were performed on two Dell OPTIPLEX 9020 workstations, each of which created three virtual machines as work nodes, and each node had a 4-Core Intel I7-4790M 3.60 GHz CPU with 8 hardware threads, 4 GB RAM, and a 2 TB disk. The two workstations communicated via a gigabit network. All six nodes used a 64-bit Linux operating system (CentOS Enterprise Server). 64-bit MongoDB, 64-bit Redis, and 64-bit MySQL were used to implement the storage management method for the heterogeneous GeoVideo data.

In order to validate our proposed method, we carried out a series of experiments. The first experiment instantiated the GeoVideo MVO through integrated organization of the heterogeneous GeoVideo data in the experimental area. The second experiment tested the spatiotemporal query efficiency of the GeoVideo MVO. The third experiment compared updating and query performance between different storage repositories in order to check whether our hybrid database outperformed traditional RDB in managing ‘big’ GeoVideo data.

5.2.1. Instantiation of GeoVideo MVO

The GeoVideo MVO data was collected from the spatiotemporal process of the trajectory of moving objects, including pedestrians, vehicles, and others in the video monitoring area. This was done under the modelling rules in Section 2.1. The small-scale GeoVideo MVO in the single camera (SC) were combined with large-scale GeoVideo MVO covering multiple cameras (MC), under diverse spatiotemporal semantic relationships such as the circumstances of moving from the FOV of one camera

into the FOV of an adjacent one. The spatiotemporal indexes were created for each type of GeoVideo MVO respectively. Considering that the spatiotemporal process of a trajectory in an indoor scene supports richer change semantics, we took the indoor GeoVideo MVO named MVIndoorTrajectory as a typical example to illustrate the instantiation process.

The feature object in MVIndoorTrajectory is composed of a person and an indoor scene. The constraint condition of the person can be the extracted contour and high-resolution image while the constraint condition of the indoor scene is a 3D model with a high level of detail. The feature instances of them are key frames recording the principle characteristics. Being the spatial reference for the location semantics and trajectory semantics of the geographic entity, the definition of the indoor scene's feature semantics follow the international standards of OGC CityGML and OGC IndoorGML, which describe the geometric components of an indoor scene using the spatial and functional structures, such as the meeting room, corridor, entrance gate, stair, and exit, etc. The feature semantics of people can be classified as staff member and outsider, according to their identity attributes. Meanwhile, staff members can have additional identity feature attributes attach, such as teacher, student, or support personnel, while outsiders can have additional identity feature attributes appended, such as trainer, visitor, or suspicious individual. For enriching the semantic description of the feature object of a person, the individual feature attributes, location semantics, key frame, and so forth can be supplemented, as illustrated in Table 2.

Table 2. The change semantics of the feature object in MVIndoorTrajectory.

ID	Identity Attribute	Identity Feature Attribute	Individual Feature Attribute	Location Semantics	Key Frame
1	staff member	student	male/young/blue pants	entrance gate	KF-001
2	staff member	student	male/young/blue pants	stair on first floor	KF-002
3	staff member	teacher	male/old/grey shirt	meeting room	KF-003
4	staff member	student	female/young/black pants	entrance gate	KF-004
5	outsider	visitor	male/old/blue shirt	corridor	KF-005/KF-006

The behavioral process object in MVIndoorTrajectory records the movement process of the person, instantiated as the key frames of the process and the trajectory of the person. The change semantics of the behavioral process object in MVIndoorTrajectory is composed of action semantics such as walk towards, run past, run away from, etc. and the trajectory semantics are in the form of “{<geographic entity>-<behavior predicate>-<geographic scene>}”, which describes the procedure schema of the trajectory, as illustrated in Table 3.

Table 3. Change semantics of the behavioral process object in MVIndoorTrajectory.

ID	Feature Object		Behavioral Process Object
	Geographic Entity	Location Semantics	Trajectory Semantics
1	outsider (A)	entrance gate	<A>walk towards<office building>
2	outsider (B)	entrance gate	wander around<entrance gate>
3	outsider (A)	reception lobby	<A> walk from <front entrance> into <reception lobby>
4	staff member (C)	corridor	<C> run past <corridor> into <stair>
5	outsider (A)	entrance gate	<A> ran away from <office building>

The event object in MVIndoorTrajectory records trigger conditions, restrictions, and an ordered set of behavioral process objects of the movement. Appearance in non-working time, crossing warning line, fire hazard and so forth are all trigger conditions of an event object and are stored in the event template library, while the abnormal shutdown of a channel or the activation of emergency access will restrict the route selection. The change semantics of the event object in MVIndoorTrajectory intuitively describes someone's understanding of the trajectory process from the perspective of a thematic application. For example, the trajectory process of “<A> walks from <front entrance> into <reception lobby>” belongs to “implementation phase of theft incident”, and the trajectory process of “<A> ran away from <office building>” belongs to “escape phase of theft incident”, and so forth. Meanwhile, the

spatiotemporal information of the trajectory process can be abstractly described as “noon/campus” or “evening/street”, that is, depicting the trajectory process from a macroscopical level.

In this experiment, we took the 3D model of an office building as the spatial framework, captured the indoor space activities of persons in the video monitoring network, and constructed the MVIndoorTrajectory of 113 trajectory processes. The visualization of GeoVideo MVO on Video GIS platform is illustrated in Figure 7.

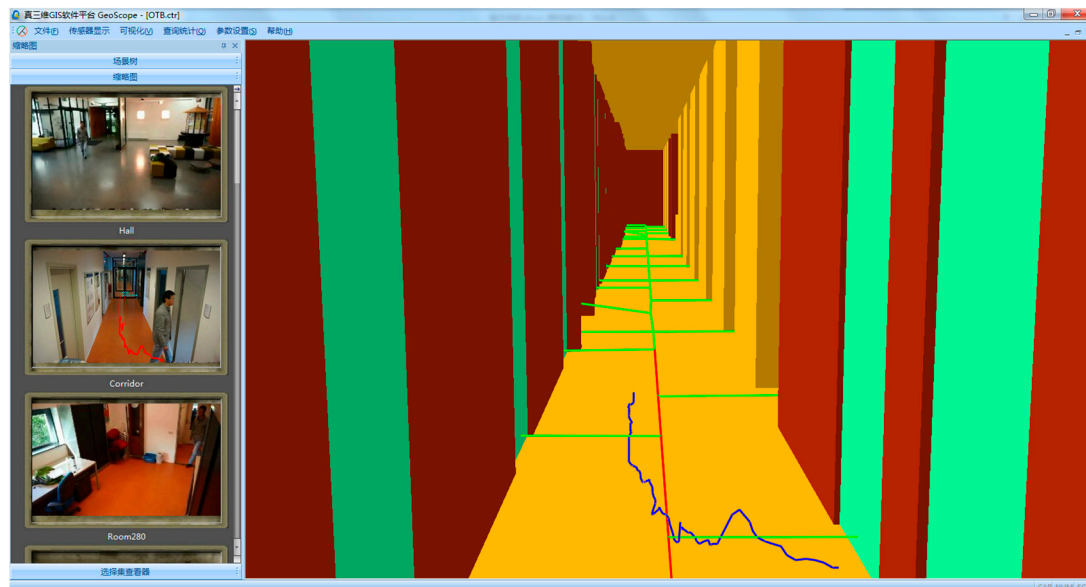


Figure 7. The visualization of GeoVideo movement object (MVO) on Video GIS platform.

5.2.2. Spatiotemporal Query of GeoVideo MVO

In order to verify the query efficiency of the GeoVideo MVO in large scale, a test was performed by making a spatiotemporal range query roaming along a fixed path in the video geographic scene. The length of the specified time range $[t_s, t_e]$ of the query was set to 30 min (which can have a temporal intersection with those MVOs having a longer life cycle) and the spatial range of the two-dimensional bounding box was set from 200×200 m to 500×500 m. The Video GIS platform sent query requests of the spatiotemporal range to the GeoVideo database and calculated the average query time, including the index search time, the scheduling time of the GeoVideo MVO in MySQL, and the transmission time of the multi-granularity video data in MongoDB, over multiple tests. The test results of the query time and the query time percentage spent at each stage are shown in Figures 8 and 9.

As shown in Figure 9, the index search time and the scheduling time of the GeoVideo MVO in MySQL accounts for less than 0.01% and 11% of the entire query time respectively, while the transmission of the multi-granularity video data is time-consuming. According to human cognitive theory, the composition structure of the query time satisfies the query requirements of the quick response of GeoVideo MVO in real-time. Meanwhile, compared with traditional grid-based 3D model management methods, the movement-based organization method in this paper organizes 3D models by trajectories, which greatly reduces the pressures of the scheduling, transmission, and drawing of invalid 3D models. As shown in Figure 10, the grid-based 3D model management method (left one) schedules all the 3D models in the grid range that cover the spatiotemporal range of trajectories, but the movement-based 3D management method (right one) only schedules the 3D models related to the movement processes.

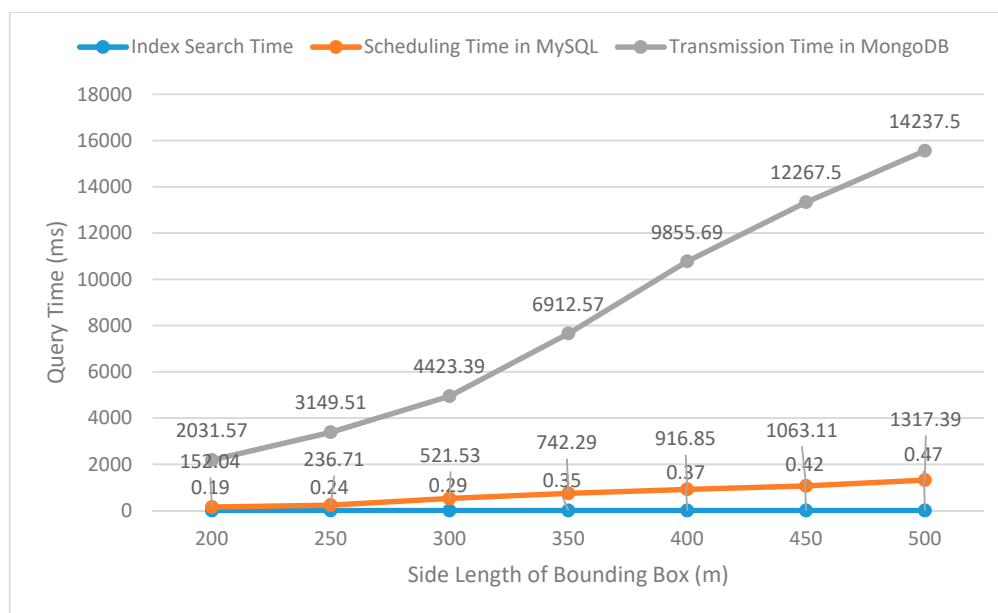


Figure 8. The query times for GeoVideo MVO at different stages.

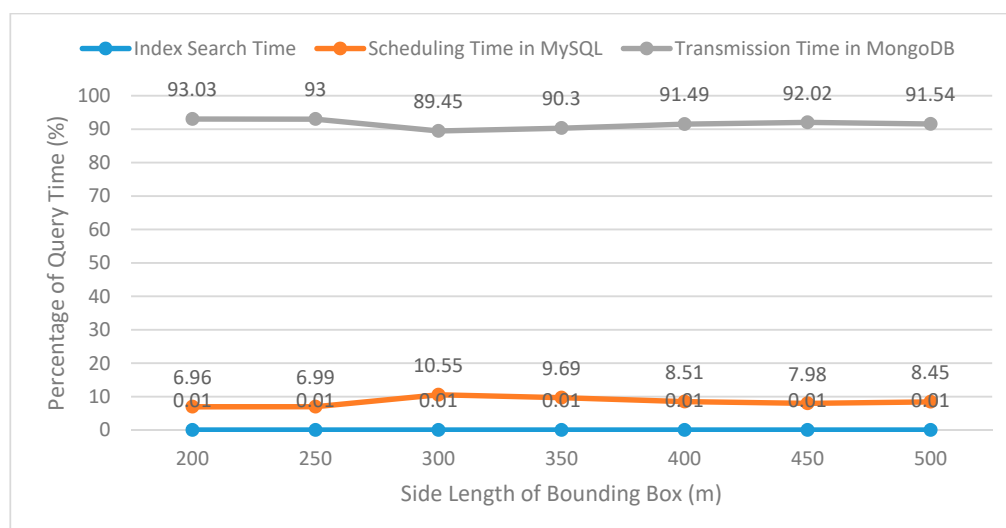


Figure 9. The query time percentages of the GeoVideo MVO at different stages.



Figure 10. The comparison between the grid-based (on the left) and the movement-based (on the right) 3D model management method.

We then took the CST following the data structure listed—defined before as a key word—constructed the CSV (as well as the corresponding change semantics inverted index of MVIndoorTrajectory), and integrated the functions of the change semantics keyword query and similarity query of the change semantics of MVIndoorTrajectory in the GeoVideo database management tool. The query interface and query result of the GeoVideo MVO by change semantics are shown in Figure 11.

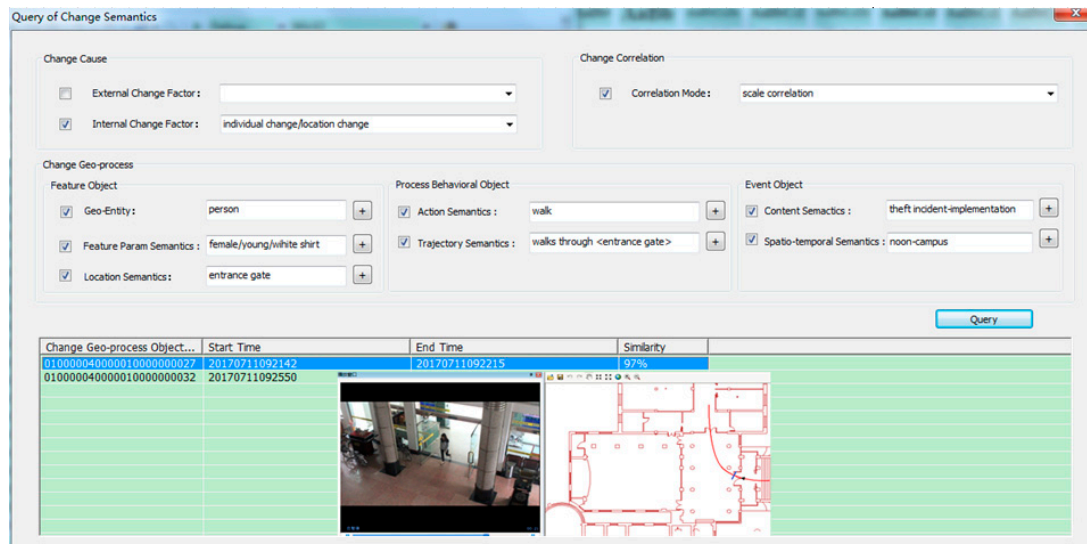


Figure 11. The query interface and query result of the GeoVideo MVO by change semantics.

5.2.3. Access Performance of GeoVideo Database

To demonstrate the access performance of the GeoVideo database under ‘big’ data, we compared updating and query performance of GeoVideo stream (data granularity is video frame) between Redis-MongoDB-based repository, MongoDB-based repository and MySQL-based repository, in order to check whether our NoSQL-SQL hybrid GeoVideo database outperformed independent representative NoSQL and SQL databases in accessing operations [30].

From Figure 12, we can see that the access efficiency of the NoSQL-SQL hybrid repository was significantly better than the other two, and was stable under big database volume. Meanwhile, the MongoDB cluster Mongos guaranteed an even distribution of GeoVideo data under multiple data nodes and successfully supported the total number of a 6TB GeoVideo stream. Therefore, the NoSQL-SQL hybrid GeoVideo database can efficiently manage massive GeoVideo data.

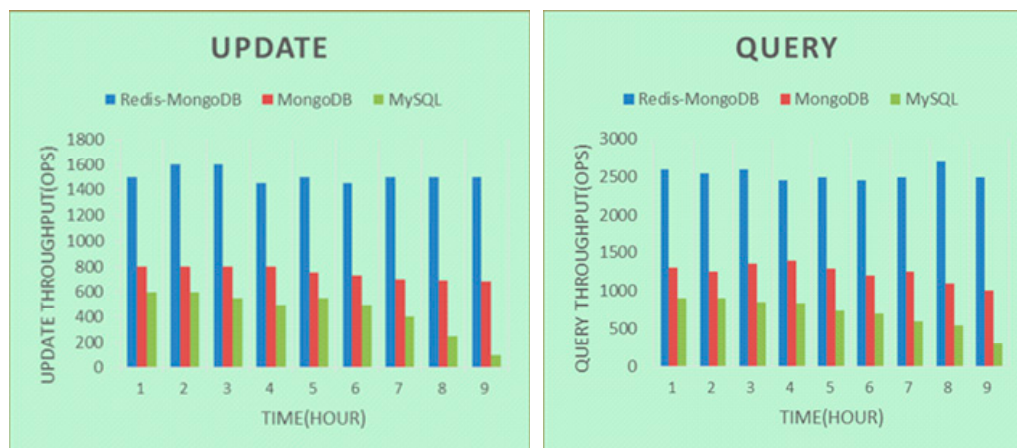


Figure 12. The access performance of the GeoVideo database.

6. Conclusions

The wide use of the GeoVideo sensor network and the rapid increase of the volume of GeoVideo data brings significant challenges to traditional GeoVideo organization methods. Understanding the value of the GeoVideo big data relies on the changes presented in the GeoVideo monitoring scenarios. This paper treats movements (processes of change within GeoVideo data) as the basic organizational object. Around the GeoVideo MVO, this paper systematically studies the objectified organization strategy of heterogeneous GeoVideo data; explores the key technologies of the NoSQL–SQL hybrid GeoVideo database model, the hybrid spatiotemporal indexes, and the multi-modal retrieval methods; and verifies the feasibility and effectiveness of the approach by a performance analysis and applications test.

Author Contributions: As the first author, C.W. wrote the first draft of this article and made the key experiments. Q.Z. and Y.Z. (Yeting Zhang) contributed to the concept and framework of the organization and retrieval approach. X.X. contributed to the concept of change semantics and idea of the experiment. H.Q. updated the manuscript and polished the language. Y.Z. (Yan Zhou) contributed a lot to the modification of change semantics description and updated the manuscript. P.Z. and W.Y. updated the manuscript and contributed to the idea of the experiments.

Funding: This paper was supported by the National Natural Science Foundation of China [Grant No. 41571392, 41471320, 41471332, 41701466], and the Smart Guangzhou Spatio-temporal Information Cloud Platform Construction (Grant No. GZIT2016-A5-147).

Acknowledgments: The authors would like to thank the anonymous reviewers for their constructive comments and suggestions.

Conflicts of Interest: The authors declare no conflicts of interest.

References

- Lewis, P.; Fotheringham, S.; Winstanley, A. Spatial video and GIS. *Int. J. Geogr. Inf. Sci.* **2011**, *25*, 697–716. [[CrossRef](#)]
- Lu, Y.; Shahabi, C.; Kim, S.H. Efficient indexing and retrieval of large-scale geo-tagged video databases. *Geoinformatica* **2016**, *20*, 829–857. [[CrossRef](#)]
- Milosavljević, A.; Rančić, D.; Dimitrijević, A.; Predić, B.; Mihajlović, V. Integration of GIS and video surveillance. *Int. J. Geogr. Inf. Sci.* **2016**, *30*, 2089–2107. [[CrossRef](#)]
- Kong, Y. Design of GeoVideo data model and implementation of Web-Based VideoGIS. *Geom. Inf. Sci. Wuhan Univ.* **2010**, *35*, 133–137.
- Xie, Y.; Wang, M.; Liu, X.; Wu, Y. Integration of GIS and moving objects in surveillance video. *ISPRS Int. J. Geo-Inf.* **2017**, *6*, 94. [[CrossRef](#)]
- Xie, X.; Zhu, Q.; Zhang, Y.; Zhou, Y.; Xu, W.; Wu, C. Hierarchical semantic model of Geovideo. *Acta Geod. Cartogr. Sin.* **2015**, *44*, 555–562.
- Wang, Z.; Wang, M.; Liu, X.; Zhang, Y. Deployment optimization for camera network coupling with geographic environment. *Ann. GIS* **2018**, *24*, 9–17. [[CrossRef](#)]
- Milosavljević, A.; Rančić, D.; Dimitrijević, A.; Predić, B.; Mihajlović, V. A method for estimating surveillance video georeferences. *ISPRS Int. J. Geo-Inf.* **2017**, *6*, 211. [[CrossRef](#)]
- Wu, Y.; Wang, M. Video spatialization method and its application. *J. Nanjing Normal Univ. (Nat. Sci. Ed.)* **2014**, *37*, 119–125.
- Sakaino, H. Video-based tracking, learning, and recognition method for multiple moving objects. *IEEE Trans. Circuits Syst. Video* **2013**, *23*, 1661–1674. [[CrossRef](#)]
- Sourimant, G.; Morin, L.; Bouatouch, K. GPS, GIS and video registration for building reconstruction. In Proceedings of the IEEE International Conference on Image Processing, San Antonio, TX, USA, 16 September–19 October 2007; Volume 6, pp. 401–404.
- Lougher, P.; Pegler, D.; Shepherd, D. Scalable storage servers for digital audio and video. In Proceedings of the International Conference on Storage and Recording Systems, Keele, UK, 5–7 April 1994; pp. 140–143.
- Ma, H.; Ay, S.; Zimmermann, R.; Kim, S. Large-scale geo-tagged video indexing and queries. *Geoinformatica* **2014**, *18*, 671–697. [[CrossRef](#)]

14. Xu, Z.; Liu, Y.; Mei, L.; Hu, C.; Chen, L. Semantic based representing and organizing surveillance big data using video structural description technology. *J. Syst. Softw.* **2015**, *102*, 217–225. [[CrossRef](#)]
15. Wang, Y.; Zhou, L.; Xing, C. Video semantic models and their evaluation criteria. *Chin. J. Comput.* **2007**, *30*, 337–351.
16. Kosch, H. Mpeg-7 and multimedia database systems. *ACM SIGMOD Rec.* **2002**, *31*, 34–39. [[CrossRef](#)]
17. Rodriguez-Silva, A.; Adkinson-Orellana, L.; Armino-Franco, I. Video surveillance based on cloud storage. In Proceedings of the IEEE International Conference on Cloud Computing, Honolulu, HI, USA, 24–29 June 2012; pp. 991–992.
18. Xiang, L.; Huang, J.; Shao, X.; Wang, D. A mongodb-based management of planar spatial data with a flattened R-tree. *ISPRS Int. J. Geo-Inf.* **2016**, *5*, 119. [[CrossRef](#)]
19. kumar, A. Mean-variance analysis of the performance of spatial ordering methods. *Int. J. Geogr. Inf. Sci.* **1998**, *12*, 269–289.
20. Kemper, A.; Neumann, T. Hyper: A hybrid OLTP&OLAP main memory database system based on virtual memory snapshots. In Proceedings of the 2011 IEEE 27th International Conference on Data Engineering (ICDE), Hannover, Germany, 11–16 April 2011; pp. 195–206.
21. Garciamolina, H.; Salem, K. Main memory database systems: An overview. *IEEE Trans. Knowl. Data Eng.* **1992**, *4*, 509–516. [[CrossRef](#)]
22. Hornsby, K.; Egenhofer, M. Identity-based change: A foundation for spatio-temporal knowledge representation. *Int. J. Geogr. Inf. Sci.* **2000**, *3*, 207–224. [[CrossRef](#)]
23. Tung, W.; Lee, T. Rank-mediated collaborative tagging recommendation service using video-tag relationship prediction. *Inf. Syst. Front.* **2013**, *4*, 627–635. [[CrossRef](#)]
24. Yattaw, N. Conceptualizing space and time: A classification of geographic movement. *Cartogr. Geogr. Inf. Sci.* **1999**, *2*, 85–98. [[CrossRef](#)]
25. Pelekis, N.; Theodoulidis, B.; Kopanakis, I. Literature review of spatio-temporal database models. *Knowl. Eng. Rev.* **2004**, *3*, 235–274. [[CrossRef](#)]
26. Xue, C.; Zhou, C.; Fenzhen, S.; Dong, Q. Research on process-oriented spatio-temporal data model. *Acta Geod. Cartogr. Sin.* **2010**, *1*, 95–101.
27. Pultar, E.; Cova, T.; Yuan, M. EDGIS: A dynamic GIS based on space time points. *Int. J. Geogr. Inf. Sci.* **2010**, *3*, 329–346. [[CrossRef](#)]
28. Xie, J.; Xue, C.; Zhang, F. Process-oriented TGIS semantics and has representation framework. *Geogr. Geo-Inf. Sci.* **2011**, *4*, 1–7.
29. Xu, W.; Zhu, Q.; Zhang, Y.; Ding, Y.; Hu, M. Real-Time GIS and its application in indoor fire disaster. *ISPRS Arch.* **2013**, *XL-2/W2*, 121–127. [[CrossRef](#)]
30. Wu, C.; Zhu, Q.; Zhang, Y.; Du, Z.; Ye, X.; Qin, H.; Zhou, Y. A NOSQL–SQL hybrid organization and management approach for real-time geospatial data: A case study of public security video surveillance. *ISPRS Int. J. Geo-Inf.* **2017**, *6*, 21. [[CrossRef](#)]
31. Zhu, Q.; Gong, J.; Zhang, Y. An efficient 3D R-tree spatial index method for virtual geographic environments. *ISPRS J. Photogramm. Remote Sens.* **2007**, *62*, 217–224. [[CrossRef](#)]
32. Melnik, S.; Raghavan, S.; Yang, B.; Garcia-Molina, H. Building a distributed full-text index for the Web. In Proceedings of the 10th International Conference on World Wide Web, Hong Kong, China, 1–5 May 2001; Volume 19, pp. 396–406.
33. Guo, Q. The similarity computing of documents based on VSM. In Proceedings of the 2008 32nd Annual IEEE International Computer Software and Applications Conference, Turku, Finland, 28 July–1 August 2008; pp. 585–586.
34. Xiang, L.; Wu, T.; Gong, J. A geo-spatial information oriented trajectory model and spatio-temporal pattern querying. *Acta Geod. Cartogr. Sin.* **2014**, *43*, 982–988.

