A METADATA-BASED APPROACH FOR USING CCTV TO FACILITATE
EMERGENCY RESPONSE

J-H. Hong\textsuperscript{1}\textsuperscript{*}, Y-T. Shi\textsuperscript{2}

\textsuperscript{1} Department of Geomatics, National Cheng Kung University - junghong@mail.ncku.edu.tw
\textsuperscript{2} Department of Geomatics, National Cheng Kung University - p68021021@gs.ncku.edu.tw

Commission IV

KEY WORDS: Emergency response; Metadata; CCTV Monitor; FOV Viewshed; Data Sharing

ABSTRACT:

The installation of closed-circuit television monitors (CCTV) has been rapidly increasing ever since the September 11 attacks and has become one of the most widely used types of sensors for the tasks that require instantaneous and long-term monitoring. With the distinguished characteristics of direct visual inspection of the subject of interests, the availability of CCTVs offers the EOC commanders a quick way to validate and access the reported disaster incidents during emergency response. However, the heterogeneity of CCTV systems and the lack of appropriate descriptions precludes the optimized use of CCTV and causes immense difficulties in effectively coordinating or appraising the use of CCTV systems. By arguing standardized metadata plays the most crucial role in quickly and precisely finding the needed CCTV during emergency response, regardless of its specifications, suppliers, or locations, we propose an integrated operational framework based on the CCTV metadata specifically designed for emergency response in this paper. This framework standardizes the metadata to unify the descriptions for heterogeneous CCTV systems and fulfill the requirements for searching and selecting CCTV. Instead of the 2D point-based location or sector-based CCTV FOV coverage, we also propose to extend the FOV from 2D to 3D to precisely describe the area and even the features that can be monitored by the CCTV. This not only improves the precision and efficiency of selecting CCTV but also increase the commander’s ability to make quick and accurate responses to disasters. The advantages of integrating heterogeneous CCTV systems can tremendously improve the possibility of monitoring and updating the real-time status caused by hazards.

1. INTRODUCTION AND RELATED WORKS

One of the purposes of establishing standard-based systems is to facilitate a barrier-free data circulation and exchange platform (Panta et al., 2020). Since data produced by various business units are often different, cross-unit and cross-field circulation have to specifically address the interoperability issues caused by the various types of heterogeneity among different systems. In the domain of GISs, the International standards organization (ISO/TC 211) and the Open GIS Consortium (OGC) have been committed to the formulation of technical standards and specifications for almost three decades (Wang & Yuan, 2010). With the standardization frameworks from ISO/TC211 and OGC, data and service are developed and distributed in a transparent way via consensus standards and specifications (Khan et al., 2020). For example, the standards of the Sensor Web aim to facilitate an open framework for linking IoT devices deployed at different places and share their collected observations in a real-time fashion. Without standards, communication between different IoT systems or services is surely not an easy task. For example, to uniquely identify a particular sensor device from heterogeneous IoT platforms is difficult due to the varying design of described resources (such as device identifiers) between platforms (Koo et al., 2019). An ideal scenario is the heterogeneous types of sensors scattered at different locations are smartly organized as an effective internet-based operating mechanism with built-in knowledge, so as to improve the quality of human lives.

CCTV-based monitoring systems have become increasingly popular in recent years. By deploying CCTVs at the chosen locations, operators can continuously collect visual images and monitor reality changes to support their application decisions, e.g., the traffic flow of a busy street, speeding photography, entrance monitors. CCTV systems not only enable the real-time monitoring of phenomena of interests but also help to rebuild the situations in the past, provided historical data is recorded and made available. With the overwhelming popularity of CCTV systems (Yeganeg et al., 2020), it is extremely advantageous if all available CCTVs can be effectively managed and accessed when an emergency situation occurs (Gene et al., 2013). Due to individual CCTVs that may have their own specifications, connection methods, and processing procedures, to effectively integrate heterogeneous CCTV systems must take “standardization” into consideration. The metadata design in this research adopts the 5WH perspectives, namely, where, when, what, why, who and how, for metadata design, such that the essential information required for selecting CCTV during emergency response can be promptly included. Each CCTV, regardless of its specification or responsible agencies, is required to create standardized metadata to improve the interoperable usage of CCTV information (Figure 1). The standardization enables the commanders in CECO to easily examine the metadata of all the available CCTV and select the best ones for validating the situations in reality. Furthermore, the standardization also simplifies the task of flexibly adding new CCTV systems to the monitoring mechanism whenever necessary, e.g., adding CCTV systems that are only allowed to access during extreme emergency situations.

\textsuperscript{*} Corresponding author
As this research mainly focuses on what is visible via CCTV, the CCTV surveillance system's hardware consists of a display device, storage device, transmission device, and monitor body. As this research mainly focuses on what is visible via CCTV, only the CCTV monitor body is analyzed. Two types of CCTV are identified according to if the captured images can be adjusted during operations.

1. CCTV with fixed image range:
   The image range of this type of CCTV is fixed and cannot be adjusted. This type of CCTV is highly adopted by a large number of system integrators, security companies, and monitor equipment sellers due to the low cost. Since the image range cannot be adjusted, the deployment of this type of CCTV must consider the subjects being monitored beforehand, such that the best-deployed location can be determined. Sometimes more than one CCTV is required to provide better visual coverage for the subject of interests.

2. Variable image range CCTV monitor:
   According to the application needs, the captured images of this type of CCTV can be adjusted to a certain degree. Two major functionalities are considered, zoom in/out and orientation adjustment. With the addition of a retractable lens, operators are able to flexibly zoom in or out to get better inspections for the subjects of interest. However, whether the subjects of interest can be vividly interpreted also depends on other factors, e.g., resolution, weather, distance to the phenomena of interests. The capability of orientation adjustment additionally allows the operator to change the orientation of the CCTV other than its original deployed orientation. By horizontally or vertically rotate the yaw and pitch angles, the visual scope can be dynamically adjusted. The maximum range of FOV is nevertheless still restricted by the hardware specification. The EOC commander needs information to contact the authorized agencies of the CCTV to operate the CCTV according to the requests.

In addition to the spatial coverage aspect, another essential property of CCTV is its capability to operate during the night time. The night vision type is further subdivided into low illumination and infrared night vision. Its major function is to automatically change the color to black and white and increase the resolution when the illumination is low at night to obtain the recognition of night images.

Regardless of whether the shooting scope can be adjusted or not, every CCTV must provide the information about its Field of View (FOV) because it is the necessary information to determine if a particular CCTV can suffice the need for visually inspecting the reported disasters. In GIS-based applications, the simplest way of showing the geographic distribution of CCTV is to represent each CCTV as a point symbol. Despite the fact that this approach shows the location of CCTV, but it fails to provide information about the area CCTV covers. An improved approach is to enhance the visualization by adding a circular sector to represent the FOV of CCTV. An obvious advantage is that some unqualified CCTV can be quickly excluded, manually, or computationally. However, this is often conducted in the 2D environment with map-overlay technique, so the major improvement focuses on the illustration of the horizontal orientation (i.e., azimuth), but does not take the effect of "deployed height," "pitch angle," and "blocked sight" into consideration. All of the three factors have a significant influence on determining if the subject of interests in the real world can be seen or not. If a CCTV cannot provide direct visual inspection on the subjects of interest, it should be quickly excluded, regardless of how excellent its hardware specification is. Despite multicamera systems are progressively becoming prevalent (Demir, Bilal; Leblebici, Yusuf; Thiran, Jean-Philippe, 2020), its best advantage can only be demonstrated after the above three issues are removed. We argue that the integration of 3D FOV analysis and 3D urban environment is the key to this challenge. When the location and orientation of a CCTV are determined after deployment, a 3D FOV can be generated. This technique has been extensively used for choosing the deployed location of CCTV (including multiple CCTVs); designers simulate the deployed location and visually inspect the area visible in the captured images and continuously adjust it until the subjects of interest can be seen (with satisfactory image resolution). Typical software for aiding CCTV deployment includes ArcGIS Pro, Revit, Sketchup pro, et al. The 3D representation of indoor/outdoor environment with an acceptable level of details and accuracy is necessary to conduct a good deployment plan (especially the locations available for deployment in the indoor environment is often restricted.).

Despite its durable and flexible capability for supporting the deployment of a single CCTV, the CCTV deployment software rarely focuses on the management of CCTV systems, not to mention the integration of multiple heterogeneous systems.

To meet the various emergency response demand using CCTV systems to support instantaneous and validation monitoring, we believe a comprehensive approach based on standardized metadata (Panta, F. J., Péninou, A., & Sèdes, F. 2019). 3D FOV analysis is not only an aggressive strategy for helping EOC commanders to make prompt and accurate decisions but also a universal solution for effectively taking advantage of the heterogeneous CCTV systems. This is especially beneficial for integrated and cross-domain applications like smart cities.

### 2. METADATA FOR CCTV CAMERA AND FOV

To enable the commanders to establish complete and comprehensive control over the heterogeneous CCTV systems. A metadata-based sharing mechanism is proposed in this research. Focusing on determining the required CCTV metadata during emergency response, the well-known 5W1H approach (i.e., Where, When, Who, What, Why, How) is adopted in this research to ensure the various essential characteristics of CCTV are fully examined and included in the metadata design if necessary. As the scope of metadata only focuses on the selection of CCTV during emergency response, this implies the metadata discussed in this research is only a subset of the metadata for describing CCTV. The 5W1H approach typically begins with asking questions from one of 6 perspectives, e.g., the “WHO” aspect asks who will be involved in the emergency response. Even if many stakeholders may respectively play a particular role, the responsible agencies of the CCTV are definitely the key stakeholders because they have a dominant influence on if the CCTV signals are available, the formats of the streaming or captured images (WHAT aspect), and how they are distributed (HOW aspect). Hence not only the...
metadata for a single aspect must be considered; their interrelations are also very important (Algu, S. P., Bhat, P., & Jain, S. 2014). The following discussion respectively, explores every aspect of the 5WH approach and proposes their metadata elements.

1.WHO
The metadata related to the “who” aspect mainly includes the descriptions about the stakeholders that establish, maintain, and provide service to the CCTV systems, as they will be able to provide required accurate specification information and grant access to the CCTV system to EOC commanders whenever necessary. The metadata should record at least the official name and contact information of the stakeholders and can be used for both government agencies and private sectors. In Table 1, the metadata element named Authority is designed for recording the information. (The name of the organization is the contact information, such as telephone, email, or webpage)

2.WHAT
The metadata related to the “what” aspect is mainly designed to describe the hardware specifications, working environment, and use limitation of the CCTV systems, such that commanders can determine if a CCTV can be used for the intended applications. Proposed metadata elements include the Fixed coverage, Night vision, Resolution, Type, Manufacturer, Identifier, Magnification number, Maximum operating temperature, Minimum operating temperature, Maximum operating humidity, Minimum operating humidity, etc. The metadata about the working environment is used to determine the available CCTV according to the situation reality dynamically, e.g., if the CCTVs are usable at night. The element of Identifier is used to identify an individual CCTV uniquely. It may also be used for identifying the internet-based streaming service for the chosen CCTV.

3.WHY
The establishment of any CCTV system has its intended applications, which often have impacts on the chosen types and deployed location of the CCTV. The metadata element of Purpose is designed to explain why the CCTV system is developed. As the selection of CCTV in emergency response may start by looking for all available CCTV systems, the aspect of “why” helps to exclude those CCTV systems that may not qualify the demands.

4.WHERE
The metadata related to the “where” aspect is mainly concerned where the CCTV is located and the area visible with CCTV. This is the most important metadata as it determines if the described CCTV is usable for the reported disaster incidents. Proposed metadata elements include the Street address, Landmark, Location, Orientation, FOV, Visible objects, etc. While the first two are used to describe the approximate location of the CCTV with familiar geocoding or names, the elements of Location and Orientation use quantitative coordinates and measurements to describe the status of the CCTV. The coverage information includes two major perspectives, the element of FOV describes the coverage from a 3D geometric representation perspective while the element of Visible Object the information about the features that are visible via the CCTV. These features can be identified either by their name or IDs.

5.WHEN
The metadata related to the “when” aspect is mainly designed to provide information about the temporal information about the availability of CCTV recordings. The proposed elements include the Deployment time, Daily beginning shooting time, Daily ending shooting time and Historical data. The Deployment time implies no recordings will be available before the specified data, while the element of Historical data implies the time the past recordings are available (e.g., within past 3 days). The elements of Daily beginning shooting time and Daily ending shooting time record the operational time period for the CCTV, if it is not operated 24/7.

6.HOW
The metadata related to the "how" aspect is mainly related to the information of accessing the CCTV. The proposed elements include the Available services, Access instruction, Online access. The element of Available service is used to indicate if the service is available to organizations other than its responsible agencies. The content of Access instruction is given by the Authority for instructing how to get the access of CCTV streaming or recordings. Finally, element of Online access provides URL of the CCTV, provided the signal is available online.

Table 1 summarizes the name, data type and definition of the proposed metadata elements.

| Metadata element | Data type | Definition |
|------------------|-----------|------------|
| WHO              | Authority | string     | The name and contact information of the organization that deploys CCTV. |
| WHAT             | Fixed coverage | boolean | If the coverage of CCTV can be adjusted upon requests. |
|                  | Night vision | boolean | If the CCTV can capture images during night time |
|                  | Resolution  | string     | 108MP/480TVLines/380TV Lines |
|                  | Type        | string     | Different types of the CCTV: Speed Dome/Body Camera/Zoom Camera |
|                  | Manufacturer| string     | Name of the CCTV manufacturer |
|                  | Identifier  | String     | Information to uniquely identify a CCTV, e.g., series number. |
|                  | Magnification| numbers   | The ratio of visually enlarging the visible objects, e.g., 18X, 22X, 36X . |
|                  | Maximum operating temperature | integer | The maximum temperature for the CCTV to remain working, recorded in Celsius degree |

Table 1
| **Minimum operating temperature** | integer | The minimum temperature for the CCTV to remain working, recorded in Celsius degree |
|-----------------------------------|---------|------------------------------------------------------------------|
| **Maximum operating humidity**    | percentage | The maximum humidity for the CCTV to remain working, recorded in percentage |
| **Minimum operating humidity**    | percentage | The minimum humidity for the CCTV to remain working, recorded in percentage |
| **WHY**                           | Purpose | string | The purpose for deploying the CCTV. |
| **WHERE**                         | Street address | string | Street address of the building where the CCTV is deployed. |
|                                  | Landmark | string | The landmark (e.g., road junction) where the CCTV is deployed. |
|                                  | Location | string | 3D coordinates of the CCTV. |
|                                  | Orientation | numbers | The orientation of the CCTV, usually represented as $\omega$, $\phi$ and $\kappa$ or Roll, Pitch and Yaw. |
|                                  | FOV | geometry | The field of view of the CCTV, represented by 3D patch. |
|                                  | Visible objects | string | The ID of the features that are visible via CCTV. |
| **WHEN**                          | Deployment time | Time | The time when the CCTV is deployed. |
|                                  | Daily beginning shooting time | String | Used when the CCTV is only working in certain time period everyday, e.g., 10AM. |
|                                  | Daily ending shooting time | String | Used when the CCTV is only working in certain time period everyday, e.g., 3PM. |
|                                  | Historical data | boolean | If historical data is available. |
| **HOW**                           | Available service | boolean | If the CCTV is open to the public. |
|                                  | Access instruction | string | Instructions about how to access the streaming service of the CCTV. |

Table 1. Metadata of CCTV monitor designed based on 5W1H

**FOV Analysis of CCTV System**

In 2D maps, CCTV is usually illustrated by point symbols based on their location (Kim et al., 2004). The EOC operators click the point symbols to obtain the streaming image of the CCTV monitor. However, this method only provides the information about the CCTV's geographic distribution, but not the spatial coverage of CCTV. The EOC commanders know where the CCTV are, but not where (and what features) the CCTV is shooting at. Even if the coverage is added to the 2D map to indicate the direction and FOV of the CCTV, it still lacks the CCTV's height and orientation information to determine its precise FOV in reality. Figure 2 shows the analyzed outcomes of two pairs of CCTVs using IPVM (https://ipvm.com/about), overlapped areas are found in both pairs of CCTVs, which indicates the overlapped area are visible from more than one CCTV. Multiple coverage is often considered to be advantageous, as the information collected from two different CCTVs can be compared to validate the situations in reality.

![Figure 2. CCTV monitor field of view in 2D FOV Map (IPVM)](image)

Different from Figure 2, Figure 3 and Figure 4 show the FOV analyzed results after introducing the 3D perspectives using the viewshed analysis function of ArcGIS Pro.
3. SCENARIO

Three scenarios are further discussed in this section. Two CCTV systems are respectively simulated (Table 2 and Table 3):

### System 1 CCTV’s Metadata

| 5W | Metadata     | Content                                      |
|----|--------------|----------------------------------------------|
| WHO| Authority    | System 1                                     |
| WHAT| Fixed coverage | True                                         |
|     | FOV          | Represented by 3D patch.                     |
|     | Night vision | True                                         |
|     | Resolution   | 180MP                                        |
|     | Type         | Body Camera                                  |
|     | Identifier   | ID 1,2,3,4                                   |
| WHY| Purpose      | Crime prevention                             |
| WHERE| Street address | No.1,Daxue Rd.,East Dist., Tainan City 701, Taiwan |
|     | Visible objects | Building ID 37,70,91,35                     |
| WHEN| Deployment time | 2018/10/01 20:14:00                          |
|     | Daily ending shooting time | continue                                    |

Table 2. System 1’s CCTV Metadata designed

### System 2 CCTV’s Metadata

| 5W | Metadata     | Content                                      |
|----|--------------|----------------------------------------------|
| WHO| Authority    | System 2                                     |
| WHAT| CCTV video format | True                                        |
|     | FOV          | Represented by 3D patch.                     |
|     | Night vision | False                                        |
|     | Resolution   | 108MP                                        |
|     | Type         | Body Camera                                  |
|     | Identifier   | ID 1,2                                       |
| WHY| Disaster Types | Safety                                |
| WHERE| Street address | No.1,Daxue Rd.,East Dist., Tainan City 701, Taiwan |
|     | Visible objects | Building ID 3,35,36,37,52,67,69,70,87,91,142 |
| WHEN| Deployment time | 2016/04/18 12:15:00                          |
|     | Daily ending shooting time | 20:00:00                                   |

Table 3. System 2’s CCTV Metadata designed

### 3.1 Scenario 1

Scenario 1 is for simulating the coverage of a selected CCTV system. Every CCTV system has its own coverage design. By selecting all CCTVs in the same system, the 3D coverage from the same authority can be shown to the commanders. Figure 5 shows the analysed outcome where the Authority is System 1. The objects visible by the chosen CCTV system can then be determined accordingly and recorded with the element of Visible Object.
3.2 scenario 2

Scenario 2 demonstrates the searching of available CCTVs from the perspective of objects in reality. By specifying Feature ID (Building 1) as the constraint, the recording data in the element of Visible objects are compared and two CCTV respectively belong to two different CCTV systems are found. However, this implies Building 1 is visible. Further 3D FOV analysis (Figure 6) shows that a portion of the Building 1 is not visible due to barriers between the CCTV and the building. But there are also a portion of the Building 1 that is visible from both CCTVs. This case clearly demonstrates the advantageous of managing multiple CCTV systems via standardized metadata and providing better coverage than a single CCTV system.

3.3 scenario 3

As some types of CCTVs may not work at night, scenario 3 is designed to demonstrate the different visual coverage in daytime and night time. Based on the elements of Night vision, CCTVs unable to work at night are removed from the list when the analysed scenario is at night. Figure 7 and Figure 8 respectively show the analysed outcome at day time and night time.

4. CONCLUSION AND FUTURE WORKS

The sharing of standardized metadata enable better use of CCTV systems deployed by different stakeholders. This is especially advantageous for emergency response assignment as we are looking for all available CCTVs to validate the disaster reports. By proposing a set of metadata elements specifically focusing on the needs of using CCTV in disaster management applications, we demonstrate the search, management, selection and access of CCTV streaming service can successfully conducted. This may serve as the basis of establishing a new collaboration framework of heterogeneous CCTV systems across different stakeholders. The extension of FOV analysis from 2D to 3D provides a new insight to the relationships between CCTV and objects in reality. In the future, we will finish the following works: Discussion on the scope of CCTV 3D FOV, which can be the actual viewing area. In disaster response, the commander can be uniformly the meaning of CCTV. The FOV of indoor CCTV monitors can be further analyzed.

REFERENCES

Koo, J., Oh, S. R., & Kim, Y. G. (2019). Device identification interoperability in heterogeneous IoT platforms. Sensors, 19(6), 1433.
Yeganegi, K., Moradi, D., & Obaid, A. J. (2020). Create a wealth of security CCTV cameras. In Journal of Physics: Conference Series (Vol. 1530, p. 012110). Demir, B. (2020). Beyond Cameras: Real-Time High-Resolution 3D and Panoramic MultiView Vision Systems with Their Applications (No. THESIS). EPFL.

Panta, F. J., Péninou, A., & Sèdes, F. (2020, August). Negative filtering of CCTV Content-forensic video analysis framework. In Proceedings of the 15th International Conference on Availability, Reliability and Security (pp. 1-10).

Khan, P. W., Byun, Y. C., & Park, N. (2020). A Data Verification System for CCTV Surveillance Cameras Using Blockchain Technology in Smart Cities. Electronics, 9(3), 484.

Panta, F. J., Péninou, A., & Sèdes, F. (2019, December). An Approach for CCTV Contents Filtering Based on Contextual Enrichment via Spatial and Temporal Metadata: Relevant Video Segments Recommended for CCTV Operators. In Proceedings of the 17th International Conference on Advances in Mobile Computing & Multimedia (pp. 195-199).

Kim, J., Park, N., Kim, G., & Jin, S. (2019). CCTV Video Processing Metadata Security Scheme Using Character Order Preserving-Transformation in the Emerging Multimedia. Electronics, 8(4), 412.

Lee, S., & Kim, T. (2019). U.S. Patent No. 10,271,018. Washington, DC: U.S. Patent and Trademark Office.

Kim, T., Youn, J., Kim, D., & Choi, W. (2019). DESIGN OF APPLICATION ARCHITECTURE FOR 3D SAFETY STATUS INFORMATION PLATFORM. International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences.

Baučić, M., & Medak, D. (2015). Web GIS for airport emergency response-UML model. Promet–Traffic&Transportation, 27(2), 155-164.

Algur, S. P., Bhat, P., & Jain, S. (2014). Metadata Construction Model for Web Videos: A Domain Specific Approach. International Journal of Engineering and Computer Science, 3(12).

Genc, Z., Heidari, F., Oey, M. A., van Splunter, S., & Brazier, F. M. (2013). Agent-based information infrastructure for disaster management. In Intelligent systems for crisis management (pp. 349-355). Springer, Berlin, Heidelberg.

Wang, F., & Yuan, H. (2010). Challenges of the sensor web for disaster management. International Journal of Digital Earth, 3(3), 260-279.

Kim, Y. H., Rana, S., & Wise, S. (2004). Exploring multiple viewshed analysis using terrain features and optimisation techniques. Computers & Geosciences, 30(9-10), 1019-1032.

Murray, A. T., Kim, K., Davis, J. W., Machiraju, R., & Parent, R. (2007). Coverage optimization to support security monitoring. Computers, Environment and Urban Systems, 31(2), 133-147.

Burrough, P. A., McDonnell, R., McDonnell, R. A., & Lloyd, C. D. (2015). Principles of geographical information systems. Oxford university press.