INDUSTRIAL TECHNOLOGY ADVANCES

Intelligent multimedia surveillance system for safer cities

TAKESHI ARIKUMA AND YASUNORI MOCHIZUKI

The demands for Smart City are growing in the context of the rapid urbanization in the world. In the public safety domain, law enforcement agencies are facing new challenges to protect people and properties in the cities, and they are shifting their focus from the traditional crime investigation towards preventive actions that prevent accidents and incidents before it occurs, and inter-agency collaboration that brings the situation under control swiftly when it occurs. In this paper, we discuss new challenges that cities are facing, and describe NEC's solution based on intelligent multimedia surveillance architecture for Safe Cities to tackle these challenges.

I. INTRODUCTION

In accordance with the fast expanding urbanization, expectations on Smart City are growing rapidly all over the globe. Presently, more than half the world's population lives in the urban area and this number will grow to 66% by 2050 according to the United Nations [1]. To meet the growing demands from the new citizens, it is necessary for cities to improve the performance and quality of their urban services. As a consequence, expectations on Smart City, in which ICT is fully deployed to improve the urban services, catch a great deal of attention in many parts of the world. According to an estimation done by Arup, the global market of Smart City can reach $408 billion per annual by 2020 [2].

One of the most important aspects of Smart City is the safety because the sense of personal and property security is mandatory to make institutions and businesses be functioned, and society be flourished. Ensuring citizens feel safe is a multidimensional task requiring governments to take a long-term and broad perspective. Therefore, the traditional systems such as the crime investigation system for police agency should evolve into intelligence solutions, which will play one of the major functions of Smart City along with other functions of Smart City in a coordinated manner.

NEC has proposed a vision called "Safer Cities" where people can live, work, and play in safety and comfort (Fig. 1) [3]. Safer Cities consists of seven domains: Law Enforcement, Citizen Services & Immigration Control, Public

NEC Corporation, 7-1, Shiba 5-chome, Minato-ku, Tokyo 108-8001, Japan. Phone: +81 3 3454 1111

Corresponding author:
T. Arikuma
Email: t-arikuma@ce.jp.nec.com

Administration Services, Critical Infrastructure Management, Information Management, Emergency & Disaster Management, and Inter-Agency Collaboration. NEC provides various solutions for these domains to support the government agencies.

Among the seven domains, Law Enforcement domain plays one of the key functions to protect people and properties in the city. The Law Enforcement should be equipped with capabilities of preventive actions and swift resolution of situations by inter-agency collaboration on the top of existing systems such as crime investigation systems.

NEC has been providing intelligence solutions for Law Enforcement domain to prevent incidents and accidents to support police and other agencies involved in the maintenance of public order. One of the examples is Safe City Test Bed led by the Singapore government in January 2013. The Singapore government invited companies to participate in its Safe City Test Bed. A number of government agencies also participated in this test-bed to collaborate with the selected companies to validate solutions that drive timely and accurate intelligence to the ground for better situation awareness of fast-developing situations [4]. NEC led one of the consortia and demonstrated intelligent multimedia surveillance solution in the project.

In this paper, we focus on the intelligent multimedia surveillance solution for Law Enforcement domain. We discuss the changes on requirements and technical challenges toward advanced law enforcement operations and discuss NEC's solution based on intelligent multimedia surveillance architecture.

The rest of the paper is organized as follows. In Section II, we give an outline of the trends in surveillance for public

safety in Safer Cities. In Section III, we explain Singapore Safe City Test Bed project as a case study. In Section IV, the requirements and technical challenges toward Safer City are explained. Section V introduces NEC's intelligent multimedia surveillance system as a solution of Safer Cities and some technologies, which play the key role for Safer Cities.

II. TRENDS IN LAW ENFORCEMENT DOMAIN FOR SAFER CITIES

In recent years, there are growing interests on quick response and preventive action to prevent incidents and accidents on the top of the traditional crime investigation, and the inter-agency collaboration, which plays an important role for the swift resolution of incidents and accidents. In this section, recent activities and their backgrounds are explained.

A) Incident and accident prevention

To tackle increasing crime rate, severe incidents and accidents, some advanced police forces had already started deployment of crime prevention systems in the world. For example, Memphis police is taking a statistical approach to prevent crime. Memphis, Tennessee use to be ranked as one of the most dangerous cities in the USA by Forbs Magazine in 2009. The Memphis Police Department deployed a statistical data analysis technology called Blue CRUSH, which allows analysts to rapidly evaluate incoming patrol data against historical trends [5]. This enabled the police to able to respond to the predicted threats before a criminal

act is committed. The crime rate in the Memphis area was reduced by 31% since the deployment of the system [5].

Another example is an intelligent video surveillance approach taken by Surat City Police. Surat City Police is the first city police which established a Picture Intelligence Unit equipped with face recognition systems in India [6]. Surat is the eighth largest city in India with population of 4.6 million. The police deployed a face recognition technology called NeoFace Watch [7] that works with live surveillance cameras or "closed-circuit television" (CCTV) in the city. The technology alerts the police real time for suspicious activity/criminal spotted in the surveillance zone and thus enables speedy action by the police. The crime rate has dropped 27% and 150 cases have been solved after the deployment of the system [6].

B) Inter-agency collaboration

Another trend in the public safety is inter-agency collaboration. For the swift resolution of incidents and accidents, multiple government agencies collaborates and work together to bring the situation under control as soon as possible.

Some of the advanced cities have already started activities to share vital information among agencies. For example, Tigre City in Argentina deployed a city operation center which gathers and provides vital information for its public safety and security agencies. Tigre, which is located 32 km from the Argentinean capital Buenos Aires, had experienced a strong growth since 1990s. The city rolled a command and control center equipped with a sophisticated video surveillance system with CCTV network and intelligent video analysis in 2013 [8]. The operation enables

Fig. 1. Seven domains of "Safer Cities".

Tigre City to gather critical information for its agencies. For example, the video footage from CCTV is used not only for finding person of interest (POI) by police, but also for finding missing children by the Missing Children Argentina.

III. CASE STUDY: SINGAPORE SAFE CITY TEST BED

Singapore Safe City Test Bed is one of the good examples of the activities by advanced governments toward Safer Cities, and it provides a source of perspectives on requirements for an advanced law enforcement system. In this section, we summarized the detail of the project reported by Wang *et al.* [4] as a case study.

A) A sophisticated Test Bed

The Safe City Test Bed project was one year pilot project conducted by Singapore government in 2013, which involved a number of government agencies. NEC Asia Pacific led a consortium consisting seven partner companies [4]. The Safe City Test Bed comes with great complexity and sophistication. For example, police force would like to have capability of not only early detection of aggression or fighting, but also suspicious activities such as loitering persons, so that officers on patrol can be alerted more swiftly. Singapore agencies also required situation awareness such as traffic situation, so that they can react to traffic accidents or occasional congestion quickly.

Furthermore, the capability of providing the information and manage them in a holistic way was also required, so that commanders can provide better support officers on the ground with improved assessments of knock-on effects from an incident.

B) Upgraded capabilities

NEC demonstrated a comprehensive system, so that agencies can comprehend suspicious activities for the quick response. The system integrated various technologies from NEC and the consortium members (Table 1). The system provides vital information such as alerts for early detection of incidents or alert on suspicious activities. The system also provided not only alerts but also surrounding situations such as traffic condition around the incident or accident location on a GIS to support the officers to plan the reaction to the situation.

The results from the Test Beds became clear soon after the first deployments went online in late 2013. Video analysis helped detecting particular behavior with aggressive actions such as fights and snatch theft thronging through a crowd. Audio analysis then reinforces the confidence of the detected event by detecting shouting and screams, which often accompanies aggressive actions. The system also used facial recognition and video analytics to detect unusual activities such as loitering. Officers can make police questioning to the person as preventive action if they receive

Table 1. Technologies and solutions provided by the consortium members.

Vendor	Technology/solution
NEC	Acoustic analytics
	Alert event reasoning
	Cyber information surveillance system
	Device integrity
	Inter-agency information governance appliances
	M2M gateway
	Sensor node
	Video analytics
	Video non-repudiation
Esri	ArcGIS Platform
FORCE21	Video cameras
	Video management system
G element	3D situation visualization
Greenfossil	Inter-agency information governance appliance
iOmniscient	Acoustic analytics
	Smell analytics
	Video cameras
	Video management system
	Video analytics
Oracle	Event processing platform
	Database
ZweecAnalytics	Acoustic analytics and acoustic capturing system
	Hemispheric camera (HemCam)
	Video analytics

alert from the system. This provides vital information to officers to react the situation. Not only incident detection, the system can also provide surrounding situation such as potential points of congestion or blockage in the area where the traffic flow might be affected. It would be able to provide this situation awareness information to both ambulance and security services in real time on a map, so that they can plan best route to reach the incident scene.

IV. REQUIREMENTS FOR THE ADVANCED LAW ENFORCEMENT SYSTEMS

As we saw in the case study, advanced government agencies has already looking for an advanced law enforcement system that allow them to prevent incidents and accidents, and to bring the situation under control swiftly if it occurred. In this section, we generalized the requirements on the advanced law enforcement systems and discuss the technical challenges toward the advanced law enforcement systems.

A) Requirements

Toward the prevention of incidents and swift reaction to the incidents, there are three requirements (Fig. 2) in the general three phase of law enforcement operation, i.e. detecting suspicious activities, making decisions for proactive actions, and reacting to the situation [9].
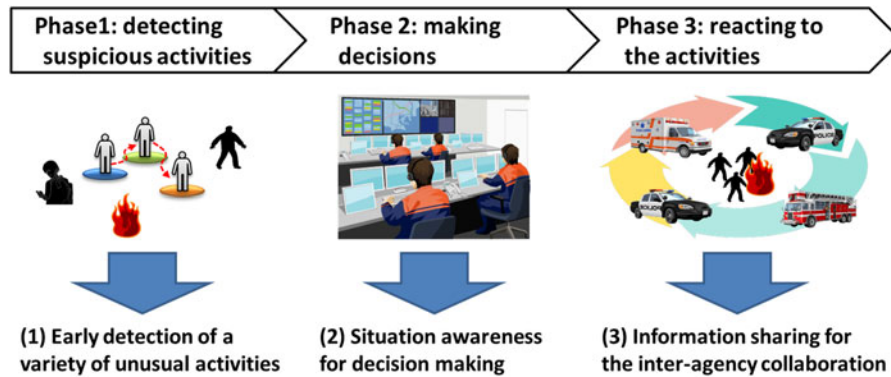
Fig. 2. Requirements in the advanced law enforcement activities.

1) EARLY DETECTION OF A VARIETY OF UNUSUAL ACTIVITIES

Government agencies would like to find incidents or accidents as early as possible. An incident is usually accompanied or composed by a variety of "unusual activities", and it is a key to detect these unusual activities on the ground for the early detection of the incident. Therefore, the system should provide a capability of detecting various types of unusual activities that can be presages of incidents or accidents.

2) SITUATION AWARENESS FOR DECISION MAKING

Government agencies would like to know not only the occurrence of unusual activity, but also the situation surrounding the activity, so that the offices can make decisions correctly and quickly. The system should provide a capability of providing the surrounding situation and development of the unusual activity.

3) INFORMATION SHARING FOR THE INTER-AGENCY COLLABORATION

Government agencies would like to collaborate closely to bring the situation under control as soon as possible. The system should provide a capability that various agencies which collaborating on a situation can access to a set of data on a "need to know" basis.

B) Technical challenges

The requirements listed in the previous section cannot be fulfilled by just applying existing surveillance systems because of following four major technical challenges.

1) DETECTING LARGE-VARIETY OF RARELY OCCURRING ACTIVITIES

Detecting large-variety of rarely occurring activities from CCTV video feeds is important for the early response to incidents. For example, the official report of the London riot in 2011 [10] reported that several unusual activities were captured by the CCTVs deployed in the city as listed in Table 2. If the surveillance system can automatically detect these variety of rarely occurring activates, then the police force can take actions such as sending police officers to the scene, before the situation became out of control.

The key technical challenge for the detection is the lack of training data required to develop unusual activity detectors. In general, thousands of non-biased training images are necessary for each unusual activity detector development. For example, a study on congestion detection [11] reported that if the number of the training data is fewer or there are any biases in training data such as backgrounds, lightings and number of people, the mean absolute error increases 2–10 times. However, gathering non-biased training images for city surveillance is a difficult task. Firstly, it is difficult to cover the variation of scenes such as background, lighting, and angles because of the variety of camera installation locations. Secondly, it is also difficult to gather images of rarely occurring unusual activities such as smashing shop windows because the occurrence of these activities in real field is too low to cover the variations of the people behavior such as number of people and how they behave during the activity.

Therefore, it is important for the system to provide a way to provide detectors for variety of unusual activity with limited number of training data.

Table 2. Unusual activities captured by CCTV during the London riot in 2011.

Unusual activity	Description in the report [10]
Smashing shop windows	"CCTV monitoring revealed approximately 50–100 youths in Fore Street smashing shop windows"
Gathering of crowd	"CCTV monitoring in Edmonton had picked up a large crowd at Joyce Avenue"
Obstructing a road	"A CCTV operator reported that approximately 30 youths were damaging shops and obstructing the road with barriers in Enfield"
People with stolen goods	"In Ponders End youths were seen on CCTV with goods believed to have been taken from the local Tesco's store"
Set alighting cars	"CCTV monitoring witnessed another car set alight in Colyton Way"
People with weapons	"A group of youths were seen on CCTV en route to Croydon police station armed with petrol bombs"

2) NETWORK CONSTRAINS FOR DISTRIBUTED SENSOR DATA ANALYTICS

There are stronger constrains on the system deployment for inter-agency collaboration compared to systems for single agency due to the information control. For inter-agency collaboration, it is difficult to have centralized system because some of analytics require confidential data, and they must be done within the agency's facility. For example, finding POI must be done in police divisions because the list of POI is considered as confidential information. Therefore, the system became physically distributed system by nature and network will be a bottleneck for the system. For example, Singapore subway system has more than 6000 CCTV for just two lines [12] which will require 18 Gpbs to send them as Full HD video without any data reduction.

Therefore, it is important for the system to provide a way to reduce the amount of data transferred between agencies to meet these physical constrains.

3) GATHERING SITUATIONAL INFORMATION AROUND THE DETECTED ACTIVITIES

Situation awareness, to provide right information at right time, is important, so that the officers can make decision correctly and quickly. For this purpose, system should be able to provide not only individual unusual activity alerts, but also related activities which lead to a higher level of understanding of the situation of interest.

4) DYNAMICALLY CHANGING ACCESS RIGHTS FOR INTER-AGENCY COLLABORATION

It is important to manage dynamically changing access rights for inter-agency collaboration because the access rights to the data are not static but change dynamically corresponding to the developments of the situation on the ground. For example, in general, police force may not allow sharing their CCTV footage due to security concerns. However, in case of severe fire, the video feeds from the police CCTV can be seen by related agencies such as the fire department and emergency medical service. This kind of operation requires dynamic change of access rights to the resources.

V. INTELLIGENT MULTIMEDIA SURVEILLANCE FOR SAFER CITIES

To tackle the technical challenges discussed in the previous section, we propose architecture for intelligent multimedia surveillance for the advanced law enforcement in Safer Cities. The architecture is designed to overcome the technical challenges mentioned in the previous section by remote sensor analysis, integrated analysis and information presentation (application), and dynamic access control (Fig. 3).

A) Remote sensor analysis

The remote sensor analysis component detects large variety of rarely occurring activities under the network constrains for distributed sensors.

1) DETECTING UNUSUAL ACTIVITIES

The remote sensors analysis component adopts two technologies to detect unusual activities: simulation generated training data approach [13] enabling detection of rare events and, multimedia analysis middleware called MAG1C Analytics Platform enabling integration of multiple analytics engines to fulfill user's requirements.

Simulation generated training data approach enables us to detect rarely occurring activities. One of the difficulties of detecting unusual activities is to collect enough sample data for the development of detectors. For example, to detect crowd behavior, it is necessary to prepare non-biased learning sample data sets that have conditions such as number of people, density of the people, variations of backgrounds, and overwrapping of persons in the scene.

To tackle this challenge, Hiroo Ikeda *et al.* proposed a machine learning technique using simulation generated training data [13]. In this technique, the non-biased comprehensive patterns of crowd situations are simulated from prepared person images to generate number of sample images as a learning data set (Fig. 4). They generated 1.2 million simulated sample images for machine learning, and achieved mean square error (MSE) of 0.47 (error of 0.69 people in average) for head counting, having 0.028 MSE for zero people cases showing high robustness on background variation [13].
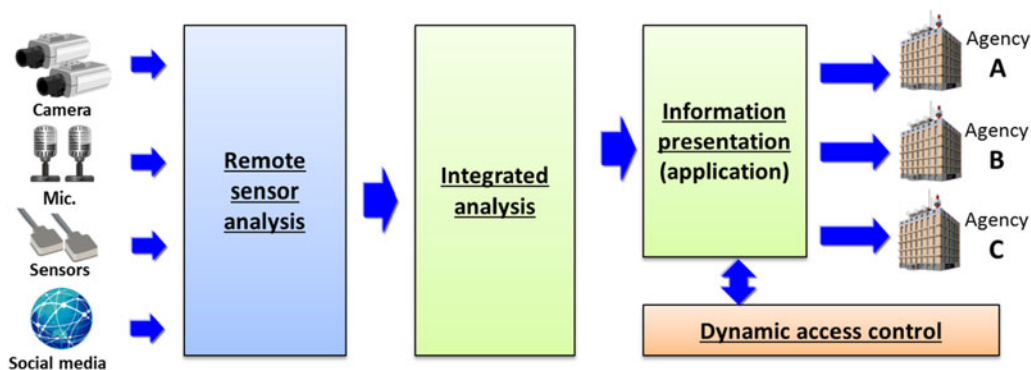

Fig. 3. The architecture of intelligent multimedia surveillances for Safer Cities.
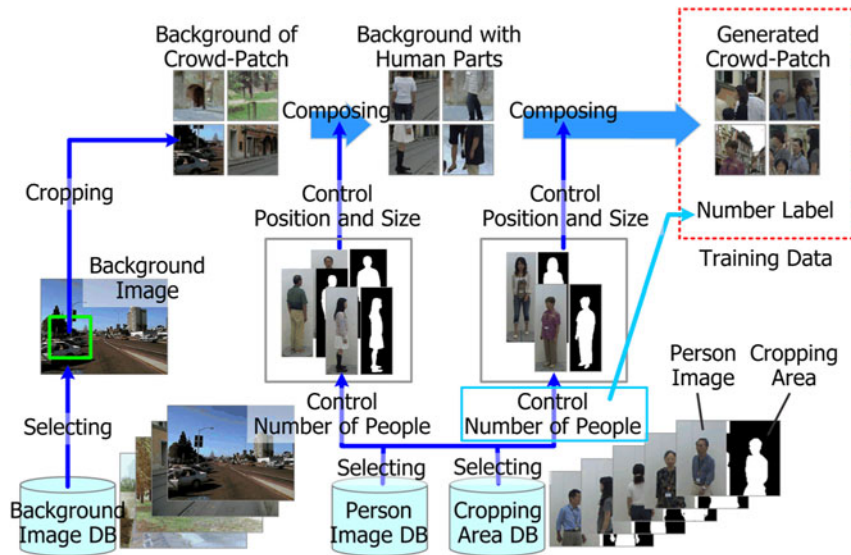
Fig. 4. Generating training data using simulation.

This simulation approach enabled us to reduce the cost of gathering sample data for the machine learning of these unusual behaviors by generating both sample image and labels of the data which can be used for the machine learning.

Multimedia analysis middleware enables us to enhancing detection capabilities. The system should provide various types of detection capabilities to detect unusual activates which relating to incidents of interest. MAG1C Analytics Platform [14] integrates various types of multimedia analysis engines, such as both video analysis and acoustic analysis, to detect a variety of unusual activities (Fig. 5). The system can integrate multiple and heterogeneous engines, so that the system can detect unusual activities by the combinations of these analysis results [15]. MAG1C Analytics Platform is also an open platform, i.e. users and venders can bring their own engine to the system by plugging into the middleware to expand the capabilities.

2) SHARING VIDEO OVER LIMITED NETWORK BANDWIDTH

Network constrains for distributed sensor data analytics can be overcame by leveraging a content-based data reduction approach with which system can reduce data amount without deteriorating the analysis quality. MAG1C Analytics Platform provides flexible framework for content-based data reduction algorisms [16]. The labeling algorism labels the video if the video contains the related contents such as number of the people or number of faces in the video, and it reduce the bitrate when there are no people or only few people at the site (Fig. 6). The framework was tested using surveillance video footages and the result shows that
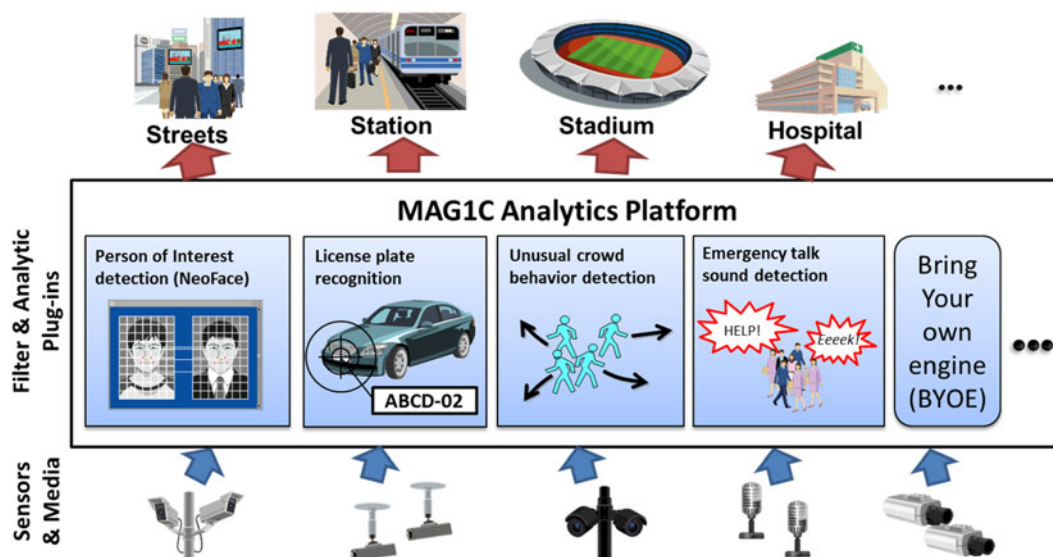

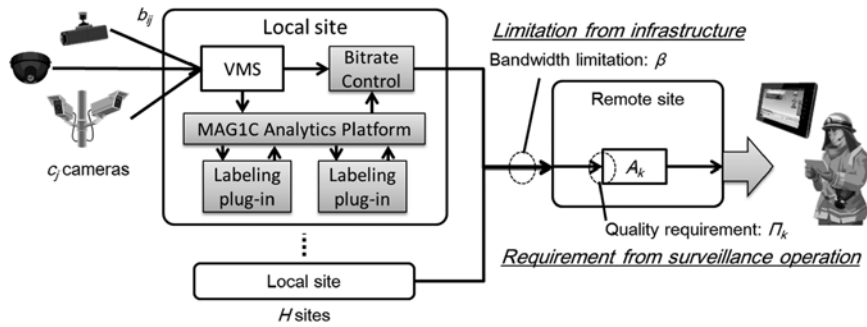**Fig. 5.** Concept of MAG1C Analytics Platform.

Fig. 6. Bitrate control by MAG1C Analytics Platform.

they system can reduced bit rate about 46 and 41% by face labeling and congestion labeling respectively without deteriorating accuracy at the remote site [16]. The platform enables system developer to select and combine labeling algorisms which is suitable for the remote site analysis.

B) Integrated analysis and information presentation

Integrated analysis and information presentation component provide situational information around detected activities. The integrated analysis provides relevant information when an unusual activity is detected by remote sensor analysis component. It uses ontology-based risk models to mix user's domain knowledge to analyze patterns and interferences of these detected events and information from other sources such as social media [4].

The information presentation component is a GIS application which shows both location of the detected events on a map and relevant information about the detected event for the situation awareness. For example, when multimedia analysis detects an unusual activity, relevant social media information is gathered and filtered by the integrated analysis component and presented to the operators via the information presentation component.

C) Dynamic access control

Dynamic access control component provide access control over multiple agencies. The component provides dynamic access control layer on the top of ordinary access control systems such as Active Directory in each agencies. This component enables various agencies to request and access information they required, while protecting it by handling dynamically changing access rights [4].

D) Use case example: early detection of unusual crowd activities

In this section, we will demonstrate how the architecture works for the early detection of unusual crowd activity scenario. Unusual crowd activity such as gathering and scattering of pedestrian are important for the accident prevention because they can be the early stage of incidents or accidents such as stampede. In the meantime, the system should be able to distinguish the difference between the unusual cases from usual one because these gathering or scattering also occur in usual situation.

The system based on the architecture described above is shown in Fig. 7. The remote sensor analysis provides early detection of the suspicious crowd activity. The simulation generated training data approach provides a robust engine for crowd behavior analysis. MAG1C Analytics Platform provide ways to combining multiple multimedia analysis, namely crowd behavior analytics using video and shouting detection using sound. By combining shouting detection results with crowd behavior analysis results, the system can selectively detect the event which has higher possibility of the severe incidents or accidents.
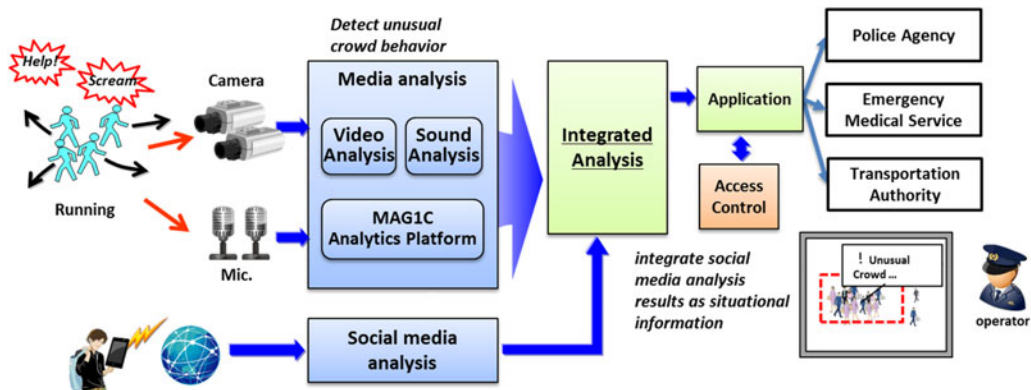

Fig. 7. Unusual crowd behavior detection on MAG1C Analytics Platform.

The system also can integrate social media information by the integrated analysis. The application shows not only the information about the detected activity itself, but also the possible eye-witness reports from social media. This can support the officers to understand what is happening on the ground from the command center.

Officers also can alter the access level for the camera in which unusual activity is detected using dynamic access control to make the video footage viewable for not only police, but also other related agencies such as emergency medical service, so that they can prepare possible ambulance dispatches.

VI. CONCLUSION

The expectation on Smart City, especially safety of the city for protecting the people and properties, is growing in accordance with rapid urbanization. Advanced government agencies have already looking at an advanced law enforcement operation that allows them to prevent incidents and accidents before it occurred by preventive actions, and to bring the situation under control swiftly if it occurred by inter-agency collaboration. NEC has proposed a concept called Safer Cities to tackle technical challenges including detecting large-variety of rarely occurring unusual activities, network constrains in physically distributed system, gathering situational information, and supporting dynamically changing access rights toward the advanced law enforcement operation. The intelligent multimedia surveillance architecture proposed in this paper overcomes these technical challenges by integrating robust media analysis using simulation generated training data generation and multimedia analysis engine integration middleware, flexible content base data reduction framework, integrated analysis for situational data fusion, and dynamic access control.

Looking forward, NEC aims to continue supporting the improvement of public safety through new and innovative technologies and partnerships that reach communities worldwide.

REFERENCES

[1] United Nations: World Urbanization Prospects The 2014 Revision, 2014.

[2] Department for Business: Smart cities – background paper, UK Government, 2013.

[3] NEC Corporation: Public Safety, 22 9 2015. [Online]. Available: http://www.nec.com/en/global/solutions/safety/ [Accessed 27 10 2015].

[4] Wang, P.; Woo, K.W.; Koh, S.K.: Building a safer city in Singapore. NEC Tech. J., 9 (1) (2015), 71–74.

[5] Perry, W.L.; McInnis, B.; Price, C.C.; Smith, S.C.; Hollywood, J.S.: Predictive Policing: The Role of Crime Forecasting in Law Enforcement Operations, RAND Corporation, 2013.

[6] Vasudevan, S.: Here is what Surat City Police did to cut crime rate by 27%, 2015. [Online]. Available: http://www.dqindia.com/here-is-what-surat-city-police-did-to-cut-crime-rate-by-27/ [Accessed 1 10 2015].

[7] NEC Corporation: NeoFace Watch, 2015. [Online]. Available: http://www.nec.com/en/global/solutions/safety/face_recognition/NeoFaceWatch.html [Accessed 27 10 2015].

[8] Vargas, J.; Bergonzelli, D.: Securing the future in Tigre. NEC Tech. J., 9 (1) (2015), 75–77.

[9] Niklas, H.; Jonas, H.; Helena, G.; Rogier, W.: Rationale for emergency management systems for local communities: a needs assessment, in Proc. of 9th Int. ISCRAM Conf., 2012.

[10] Metropolitan Police Service: 4 Days in August – Strategic Review into the Disorder of August 2011, Metropolitan Police Service, 2012.

[11] Zhang, C.; Li, H.; Wang, X.; Yang, X.: Cross-Scene Crowd Counting via Deep Convolutional Neural Networks, in IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), 2015, 833–841.

[12] March Networks: Singapore MRT moves ahead on CC TV system expansion. March Networks News – Transportation Edition October (2008), 14–15.

[13] Hiroo, I.; Ryoma, O.; Hiroyoshi, M.: Accuracy improvement of people counting based on crowd-patch learning, in 21st Symp. on Sensing via Image Information (SSII2015), 2015.

[14] NEC Corporation: Inter-agency collaboration: public safety, 28 10 2015. [Online]. Available: http://www.nec.com/en/global/solutions/safety/Inter-Agency/index.html [Accessed 28 10 2015].

[15] Arikuma, T.; Koyama, K.; Kitano, T.; Shiraishi, N.; Nagai, Y.; Kawamata, T.: Analysis control middleware for large-scale video surveillance, in 2013 10th IEEE Int. Conf. on Advanced Video and Signal Based Surveillance (AVSS), 2013, 294–299.

[16] Arikuma, T.; Jain, A.; Koyama, K.; Woo, K.W.; Kawamata, T.; Yamada, K.: content based video quality control for wide-area video surveillance systems, in Int. Conf. on Artificial Intelligence and Pattern Recognition (AIPR2014), 2014, 79–86.

Takeshi Arikuma received his B.E. and M.E. degrees in Computer Science from Tokyo Institute of Technology in 2006 and 2008, respectively. He joined NEC Corporation (Central Research Laboratories) as a researcher on large-scale systems in 2008. His main research interest is architecture design of large-scale video and audio analysis systems.

Yasunori Mochizuki received his Ph.D. degree in Electronic Engineering from the University of Tokyo in 1987. He is the Vice President of Central Research Laboratories (CRL), NEC Corporation since June, 2011 and is responsible for the management and/or coordination of cloud-related research activities covering computing architecture, networking, applications, and services. Before being appointed as VP, he was a General Manager of Information and Media Processing Laboratories in CRL from 2010, where he led the research groups on video object recognition, audio/video signal processing, speech recognition, natural language processing, data mining, and security technologies. Since April 2013, he has been also a General Manager of Corporate Technology, which is responsible for corporate-wide strategy on technologies for NEC. Dr. Mochizuki is a fellow of Japan Society of Applied Physics.