LOW-COMPLEXITY BIG VIDEO DATA RECORDING ALGORITHMS FOR URBAN SURVEILLANCE SYSTEMS

Ling Hu and Qiang Ni

School of Computing and Communications, Lancaster University, LA1 4WA, UK

ABSTRACT

Big Video data analytics and processing are becoming increasingly important research areas because of infinite generation of massive video data volumes all over the world. In this paper, by utilizing Bayesian-based importance analysis, we propose a set of novel, simple but effective video recording methodologies and intelligent algorithms to solve the so-called big video data volume problem in urban surveillance systems. The complexity of our proposed algorithms are only $O(n)$, hence they can be easily implemented in real video urban surveillance systems without complicated computational cost. The simulation results show amazing recording efficiency. Our methods can dramatically reduce the hard discs (HD) occupation space requirement so that the recording time of HDs can be greatly enlarged.

KEYWORDS

Big Video Data, Bayesian Rule, Urban Surveillance Systems

1. INTRODUCTION

Big data is used to describe a massive volume of both structured and unstructured data sets which is very large and beyond the ability of traditional database tools to capture, manage, record and process them. According to the 3V's definition, big data has three main characteristics, which are Volume, Velocity and Variety [1]. Volume refers to an extreme amount of data; the examples of big data might be petabytes (1,024 terabytes) and exabytes (1,024 petabytes) of data consisting of billions to trillions of records, from different sources (e.g. Web, sales, customer contact centres, social media, mobile data and so on). Velocity describes how fast the data is produced and must be processed. Variety means a wide range of data types. For example, data can be collected/stored in multiple formats, which may include image, audio, video, texts and other forms of data.

Among all these different varieties of big data, videos are the biggest. According to Cisco's statistics, images and videos make up about 90 percent of overall IP traffic [2]. Hence video data analysis receives an increasing attention due to its fast-growing applications and huge data volumes. For instance, the data volume of all video surveillance devices in Shanghai, China is up to 1 TB every day [3] and is expected to keep growing substantially in the future. The explosive increasing number of video resources has brought an urgent need to develop intelligent methods to reduce the pressure on the database management processes [4].

In order to dig out useful information from huge image and video data, many researchers focus on object detection from videos. For example, an application for vehicle search in crowded urban surveillance videos was proposed [5]. Another work proposed a new visual object tracking algorithm using a Bayesian Kalman filter with simplified Gaussian mixture [6].

Concerning the video data analysis, most researchers in the literature focus on compressing every single image frame. For example, some people proposed a method for simultaneously estimating
the high-resolution frames and the corresponding motion field from a compressed low-resolution video sequence [7]. Another research combines motion-compensation and transform coding schemes, and uses Bayesian method to enhance the resolution of compressed video data [8]. In order to cope with the high computational complexity, some researches focus on implementing the computation on a hardwired design, such as embedded compression engine targeting the reduction of full high definition video transmission bandwidth over the wireless networks [9]; or a CMOS image sensor for tracking the moving objects in region-of-interest and suppressing motion blur [10].

Although there is much work focusing on the detecting information from frames and videos, to the best of our knowledge, currently there is little work done to reduce data volume in the video recording processes. Even though it is important to deal with the single images, efforts need to be done on the recording processes to solve the big video data recording problem. Nowadays the big video data increases rapidly, which makes its volume much higher than the potential hard discs (HD) capacity. In this paper, we propose a set of simple and effective methods to intelligently and dramatically reduce the video frame number which are needed to be stored up in the HDs for the urban surveillance systems. To the best of our knowledge, we are the first to propose these schemes.

2. Novelty and Contributions

In this paper, we propose several new simple but effective methods to dramatically reduce the space of the HDs in order to record the huge amount of video data generated by urban surveillance systems.

Different from the current efforts in the literature which focus on single video frame compression, the novelty of our methods is to analyse and to only record those important frames according to feature requirement. Using Bayesian-based importance analysis, we propose three intelligent mechanisms, called pre-treatment scheme, post-treatment scheme and combined-treatment scheme. We first determine which are unimportant frames and then remove part of them in the recording process to dramatically reduce the video recording data volume while at the same time to keep acceptable visual quality. The simulation results demonstrate that using our proposed methods, the occupation space of the HDs can be significantly reduced.

An important advantage of our schemes is their low complexity features. The complexity of our methods is only $O(n)$. This means that they are very efficient, and hence they are easy to be adopted and deployed in real video urban surveillance systems.

3. Urban Surveillance System Model

Urban surveillance systems are widely used now to improve the intelligence of cities. The normal applications include counter surveillance in banks, entrance monitoring of safeguard regions, traffic monitoring of crossroads, and so on. In most cases, the cameras are fixed and the surveillance regions are also located constantly. A typical diagram of urban surveillance system is shown in Figure 1. In current urban video surveillance systems, there are huge number of cameras working 24 hours per day, and 7 days a week. Every camera focuses on a fixed area, in order to show the scenarios instantly, clearly and fluently. The frame per second (FPS) of every camera is normally kept at 30. In order to get higher quality for instant monitoring, some cameras increase the frame rate from 30 FPS to 60 FPS, or even higher (some cameras may even adopt 100 FPS). In these systems, digital video recorders (DVRs) are used to record the video frames; the most common number of inputs for a DVR is 1, 2, 4, 8, 16 and 32. Obviously the key components of
the DVRs are huge HDs. Although various compression techniques are currently used to reduce the size of every frame, the demand for high capacity HDs still increases rapidly due to the huge number of frames generated and to be generated explosively in the future.

A widely used way of urban surveillance systems is to monitor the traffic accident scenes of the roads. Research work has shown that there are four factors to influence safety management of traffic accident scenes, which are human factor, vehicle factor, road and environment factor, and management factor. The human factor, vehicle factor and management factor normally do not change between a 24 hour period; but the roads are more crowded at the peak time periods, it can worsen the road and environment factor and affect the psychological diathesis of drivers, which makes the probability of traffic accidents occurring to increase [11].

![Diagram of Urban Surveillance Systems](image)

**Figure 1. The Basic Diagram of Urban Surveillance Systems**

### 4. OUR PROPOSED SCHEMES

#### 4.1 Bayesian-Based analysis

Concerning the urban video surveillance systems, in order to reduce the number of recorded video frames, we propose the idea of only recording those frames which are potentially important. But, how do the systems know or determine which frames are important without extremely complex analysis of huge video frames data generated and to be generated continuously? First, we use Bayesian Theory to provide a theoretical guidance on how to design simple approaches of importance analysis. Let us assume that there is a camera installed at a bus station to monitor the traffic situations. We denote \( A_1 \) and \( A_2 \) as the events if a time duration is a peak time period and a non-peak time period, respectively. \( B \) is defined as the event if an accident happens.

\( A_1 \): The traffic is crowded (i.e. peak time).
\( A_2 \): The traffic is not crowded (i.e. non-peak time).
B: The event when an accident occurs.

Without loss of generality, we assume that the probabilities of peak time and non-peak time are the same, that is \( P(A_1) = 0.5 \) and \( P(A_2) = 0.5 \). Since evidences show that it is more prone for an accident to appear when the traffic is crowded [11], we assume the probability of an accident happening at peak time is \( P(B/A_1) = 0.03 \) and the probability of an accident happening at non-peak time is \( P(B/A_2) = 0.01 \). As defined, the probability of an accident happening is \( P(B) \). Now the problems are, at what time periods do an accident happen mostly? What is the probability that it happens in the peak time period? What is the probability that it happens in non-peak time period?

According to Bayesian rule, the probability of the accident happening in peak time period can be calculated as:

\[
P(A_1/B) = \frac{P(A_1)P(B/A_1)}{P(A_1)P(B/A_1) + P(A_2)P(B/A_2)} = \frac{0.5 \times 0.03}{0.5 \times 0.03 + 0.5 \times 0.01} = 75\%
\]

The probability of an accident happening in non-peak time period is obtained as:

\[
P(A_2/B) = \frac{P(A_2)P(B/A_2)}{P(A_1)P(B/A_1) + P(A_2)P(B/A_2)} = \frac{0.5 \times 0.01}{0.5 \times 0.03 + 0.5 \times 0.01} = 25\%
\]

The Bayesian rule is guidance for us to cut some unimportant frames in the recoding processes. From the above example we conclude that the probability of an accident happening in peak time period is much higher than that in off-peak time period even when the probabilities of peak time and off-peak time are the same. Although the high quality frames can be shown at 30 FPS or even higher instantly for the surveillants to keep watching on the scenarios in real time, it is not necessary for all the huge number of video frames to be recorded into the HDs.

4.2 Pre-Treatment Scheme

Our pre-treatment scheme is to divide a whole day into two parts: peak time period and non-peak time period. How to divide the time period is dependent on different scenarios. In the peak time period, we keep all the frames to be recorded into the HDs, which is normally 30 FPS. In this way, the comparably important frames are kept without loss, because they have higher probability to be reviewed in the future. On the other hand, we may cut some frames in the non-peak time periods, that is, to reduce the FPS. Even the FPS is as low as 1, it will effectively record the scenarios with human being's activities. Considering 1 FPS is only around 3% of the normally 30 FPS, the frames to be recorded could be reduced dramatically. The pseudo-code of our proposed pre-treatment is outlined in Algorithm 1.

Algorithm 1: Pre-treatment algorithm

Input: The video frames generated by cameras

Return: The video frames needed to be recorded to HD

1. Check current time \( t \)
2. if \( t \) belongs to peak hours
3. Goto 6
4. Else
5. Cut the FPS to the pre-defined number of non-peak time period
6. End
7. Record the video frames to HD
As we already know, a typical camera generates 5.4 G bytes video data per day. Hence, with a commonly 25 inputs DVR, it will need 4050 G bytes space to record the videos only for 30 days, which is a tremendous volume. In order to save the space of HDs, currently the DVRs have the program for the surveillants to choose which days’ video should be recorded. This is obviously a simple but coarse method, which may cause some important frames to get lost if we cut frames randomly.

Using our scheme, the calculation could be low-cost. Our model is simple to be implemented within the video processor, concerning the significant efficiency of cutting big video data. With the same HD space, obviously the total recording time will be extended.

4.3 Post-Treatment Scheme

In some cases, it is not practical to divide a day into two parts, as every time instant along the whole day could be equally important. Hence we propose a new post-treatment scheme. In most circumstances, people need to review the recorded video shortly after something happened. In practice, the probability of the need to review the recorded videos in the near several days is the highest. When the time flies away, the need to review the history videos reduces rapidly. By taking this into account, we propose a post-treatment scheme where our idea is to keep all the generated frames recorded into the HDs initially, which is normally 30 FPS. Since the importance of the recorded video frames reduces rapidly with time passing, we thus introduce a cut-off day. For instance, if we define the cut-off day as 3 days, this means that after the 3 days we could discard some of the recorded video frames, from normal 30 FPS to a lower FPS. In this way, the re-recorded files will then be significantly smaller to be saved concerning the space of the HDs.

The process of our proposed post-treatment is depicted in Algorithm 2.

Algorithm 2: Post-treatment algorithm

Input: The recorded video frames
Return: The video frames needed to be re-recorded to HD

1. Check current day d after the frames recorded day
2. if d = cutoff day
3. Discard some frames according to pre-defined FPS
4. Re-record the video frames to HD
5. Else
6. End

4.4 Combined-Treatment Scheme

Finally, our pre-treatment and post-treatment may be combined together to achieve better results. The process of the combined-treatment scheme is shown in Figure 2.
When an HD receives the video frames occurred from the surveillance cameras, it can begin to run the pre-treatment scheme, that is, to cut some frames from the non-peak hours and save the resulted data into the HD to generate the current recording files.

At later days, the system will check the recorded files to see whether or not the cut-off day is due. When the cut-off day is due, the post-treatment scheme will begin to work. The recorded files may be re-recorded with the pre-defined lower FPS; hence the size of the files will be reduced again.
By combining intelligent video analysis with high performance computing technology, our proposed methods will help to develop intelligent surveillance video systems and result in huge economic benefit.

4.5 Complexity Analysis

Let n denote the number of video frames. As for the time complexity of the pre-treatment scheme, the calculation time cost is only linear with the off-peak time's number of video frames, hence the time complexity of this scheme is only O(n).

As for the post-treatment scheme, the calculation time cost is linear with the pre-defined re-record frames' number after cut-off day, hence the time complexity of this scheme is also O(n).

For our combined-treatment method, the calculation time cost is also linear with the frame number at off-peak time or after the cut-off day, hence the time complexity of our combined-treatment scheme is still O(n).

5. Simulation Result

5.1 Pre-Treatment Scheme

We simulate our schemes and plot the simulation results in figures. In order to reflect the most common scenarios, we set the FPS at peak time period as 30. The FPS at off-peak time period may be set from 1 to 30. The peak time hours per day could be chosen from 0 to 24. The simulation results are illustrated in 3-dimensional plots. They are shown in Figure 3 and Figure 4.

Figure 3. Occupation Rate Comparing with the Original HD Occupation Space (3D)
Figure 3 illustrates the occupation rate comparing with the original HD occupation space for a 24 hours period. X-axis is the frames per second of non-peak hours; Y-axis shows the peak time hours per day; and Z-axis shows the results of the HD occupation rate in percentage comparing with the original HD occupation space.

Figure 4 shows the gains in recording time comparing with the original time period with the same HD. X-axis is the frames per second of non-peak hours, which ranges from 1 to 30. Y-axis shows the peak time hours per day, which is from 0 to 24. Z-axis shows the gains in recording time with the same HD, which are displayed in percentage.

Figure 3 and Figure 4 show significant improvement of the urban video surveillance systems. When we choose less peak-time per day and fewer frames per second of non-peak time period, the occupation rate comparing with the original HD occupation space falls down sharply. At the same time, the gains in recording time rise up at very high speed, concerning the same HD space. While our scheme is very simple and easy to be implemented, the performance gain is amazing. Since Figures 3 and Figure 4 are 3-dimensional figures, it is not easy for people to investigate further details. We choose the situations of peak time from 0 to 8 hours per day, which may cover most common scenarios, to plot the results in 2-dimensions. The 2-dimensional results are plotted in Figures 5 and Figure 6.

Figure 5 is the occupation rate comparing with the original HD occupation space for a 24 hours period. X-axis is the frames per second of non-peak hours, and Y-axis shows the result of the HD occupation rate in percentage comparing with the original HD occupation space. For different peak-time hours per day, the results are plotted in different lines.
Figure 6 illustrates the recording time gain comparing with the original time period with the same HD. X-axis is the frames per second of non-peak hours, and Y-axis shows the gains in recording time in percentage. For different peak-time hours per day, the results are plotted in different lines.

Figure 5. Occupation Rate Comparing with the Original HD Occupation Space

Figure 6. Gains of Recording Time
As shown in Figures 5 and Figure 6, at each different peak time hours per day, the occupation rate comparing with the original HD occupation space is a linear result, the lower FPS of non-peak time hours leads to the lower occupation rate.

At the same time, the time gain comparing with the original time period shows the amazing results. For example, when a camera watches at a never crowded road, the peak time may be set at 0. When the recorded FPS is set to 1, the system will not lose the main activities of cars and pedestrians. The occupation rate may be as low as around 3.3% of original HD space for a 24 hours period; and the recording time gain may be around 30 times compared to the original time with the same HD space.

5.2 Post-Treatment Scheme

Again, we plot here the simulation results. We still set the original FPS as 30. The cut-off days could be chosen from 1 to 30. The FPSs after the cut-off days are set from 1 to 30. The calculation is based on the original recorded period of 30 days. The simulation results are also plotted in 3-dimensions. They are illustrated in Figure 7 and Figure 8.

Figure 7 is the occupation rate comparing with the original HD occupation space for a 30 days period. X-axis shows the pre-defined cut-off days. After the cut-off day, the recorded video frames will be discarded partly. Y-axis shows the re-recorded FPSs after the cut-off day and Z-axis are the results of the HD occupation rate in percentage comparing with the original HD occupation space.
Figure 8 shows the gain of recording time in percentage using our post-treatment method, concerning using the same HD occupation space of the original period of 30 days. X-axis are the cut-off days. Y-axis shows the re-recorded FPSs after the cut-off day. Z-axis shows the gain of recording time in percentage.

Figures 7 and Figure 8 also show significant improvement of the urban video surveillance systems. The occupation rate comparing with the original HD occupation space falls down sharply when we choose smaller cut-off day and smaller re-recorded FPS after the cut-off day. At the same time, the gain in recording time rises up sharply in comparison with the same HD space of original 30 days’ images.

Again we choose the cut-off days from 1 day to 8 days to plot the 2-dimensional simulation results as shown in Figures 9 and Figure 10.

Figure 9 shows the occupation rate comparing with the original HD occupation space for a 30 days period using our post-treatment method. X-axis are the re-recorded FPSs after a cut-off day, and Y-axis shows the results of the HD occupation rate in percentage comparing with the original HD occupation space. For different cases of cut-off day settings, the results are plotted in different lines.

Figure 10 illustrates the gain of recording time, concerning using the same HD space of the original period of 30 days. Again X-axis shows the re-recorded FPSs after the cut-off day and Y-axis are the results.
As shown in Figures 9 and 10, at each different cut-off day, the occupation rate comparing with the original HD occupation space is a linear result. The gains of recording time again show
amazing results. For instance, when the recorded FPS after cut-off day is set at 1, it will not lose the main activities of human being and cars. In these cases, the gain of recording time in percentage are extremely high, which may easily reach several hundred, or even higher than one thousand percentage in common use scenarios.

5.3 Combined-Treatment Scheme

For our combined-treatment scheme, in order to get the occupation rate comparing with the original HD occupation space for a 30 days period, we may use X-axis to represent the peak-time hours per day, and use the Y-axis to show the cut-off day. At peak time hours and before the cut-off day, the original FPS is still set at 30. At off-peak time or after the cut-off day, the FPSs are all set to 1. Again, the Z-axis shows the occupation rate comparing with the original HD occupation space for a 30 days period. The 3-dimensional simulation results are shown in Figure 11.

Figure 12 shows the gain of recording time using our combined-treatment method, as compared to concerning using the same HD occupation space of the original period of 30 days. The X-axis represents the peak-time hours per day, and the Y-axis shows the cut-off days. The Z-axis shows the results in percentage, which are the gains of recording time comparing with using the same HD space of the original video frames of 30 days.

Figures 11 and Figure 12 show significant improvement of the urban video surveillance systems. The occupation rate comparing with the original HD occupation space falls down quickly when we choose smaller peak-time hours per day and smaller cut-off days, and the corresponding gains of recording time again increase sharply. The figures show that the proposed combined-treatment scheme achieves better efficiency than the pre-treatment method or post-treatment method.
Now we choose the cut-off day from 1 to 8 days situations, which may cover most common use scenarios, to plot the results in 2-dimensions. The results are shown in Figure 13 and Figure 14.
Figure 13 shows the occupation rate comparison by varying peak hours per day and cut-off day duration using our combined treatment method. During the peak time hours and before the cut-off day, the original FPS is set at 30. During the off-peak time or after the cut-off day, the FPSs are all set at 1 FPS. X-axis is the peak-time hours per day, and Y-axis is the results of the HD occupation rate in percentage comparing with the original HD occupation space. For different settings of cut-off days, the results are plotted in different lines.

Figure 14 illustrates the gain of recording time between the proposed combined-treatment scheme and the existing method. Again during the peak time hours and before the cut-off day, the original FPS is set to 30. During the off-peak time or after the cut-off day, the FPSs are set to 1. X-axis is the peak-time hours per day, and Y-axis shows the results, which are the gain of recording time in percentage. For different sets of cut-off days, the results are plotted in different lines. The gains of recording time in percentage reach several thousand. This indicates that our scheme can help to save the HD storage space dramatically.

6. CONCLUSIONS

In this paper we proposed several simple yet efficient video data recording algorithms to be used in urban surveillance systems. The main idea of our schemes is to intelligently record the important video frames with the number of unimportant video frames to be reduced as small as possible. Our proposed methods are very easy to be implemented on real time urban surveillance systems with low complexity and can efficiently save huge amount of the HDs storage space. The simulation results demonstrated that, using our combined-treatment scheme, the occupation rate may be as low as around 3% of original HD space. The gains of recording time may easily reach several thousand percentages, comparing with the same HD space recording the original video frames of 30 days period.
ACKNOWLEDGEMENTS

This work was supported by Institute of Software Application Technology, Guangzhou & Chinese Academy of Sciences under Smart City Urban Surveillance Systems project CSA7035.

REFERENCES


Authors

Ling Hu is currently a PhD student of Computing and Communications, Lancaster University, Lancaster, U.K. She got Bachelor of Engineering from Huazhong University of Science and Technology, China, Master of Science from École Supérieure d'Ingénieurs en Électronique et Électrotechnique, France and Master of Engineering from Dublin City University, Ireland. She has both work experiences of industrial world and academic world as engineer and research assistant.

Qiang Ni is a Professor at School of Computing and Communications, Lancaster University, Lancaster, U.K. He is also a member of Data Science Institute at Lancaster University. He is a Fellow of IET and Senior Member of IEEE. His main research interests lie in the area of future generation networks, clouds and big data analytics. He had already published over 150 papers that have attracted over 4700 citations.