

Emerging Trends in Video Surveillance Applications

Rajesh. N¹ , Dr.H.Sarojadevi²

¹Dept of ECE, Nitte Meenakshi Institute of Technology, Bangalore, India

²Dept of CSE, Nitte Meenakshi Institute of Technology, Bangalore, India

Abstract. This paper presents a review of emerging trends in video surveillance applications. Various techniques are highlighted and compared. Detection and tracking of moving objects is important in the analysis of video data and higher level security assessment. Image registration plays a key role in designing video surveillance system. The purpose of image registration in surveillance application is to geometrically align two or more images to enable in-depth data analysis such as feature detection, or motion sensing. Image registration is especially necessary to compare and integrate image data that comes from different sources, times, or perspectives. Use of high performance FPGA that can be dynamically reconfigured for video surveillance applications, together with camera interface can offer improved performance.

Keywords: Surveillance, Image registration, Feature detection and matching, Reconfigurability, FPGA.

1. Introduction

There is growing demand in aerial surveillance using video cameras. Compared to traditional framing cameras, videos provide the capability to observe ongoing activity within a scene and to automatically control the camera to track the activity. Separating an aerial video into the natural components corresponding to the scene is essential. The major components of the scene are the static background geometry, moving objects, and appearance of the static and dynamic components of the scene. The geo-location of video and tracked objects can be estimated by registration of the video to controlled reference.

With the development of video coding technologies, digital video surveillance has become a very hot application in the recent few years. In most of the present solutions for digital video surveillance system, video coding standards for consumer electronics, such as MPEG is widely used. One possible reason is that the ASICs containing these video coding algorithms are of low cost. However, due to different application backgrounds, when utilized in the digital video surveillance systems, these standards are not the optimal choices[1]. A good video coding algorithm for surveillance should support multiple channels (usually a central station connects to more than one video source), it should have alarming function, the video quality should be better for the moving objects as the security is a basic requirement, high compression ratio and fast processing speed is a must (usually the video surveillance system will run day and night, frequent changes of storage media are tedious and not acceptable) and code stream should be protected from easy modification because the video data may be a proof in the court.

The paper is organized as follows. Section 2 provides an overview of video surveillance application and associated imaging methods. Section 3 deals with application of image registration, different steps in image registration and methods. Section 4 explains about the implementation trends in surveillance applications and Section 5 gives an idea about accuracy tests in industrial and video surveillance applications.

2. Video Surveillance

¹ Corresponding author. Tel.: + 919448912098; fax: +918022167805.
E-mail address: nrajesh7@gmail.com.

Aerial surveillance has a long history in the military for observing enemy activities and in the commercial world for monitoring resources such as forests and crops. Similar imaging techniques are used in aerial news gathering and search and rescue. Until recently, aerial surveillance has been performed using framing cameras to gather high-resolution still images of an area under surveillance that could later be examined by human or machine analysts to derive information of interest. Currently, there is growing interest in using sophisticated video cameras for these tasks. Video captures dynamic events that cannot be understood from aerial still images. It enables feedback and triggering of actions based on dynamic events and provides crucial and timely intelligence and understanding that is not otherwise available.

Video surveillance is typically performed with one or two video cameras that are mounted on a platform[2]. Image processing performed can be divided into two stages front-end processing and scene analysis. Front-end processing is applied to the source video to enhance its quality and to isolate signal components of interest such as electronic stabilization, noise and clutter reduction, mosaic construction, image fusion, motion estimation, and the computation of “attribute” images such as change and texture energy. Scene analysis includes operations that interpret the source video in terms of objects and activities in the scene. Moving objects are detected and tracked over the cluttered scene. Image registration is a major technique used in video surveillance applications mainly for change detection, multichannel image fusion, mosaicing, image correction etc between different image frames.

3. Image Registration

3.1. Applications of Image Registration

Applications of image registration can be divided into four main groups based on image acquisition methods:

1) Multiview analysis: Images of the same scene are acquired from different viewpoints to gain larger view or 3D representation of the scanned scene. The applications being mosaicing of images of the surveyed area and shape recovery.

2) Multitemporal analysis: Images of the same scene are acquired at different times[11], under different conditions to find and evaluate changes in the scene which appeared between the consecutive image acquisitions. The applications being monitoring of global land usage, landscape planning, automatic change detection for security monitoring, motion tracking, monitoring of the healing therapy, monitoring of the tumor evolution etc.

3) Multimodal analysis: Images of the same scene are acquired by different sensors[11] to integrate the information obtained from different source streams to gain more complex and detailed scene representation. The examples being fusion of information from sensors with different characteristics, color/multispectral images with better spectral resolution, or radar images independent of cloud cover and solar illumination, combination of sensors recording the anatomical body structure like magnetic resonance image (MRI), ultrasound or CT with sensors monitoring functional and metabolic body activities like positron emission tomography (PET), single photon emission computed tomography (SPECT) or magnetic resonance spectroscopy (MRS).

4) Scene to model registration: Images of a scene and a model of the scene are registered[11]. The model can be a computer representation of the scene, maps or digital elevation models (DEM) in geographic information systems (GIS), another scene with similar content, ‘average’ specimen etc. The aim is to localize the acquired image in the scene/model and/or to compare them. The applications being registration of aerial or satellite data into maps or other GIS layers, target template matching with real-time images, automatic quality inspection, comparison of the patient’s image with digital anatomical atlases, specimen classification.

3.2. Steps in Image Registration

The majority of the registration methods consist of the following four steps, often associated with an initial step of preprocessing/transformation.

a) Feature Detection: The feature extraction unit is responsible for detecting objects in each frame and tracking them, creating a list of low level items with all the features found [4]. Salient and distinctive objects are manually or, preferably, automatically detected. For further processing, these features can be represented by their point representatives (line endings, distinctive points), which are called control points (CPs).

b) Feature Matching: In this step, the correspondence between the features detected in the sensed image and those detected in the reference image is established. Various feature descriptors and similarity measures along with spatial relationships among the features are used for feature matching.

c) Transform model Estimation: The type and parameters of the so-called mapping functions, aligning the sensed image with the reference image, are estimated[11]. The parameters of the mapping functions are computed by means of the established feature correspondence.

d) Image Resampling and Transformation: The sensed image is transformed by means of the mapping functions. Image values in non-integer coordinates are computed by the appropriate interpolation technique.

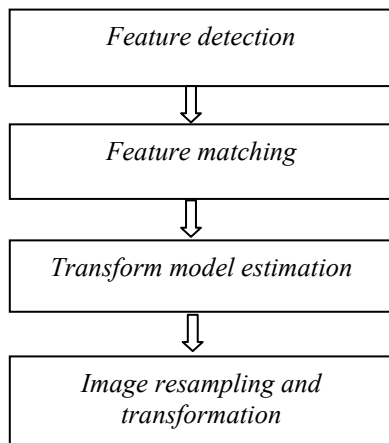


Fig. 1: Steps involved in Image Registration



Fig. 2: Example for four Steps in image registration

Fig. 2 explains the four steps of image registration: feature detection, feature matching by invariant descriptors (the corresponding pairs are marked by numbers), transform model estimation exploiting the established correspondence, image re-sampling and transformation using appropriate interpolation technique.

3.3. Approaches to image registration

Some of the approaches to image registration as mentioned in [9] are Transformations using Fourier Analysis, Cross correlation approach using Fourier Analysis, Sum of squares search technique, Eigen Value Decomposition, Moment matching techniques, Warping Techniques, Procedural approach, Anatomic Atlas, Internal landmarks, External Landmarks.

The registration methods can be broadly divided in to

a) Correlation and Sequential Methods: Cross correlation is the basic statistical approach to registration[10] which gives a measure of degree of similarity between an image and a template. For a template T and image I, where T is small compared to I, the two-dimensional normalized cross-correlation function measures the similarity for each translation.

$$C(u,v) = \frac{\sum_x \sum_y T(x,y)I(x-u,y-v)}{[\sum_x \sum_y I^2(x-u,y-v)]^{1/2}} \quad (1)$$

b) Fourier Methods: Translation, rotation, reflection, distributive and scaling properties of Fourier transform can be exploited for image registration as all these properties have their counterpart in Fourier domain.

c) Point Mapping: The point mapping technique is used to register two images whose type of misalignment is unknown. The general method consists of three stages. Computation of the feature in the image, corresponding feature points (control points) in the reference image with feature points in the data image and spatial mapping using matched feature points.

d) Elastic Model Based Matching: This method model the distortion in the image as the deformation of an elastic material.

4. Implementation Trends in Surveillance Application

Video surveillance systems are widely used nowadays. Video and image processing typically require very high computational power especially in digital video surveillance systems which make them reliable and effective alternative to traditional analog systems[3]. In addition to offering advanced video compression techniques like H.264 these systems can be augmented with algorithms, such as stabilization, panorama, and video motion detection. Typically the system can be realized in platforms such as DSPs, FPGAs, ASICs, or the combination of them. Compared to DSP, FPGA provides much higher performance and flexibility for parallel processing and pipelined operations[6] and effectively reduces the design complication, development time and ensures the real-time encoding and decoding procedure[5]. The system can utilize Ethernet as its connection between encoding side and decoding side, thus could be easily expanded to multipoint-to-point or multipoint-to-multipoint Ethernet services that would be required by different applications[8].

For networked real time video surveillance system with real time video capture, video encoding and decoding and network transmission, FPGA based system is suitable. Since the FPGA-based system architecture is very flexible and easy reconfigurable, the system can also be viewed as a system level video algorithm verification platform and is suitable to be used for H.264 and future video codec standards.

5. Accuracy Tests in Video Surveillance Applications

Registration accuracy is controlled mainly by two parameters: the number of control points selected in each frame and the tolerance used to find correspondence between control points in consecutive frames. Obtained registration accuracy also depends on frame rate of the camera as well as the distance of the camera to the scene and the variation in scene. Higher the accuracy, the higher the reliability will be. To achieve highly reliable registrations, there is a need to set the parameters of the system to obtain a highly accurate registration, which translates to a higher computational cost or a lower speed.

More and more industrial applications appeal to artificial vision to implement, analyze, acquire, and understand images. In modern surveillance applications, artificial vision is mandatory. The producers of electronic devices have launched a large number of application specific chips, but a real appreciation of their performance, compared with the claimed performance, is rarely seen [12].

Two types of accuracy test generally used for acquisition accuracy tests: static tests and online tests. Static tests [7] are performed on static images. This is useful for applications implying static image analysis like granulometry and cell counting. The average mean square error is estimated by the average sample mean-square error in the given image is given by the equation below. where : M -mean value of the image, $x(i, j)$ -grey value of pixel located in coordinates (i, j) , N - spatial resolution of the acquisition section (resolutions on x and y-axis are considered the same).

$$N_{ms}^2 \equiv \frac{\sum_{i=0}^N \sum_{j=0}^N (x(i, j) - M)^2}{N \times N} \tag{2}$$

Online tests [12] are useful in applications related with image series processing (surveillance, remote monitoring). S/N ratio is computed using the standard formula given below. Where: A_s -maximum amplitude of the signal, with values in 16-235 interval (according to ITU 601), U_N -effective value of the noise, U_{AVE} -average value, x_i - value of the i^{th} sample during noise estimation, i is variable from 0 to $N-1$.

$$S/N(dB) = 20 \log \left(\frac{A_s}{U_N} \right) \tag{3}$$

$$U_N = \sqrt{\frac{1}{N-1} \sum_{i=0}^{N-1} (x_i - U_{AVE})^2} \tag{4}$$

The calculations use statistical properties of the signal and they have measurement errors[12] leading to slight error between the estimated value of noise and the real value. The statistical error in this case is given by $E = U_N/\sqrt{N}$, where N is the number of samples.

6. Open Issues and Future Trends

Real-world automatic video-surveillance is a challenging task. Most of the present available algorithms fail to satisfy the requirements of real-time operation, flexibility, cost effectiveness and robustness to varying environmental conditions. Scene-independent algorithms such as rule-based, k-means and Adaboost show low performance due to variations in camera view and camera angle[13]. On the other hand, most scene-dependent methods require a long training process that is practically difficult for real-world deployment.

Image registration is a key stage in video surveillance and other imaging systems. Although a lot of work has been done, automatic image registration still remains an open problem. Registration of images with complex nonlinear and local distortions, multimodal registration, and registration of N-D images (where $N > 2$) belong to the most challenging tasks at this moment.

7. Conclusions

Video surveillance systems require supporting multiple channels resulting in high compression ratio, sensitive to moving objects and providing automatic alarming function. Efficient algorithms supported by hardware for fast computing is to meet the requirements are essential. The emphasis on FPGA coprocessor system architecture, complemented by software tools and implementing high performance DSP algorithms in an effective manner improves the overall performance. Parallel and array processing capabilities of FPGAs make them an attractive implementation option for highly repetitive tasks found in video and image processing.

8. References

- [1] Yan-Hai SHEN, Jin-Lan LIU. A Novel Coding Architecture for Digital Video Surveillance Applications. Proceedings of ICCT. 2003.
- [2] Weijun Zhang, Yujie Dun, Weixiang Shi. The Design and Implementation of a Networked Real-time Video Surveillance System Based on FPG. IET 2008.
- [3] P. NirmalKumar. Enhanced Performance of H.264 using FPGA Coprocessors in video surveillance. IEEE 2010.
- [4] Ginés Benet, José E. Simó, Gabriela Andreu-García, Juan Rosell and Jordi Sánchez. Embedded Low-Level Video Processing for Surveillance Purposes. IEEE 2010.
- [5] Liang Longfei and Yu Songyu. Real-Time Duplex Digital Video Surveillance System and its Implementation with FPGA. IEEE 2001.
- [6] Bei Na Wei, Yu Shi, Getian Ye, Jie Xu. Developing a Smart Camera for Road Traffic Surveillance. IEEE 2008.
- [7] Rakesh Kumar, Supun Samarasekera, Steve Hsuyanlin Guo, David Hirvonen, Michael Hansen, And Peter Burt. Aerial Video Surveillance and Exploitation. IEEE 2001.
- [8] R. Arsinte, C. Miron. Acquisition Accuracy Evaluation in Visual Inspection Systems - A Practical Approach. Proceedings of the Symposium on Electronics and Telecommunications 1996.
- [9] J.V.Chapnick, M.E.No, G.Q. Maguire, E.L.Kramer, J.J.Sanger, B.A.Birnbaum, A.J.Megibow. Techniques of Multimodality Image Registration. Proceedings of the IEEE 1993.
- [10] Lisa Gottesfeld Brown. A Survey of image registration techniques. ACM Computing Surveys (CSUR) Volume 24 Issue 4, Dec. 1992.
- [11] Barbara Zitova, Jan Flusser. Image registration methods: a survey. Image and Vision Computing 2003.
- [12] R. Arsinte, Z. Baruch. Methods for Accuracy Tests in Industrial Vision and Video Surveillance Applications. IEEE 2006.

[13] S Boragno, B Boghossian, D Makris, S Velastin. Object Classification for Real-Time Video-Surveillance Applications. IET 2008.