International Journal of Energy, Information and Communications
Vol.4, Issue 6 (2013), pp.21-28
http://dx.doi.org/10.14257/ijeic.2013.4.6.03

Sparse Representation based Target Detection in Infrared Image

Shujuan Gao, Insuk Kim and Seong Tae Jhang*

Department of Computer, The University of Suwon, Hwaseong-si Gyeonggi-do445-743, Korea

gsh360@hotmail.com, saintboy@suwon.ac.kr, stjhang@suwon.ac.kr

Abstract

Infrared target detection research has recently attracted much attention, especially in pedestrian detection. In this paper, we use compressive sensing theory, a new and emerging field, for an infrared target detection system. Based on the framework of Bayesian filter, target is expressed by sparse representation. Compressive sensing (CS) is based on the illusion that a small quantity of non-adaptive linear projection of a compressible signal contains sufficient information for signal reconstruction and processing. In this paper we give an overview of compressive sensing and our proposed method. The method firstly construct appearance model using features extracted from the data independent image feature space. The appearance model can preserve structure targets' feature since it adopts non-adaptive random projections. The experiment results show that the proposed infrared target detection system is feasible and efficient.

Keywords: *sparse representation; target detection; infrared image; night vision*

1. Introduction

Night vision is the ability to see in low illumination environment. Driven by the decreasing cost of infrared cameras in recent years, more and more interest has focused on night vision technology. Target detection in infrared images has been studied in the context of video surveillance. The lack of texture and shape information makes tracking targets in infrared image even harder.

Infrared light can be split into three categories: Near-infrared (NIR) close to visible light, has wavelengths that range from 0.7 to 1.3 microns, or 700 billionths to 1,300 billionths of a meter; Mid-infrared (MIR) has wavelengths ranging from 1.3 to 3 microns. Both NIR and MIR are used by a variety of electronic devices, including remote controls; Thermal-infrared (TIR) is also known as FIR, occupying the largest part of the infrared spectrum, TIR (or FIR) has wavelengths ranging from 3 microns to 30 microns [1]. For night vision technology FIR and NIR images are provided by Far and Near Infrared cameras. Both of them have distinct characteristic that are introduced in [2]. Infrared images are nosier than photographic images. An infrared image may include large smooth regions, which lack first- rate details, but there exits greater scope of data compression. Infrared images also can provide information that is not accessible in visual images.

Infrared technology at present is widely applied in public traffic surveillance and guidance [3], medical imaging [4], detection of high dangerous areas [5], astronomy, robotics and the military target recognition and tracking, and also achieved a great development. However in target detection and tracking in infrared imagery and video there are still many

* *Corresponding Author*

ISSN: 2093-9655 IJEIC

problems existed. This is due to a number of key factors like lighting changes (shadow vs. sunny day, indoor/night vs. outdoor), cluttered backgrounds (trees, vehicles, animals), artificial appearances (clothing, portable objects), non-rigid kinematics of pedestrians, camera and object motions, depth and scale changes (child vs. adult), and low video resolution and image quality [6]. Some problems are coped with different forms of sensors such as cameras, radar or both of them. For object detection technology there are mainly three kinds of approaches: background subtraction approach, frame difference approach, and optical flow approach. The approach of background subtraction is sensitive to variance in background and illumination changes. Frame difference approach extracts moving object by differences between two or three consecutive frames. This approach is most simple and quick compare to the others. But this approach only extracts object with relative movement. Optical flow is a famous approach for object tracking but this method is extremely sensitive to illumination changes or noise.

In our research, we are eager to develop an energy efficient infrared target detection system. To deal with the low contrast noises when tracking targets in infrared images, an algorithm based on sparse representation model is proposed which inspired from a visual surveillance system [7].

2. Compressive Sensing

Compressive sensing method integrated decomposition of matrixes format with image frames of surveillance video into low rank and sparse matrix is put forward to segment the moving targets and the background in a surveillance video. In compressive sensing, the encoding is theoretically a matrix-by-vector multiplication. To what degree a video can be recovered from compressive measurements depends on the sparsity of the signal (after transform) and the properties of measurement matrix. It is also known that if the measurements matrix satisfies the restricted isometry property (RIP), the original signal can be recovered from the measurements if the number of measurements is large enough [8].Traditional signal acquisition systems are usually based on transform coding. All samples of a signal are captured in acquisition procedure and the whole set of coefficients is computed. Large number of coefficients contain most energy of signal can be retained, but other coefficients are discarded. It also claims that the sampling rate must be at least twice the maximum frequency present in the signal. The sampling data have to be compressed for efficient storage, transmission and processing. In 2006, Candès reported a novel sampling theory: compressive sensing, also called compressive sampling (CS) [9, 10]. The theory asserts that one can recover signals and images from far fewer samples or measurements.

Background

We first briefly review some preliminaries of compressive sensing which are used in the proposed method.

Any $N\times1$ signal x in R^N space can be represented in an orthonormal basis $\Psi=[\psi_1\psi_2\cdots\psi_n]$ as follows [11],

$$x=\sum_{i=1}^{N}s_i\psi_i\ , \tag{1}$$

where s_i is the coefficient sequence of signal x. Equivalent form of signal x is $x=\Psi s$, where Ψ is the $N\times N$ matrix with $\psi_1\psi_2\cdots\psi_n$ as columns. Many real world signals can be approximated by sparse or compressible with a suitable basis. Clearly, x and s are equivalent description of signal, x in the space or time domain and s in the Ψ domain. Signal x is K-sparse if it is a linear combination of only k basis vectors; that means K of the s_i coefficients are nonzero and N-K are zero.

Compressive sensing theory has two basic principles: sparsity and incoherence, or sparsity and restricted isometry property. Sparsity means that information rate of signals can be smaller than implied by its bandwidth. If there are just a few large coefficients and many small in some basis representation, signal x is called a compressive signal.

3. Sparse Random projection

The key idea of random projection comes from the Johnson-Lindenstrauss lemma[12], uses a random $d \times k$ matrix R whose rows have unit lengths to project original high-dimensional Euclidean space to a k-dimensional ($k << d$) Euclidean subspace. Johnson-Lindenstrauss lemma: in a vector space if points are projected onto a randomly selected subspace of suitably high dimension then the distances between the points are approximately preserved.

Computation: There are random vectors $X_i \, (i = 1, 2, \cdots, N)$ in d-dimensional space, linear project X_i to k-dimensional Euclidean space vectors with random matrix R:

$$Y = XR \tag{2}$$

where R is $d \times k$ dimensional random matrix. If the random matrix R satisfies the Johnson-Lindenstrauss lemma, we can reconstruct X with minimum error from Y with high probability if X is compressive. Almost all the information in X preserved in Y.

The selection of random matrix is a key point. Gaussian distributed is often used recently, however, this matrix is dense when d is large the computation will be very complex. In the proposed method, we use a very sparse matrix that can compute efficiently. That is a simple measurement matrix defined as:

$$r_{ij} = \sqrt{3} \times \begin{cases} +1 & \text{with probability } 1/6 \\ 0 & \text{with probability } 2/3 \\ -1 & \text{with probability } 1/6 \end{cases} \tag{3}$$

Achlioptas has proved that this very sparse matrix satisfies the Johnson-Lindenstrauss lemma [13]. This very sparse matrix is very easy to compute, as it requires only a uniform random generator. It is obviously that two thirds of computation can be ignored. We use $s = d / 4$ to construct a sparse matrix. For each row of matrix not more than 4 entries are computed. Therefore, only the nonzero entries of matrix need to be stored.

4. Infrared Target Detection Algorithm

We present the tracking algorithm in details in this section. We got a numerous number of patches from original images beforehand. The truth is that we start tracking from the second frame. We assume that in the first frame target has been determined. In every frame, positive and negative patches are sampled near current target's location, and these patches will be used for updating the classifier with maximal classification values. In order to predict the target position in the next frame, we draw some samples around the current target position and determine the one with the maximal classification scores.

Actually, target detection or recognition via sparse representation has obtained significant performance improvement because of its excellent properties. The choice of features is no longer critical; the important thing is whether the number of feature is large enough and whether the sparse representation is correctly computed. Sparse representation for features: In order to compressive a high dimensional original vector $x \in \mathrm{R}^m$ to a low dimensional vector y $\in \mathrm{R}^n$, the dimensionality m is usually in range of 10^6 to 10^{10}, we use a random measurement matrix $R \in \mathrm{R}^{n\times m}$

$$y = Rx \tag{4}$$

In the matrix R, values of -1, 1, 0 represent the negative, positive and zero entries (represent as dark, gray and white rectangles) respectively. One of the nonzero entries in matrix R sensing to x_i is the same as a filter convolving the intensity at a determinate area of an original input image. This sparse representation ideology makes that the features can preserve enough or almost all of the necessary information of the original input image. For that reason the proposed method works efficiently in stage of feature extraction.

Classification: After collecting the low dimensional feature sets $y = (y_1, y_2, \ldots y_n) \in \mathrm{R}^n$, assume all elements of y are distributed independently and model representation sets with Bayes classifier

$$H(y) = \log\left(\frac{\Pi_{i=1}^{n} p(y_i|q=1)\,p(q=1)}{\Pi_{i=1}^{n} p(y_i|q=0)\,p(q=0)}\right) = \sum_{i=1}^{n} \log\left(\frac{p(y_i|q=1)}{p(y_i|q=0)}\right) \tag{5}$$

where we assume uniform prior, $p(q=1) = p(q=0)$ and $q \in \{0,1\}$ is a binary variable which represents the sample label.

Random projections of high dimensional random vectors are almost Gaussian. $p(y_i|q=1)$, $p(y_i|q=0)$ of $H(y)$ (5) are assumed that it is Gaussian distributed with parameters of $(\mu_i^1, \sigma_i^1, \mu_i^0, \sigma_i^0)$ where

$$p(y_i|q=1) \sim N(\mu_i^1, \sigma_i^1),\ p(y_i|q=0) \sim N(\mu_i^0, \sigma_i^0) \tag{6}$$

Scalar parameters of (6) updated by

$$\mu_i^1 \leftarrow \lambda \mu_i^1 + (1-\lambda)\mu$$

$$\sigma_i^1 \leftarrow \sqrt{\lambda(\sigma_i^1)2 + (1-\lambda)(\sigma^1)2 + \lambda(1-\lambda)(\mu_i^1 - \mu^1)^2} \tag{7}$$

where λ is learning parameter, $\sigma^1 = \sqrt{\frac{1}{n}\sum_{k=0|q=1}^{n-1}(y_i(k) - \mu^1)^2}$ and $\mu^1 = \frac{1}{n}\sum_{k=0|q=1}^{n-1} y_i(k)$

5. Experiment Result

We provide some experiments to validate the advantage of the proposed method over the classic method of background subtraction in this section. Our algorithms are run on a PC with 3.10 GHz CPU and 8G memory. Software platform is Matlab 2012a.

In order to validate the performance of our target detection system based on infrared images, we have tested two different infrared images sequences that contain different kinds of target. Dataset 1 consists of 80 frames (360*240 pixels) with a pedestrian. Dataset 2 is provided within the OTCBVS Benchmark Dataset Collection consisting of 131 frames (320*256 pixels) with one car.

The processed sequence is shown in following figures.

Figure 1. Experiment result of Dataset 1

Figure 2. Experiment result of Dataset 2

Table 1. Processing time of Infrared Target Detection Algorithm

Dataset	Image Size	Frame Number	Total Run time(s)	Frame Per Second (fps)
1	360*240	80	10.0469	7.96265
2	320*256	131	27.7500	4.720720

Many infrared target segmentation algorithms rely on background subtraction techniques. A background subtraction is initially performed to identify local foreground objects. Then gradient information from the objects and the background is combined to generate a contour map representing the confidence for each pixel to belong to a person's boundary [14]. We compared the performance of background subtraction based segmentation in infrared images with our method in infrared images.

Table 2. Results of Background-subtraction based pedestrian detection algorithm

Dataset	Image Size	Frame Number	Total Run time(s)	Frame Per Second (fps)
1	360*240	80	36.0236	2.22076
2	320*256	131	68.9473	1.90000

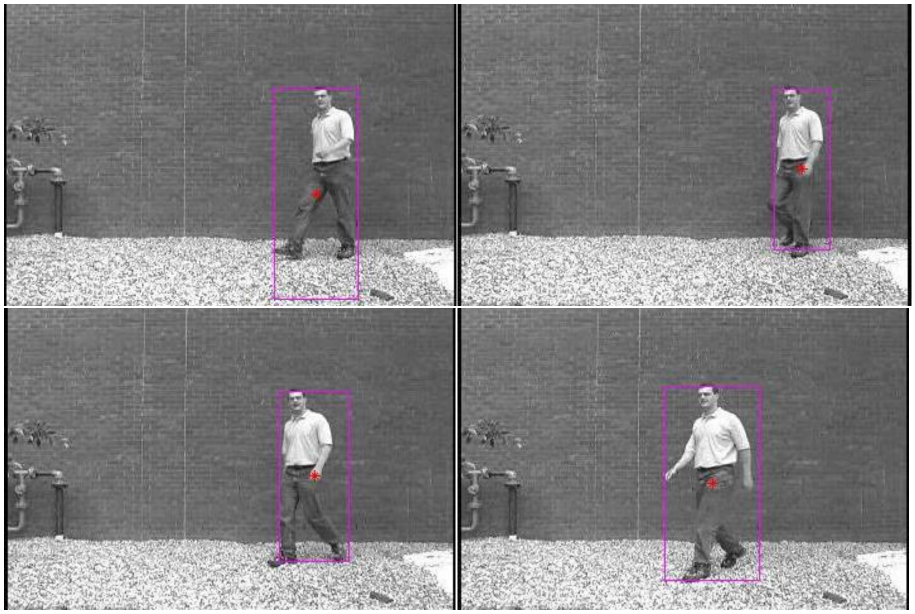

Figure 3. Results of some experimental images for Background-subtraction based pedestrian detection system

Our experimental result turns out that the method can meet the requirement of real time ability of infrared target tracking system. Our method achieves a speed that is much faster than that of the method based on the method of background subtraction. However, a current limitation of our approach is that we haven't resolved the problem of scale changing. Detection accuracy is not very good in some test frames since the target scale changes when it is moving on. We are eager to solve it the near future.

6. Conclusion and Future work

In order to deal with the conflict between complexity and image quality of infrared imaging systems, we have applied the theory of compressive sensing to research the infrared imaging systems. A measurement with less elements of the target can be obtained. The proposed target detection system can work efficiently. So far super-resolution can be implemented by the sensor with low resolution. Some researches show that super-resolution imaging can be realized by compressive sensing theory, high resolution information can be sampled by low-resolution sensors, and the mean square error was low between the reconstructed image and original image.

Our experimental results show that the proposed approach is efficient and rapid in target detection. Currently, in our work a real-time target detection based on infrared camera system is under development. In the future we will pay more attention on improving the accuracy and speed of target identification technology.

Acknowledgements

This work was supported by GRRC program of Gyeonggi province [(GRRC SUWON 2013B1) Cooperative CCTV Image Based Context-Aware Process Technology] and by Business for Cooperative R&D between Industry, Academy, and Research Institute funded Korea Small and Medium Business Administration (Grants No. C0094876) in 2013.

References

[1] O. Mahfooz and A. Shah, "Revisiting Far/Near Infrared Pyramid-Based Fusion Types for Night Vision Using Matlab", **(2011).**

[2] R. K. Parmar, "Night Vision Technology", A seminar report on Computer science and engineering in School of Engineering Cochin University, **(2008).**

[3] Y. Iwasaki, S. Kawata and T. Nakamiya, "Vehicle Detection Even in Poor Visibility Conditions Using Infrared Thermal Images and Its Application to Road Traffic Flow Monitoring", in in Lecture Notes in Electrical Engineering, vol. 151, no. 85, New York, NY: Springer New York, **(2012)**.

[4] B. B. Lahiri, S. Bagavathiappan, T. Jayakumar and J. Philip, "Medical applications of infrared thermography: A review", Infrared Physics and Technology, vol. 55, no. 4, **(2012)**, July, pp. 221–235.

[5] M. Locatelli, E. Pugliese, M. Paturzo, V. Bianco, A. Finizio, A. Pelagotti, P. Poggi, L. Miccio, R. Meucci and P. Ferraro, "Imaging live humans through smoke and flames using far-infrared digital holography", Optics Express, vol. 21, no. 5, **(2013)**, pp. 5379–5390.

[6] T. T. Zin, H. Takahashi and T. T. A. H. Hama, "Fusion of Infrared and Visible Images for Robust Person Detection", **(2011)**, January, pp. 1–26.

[7] H. Li, C. Shen and Q. Shi, "Real-time visual tracking using compressive sensing", presented at the 2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), **(2011)**, pp. 1305–1312.

[8] H. Jiang, W. Deng and Z. Shen, "Surveillance video processing using compressive sensing", Inverse Probl. Imaging, vol. 6, no. 2, **(2012)**, pp. 201–214.

[9] E. J. Candès, "Compressive sampling", presented at the Proceedings on the International Congress of Mathematicians, **(2006)**.

[10] E. J. Candès, J. Romberg and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information", IEEE Trans. Inform. Theory, vol. 52, no. 2, **(2006)**, pp. 489–509.

[11] M. Kang and M. Kang, "Compressive sensing and applications", Asia Pac. Math. Newsl, vol. 2, no. 2, **(2012)**, pp. 1–5.

[12] J. Li, W. Gong, J. Yang and W. Li, "Hybrid Classification Features-based Real-time Pedestrian Detection in Far-infrared Images", China: Opto-Electronic Engineering, vol. 36, no. 2, **(2009)**, pp. 55–61.

[13] D. Achlioptas, "Database-friendly random projections: Johnson-Lindenstrauss with binary coins", Journal of Computer and System Sciences, vol. 66, no. 4, **(2003)** June, pp. 671–687.

[14] J. W. Davis and V. Sharma, "Robust detection of people in thermal imagery", vol. 4, **(2004)**, pp. 713–716.