Computer Vision and Image Understanding 117 (2013) 1659–1671

Contents lists available at SciVerse ScienceDirect

Computer Vision and Image Understanding

journal homepage: www.elsevier.com/locate/cviu

Finite asymmetric generalized Gaussian mixture models learning for infrared object detection

Tarek Elguebaly [a], Nizar Bouguila [b,*]

[a] Electrical and Computer Engineering (ECE), Concordia University, Montreal, QC, Canada
[b] Concordia Institute for Information Systems Engineering (CIISE), Concordia University, Montreal, QC, Canada

ARTICLE INFO

Article history:
Received 18 July 2012
Accepted 11 July 2013
Available online 25 July 2013

Keywords:
Infrared
Multiple target tracking
Pedestrian detection
Foreground segmentation
Image fusion
Mixture models
Asymmetric generalized Gaussian
EM
MML

ABSTRACT

The interest in automatic surveillance and monitoring systems has been growing over the last years due to increasing demands for security and law enforcement applications. Although, automatic surveillance systems have reached a significant level of maturity with some practical success, it still remains a challenging problem due to large variation in illumination conditions. Recognition based only on the visual spectrum remains limited in uncontrolled operating environments such as outdoor situations and low illumination conditions. In the last years, as a result of the development of low-cost infrared cameras, night vision systems have gained more and more interest, making infrared (IR) imagery as a viable alternative to visible imaging in the search for a robust and practical identification system. Recently, some researchers have proposed the fusion of data recorded by an IR sensor and a visible camera in order to produce information otherwise not obtainable by viewing the sensor outputs separately. In this article, we propose the application of finite mixtures of multidimensional asymmetric generalized Gaussian distributions for different challenging tasks involving IR images. The advantage of the considered model is that it has the required flexibility to fit different shapes of observed non-Gaussian and asymmetric data. In particular, we present a highly efficient expectation–maximization (EM) algorithm, based on minimum message length (MML) formulation, for the unsupervised learning of the proposed model's parameters. In addition, we study its performance in two interesting applications namely pedestrian detection and multiple target tracking. Furthermore, we examine whether fusion of visual and thermal images can increase the overall performance of surveillance systems.

1. Introduction

Security of human lives and property has always been a major concern. Nowadays, developing video surveillance systems aimed at monitoring private and public areas has became one of the most active research fields due to the high amount of theft, accidents, terrorists attacks and riots. However, human attention is known to drop after just 30 min when engaged in monotonous and repetitive activities [1]. This is the case for security personnel tasked to monitor relatively vast environments where suspicious events are rare. Therefore, automatic video surveillance techniques were proposed to allow automatic processing of the data acquired by surveillance cameras without requiring the continuous attention of human operators. Automatic video surveillance systems are employed in controlled and uncontrolled environments [2]. In controlled or indoor environments (i.e. airports, warehouses, and production plants) monitoring is easier to implement as it does not depend on weather changes [3,4]. Uncontrolled environment is used to refer to outdoor scenes where illumination and temperature changes occur frequently, and where various atmospheric conditions can be observed [4,5].

Normally, when setting up a security system there are two major types of security cameras: visual-light, and infrared sensors. Visual-light or color cameras are employed vastly due to their lower cost compared to infrared sensors [6,7]. However, under low illumination sensing in visible spectrum becomes infeasible [8]. Thermal IR sensors measure the emitted heat energy from different objects, which make it invariant to changes in ambient illumination. Hence, IR imaging is a perfect choice for monitoring under low illumination conditions or even in darkness [9]. In order to show that thermal IR offers a promising alternative to visible imagery we will use it for pedestrian detection. Despite its robustness to illumination changes, IR has various drawbacks. One of its disadvantages is its sensitivity to outdoor temperature changes, which make it vulnerable to cold or warm air [10,11]. Some researchers decided to use both visible and infrared images together in order to increase the efficiency of surveillance systems [12,13]. It is widely known in the field of image fusion that the combination

* Corresponding author.
E-mail addresses: t_elgue@encs.concordia.ca (T. Elguebaly), nizar.bouguila@concordia.ca (N. Bouguila).

1077-3142/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.cviu.2013.07.007

of thermal infrared and visible images is not trivial. Fusion techniques can be grouped into two classes: representative and analytical. Representative fusion uses both visible and infrared features together in order to generate a new image more informative or intuitive for a human observer. It is important to understand that the generation of such an image can be of a great importance in the case of human monitoring and is not required for automated video monitoring applications. On the other hand, analytical fusion combines available information from both sensors for a more robust analysis and interpretation of the image or video content. This method is based on the idea that combining both thermal and visible information can overcome the disadvantages of both visible-light images (i.e. shadows problem, sensitivity to variations in illumination and lights) and infrared images (i.e. sensitivity to outdoor temperature changes).

Discovering and finding valuable information and patterns in multidimensional data depends generally on the selection of an appropriate statistical model and the learning of its parameters. In recent years a lot of different algorithms were developed in the aim of automatically learning to recognize complex patterns, and to produce intelligent decisions based on observed data. Finite mixture models are now among the most widely used statistical approaches in many areas and applications and allow a formal approach for unsupervised learning. In such context, classic interest is often related to the determination of the number of clusters (i.e. model selection) and the estimation of the mixture's parameters. The isotropic nature of the Gaussian distribution, along with its capability to represent the data compactly by a mean vector and covariance matrix, has made Gaussian mixture (GM) decomposition a popular technique. However, Gaussian density has some drawbacks such as its symmetry around the mean and the rigidity of its shape, which prevent it from fitting accurately the data especially in the presence of outliers. Fig. 1 shows an example of an IR image. We can notice that its intensity distribution is not symmetrical. It is clear that using the GM to represent this distribution is not efficient. In order to overcome problems related to the Gaussian assumption, some researchers have shown that the generalized Gaussian distribution (GGD) can be a good choice to model non-Gaussian data [14,15]. Compared to the GD, the GGD has one more parameter λ that controls the tail of the distribution: the larger the value of λ is, the flatter is the distribution; the smaller λ is, the more peaked is the distribution. Despite the higher flexibility that GGD offers, it is still a symmetric distribution inappropriate to model non-symmetrical data. In this article, we suggest the consideration of the asymmetric generalized Gaussian distribution (AGGD) capable of modeling non-Gaussian asymmetrical data. The AGGD uses two variance parameters for left and right parts of the distribution, which allow it not only to approximate a large class of statistical distributions (e.g. impulsive, Laplacian, Gaussian and uniform distributions) but also to include the asymmetry. As shown in Fig. 1(b) we can notice that the asymmetric generalized Gaussian mixture (AGGM) was able to accurately model the data and outperforms both the GM and the generalized Gaussian mixture (GGM).

An important part of the mixture modeling problem concerns learning the model parameters and determining the number of consistent components (M) which best describes the data. For this purpose, many approaches have been suggested. The vast majority of these approaches can be classified, from a computational point of view, into two classes: deterministic and stochastic methods. Deterministic methods, estimate the model parameters for different range of M then choose the best value that maximizes a model selection criterion such as Akaike's information criterion (AIC) [16], minimum description length (MDL) [17] and Laplace empirical criterion (LEC) [18]. Stochastic methods such as Markov chain Monte Carlo (MCMC) can be used in order to sample from the full *a posteriori* distribution with M considered unknown [19]. Despite their formal appeal, MCMC methods are too computationally demanding, therefore cannot be applied efficiently for online applications such as automatic video surveillance. For this reason, we are interested in deterministic approaches. In our proposed method, we use K-means algorithm to initialize the asymmetric generalized Gaussian mixture parameters and successfully solve the initialization problem. The number of mixture components is automatically determined by implementing MML criterion [20] into an EM algorithm based on maximum likelihood (ML) estimation. Our learning method can integrate simultaneously parameter estimation and model selection in a single algorithm and is consequently totally unsupervised. It is noteworthy that the proposed work is completely different from recent efforts published for instance in [21–23]. In fact, [21] proposed the use of the gradient and the ML methods for estimating the parameters of only a one-dimensional AGGD. The work in [22] has been devoted to image segmentation using ML estimation of one-dimensional AGGM with known number of components. In [23] a Bayesian nonparametric approach based on infinite GGM was developed for pedestrian detection and foreground segmentation.

The rest of this paper is organized as follows. Section 2 describes the AGGM model and gives a complete learning algorithm. In Section 3, we assess the performance of the new model for pe-

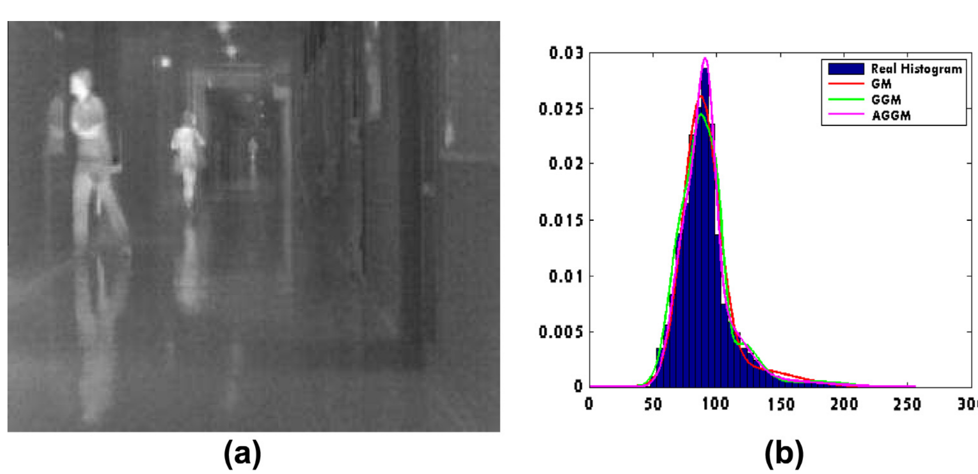

Fig. 1. (a) IR image. (b) Real and estimated (using GM, GGM and AGGM) histograms for the IR image.