Computer Vision and Image Understanding 154 (2017) 64–72

Contents lists available at ScienceDirect

Computer Vision and Image Understanding

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

Segmentation of clusters by template rotation expectation maximization

Carl-Magnus Svensson[a,*], Karen Grace Bondoc[b], Georg Pohnert[b], Marc Thilo Figge[a,c]

[a] Applied Systems Biology, Leibniz Institute for Natural Product Research and Infection Biology,Hans-Knöll Institute, Jena, Germany
[b] Institute for Inorganic and Analytical Chemistry, Bioorganic Analytics, Friedrich Schiller University, Jena, Germany
[c] Friedrich Schiller University, Jena, Germany

ARTICLE INFO

Article history:
Received 13 January 2016
Revised 1 July 2016
Accepted 10 August 2016
Available online 10 August 2016

Keywords:
Segmentation
Generative model
Expectation maximization
Template matching

ABSTRACT

To solve the task of segmenting clusters of nearly identical objects we here present the template rotation expectation maximization (TREM) approach which is based on a generative model. We explore both a general purpose optimization approach for maximizing the log-likelihood and a modification of the standard expectation maximization (EM) algorithm. The general purpose approach is strict template matching, while TREM allows for a more deformable model. As benchmarking we compare TREM with standard EM for a two dimensional Gaussian mixture model (GMM) as well as direct maximization of the log-likelihood using general purpose optimization. We find that the EM based algorithms, TREM and standard GMM, are faster than the general purpose optimizer algorithms without any loss of segmentation accuracy. When applying TREM and GMM to a synthetic data set consisting of pairs of almost parallel objects we find that the TREM is better at segmenting those than an unconstrained GMM. Finally we demonstrate that this advantage for TREM over GMM gives significant improvement in segmentation of microscopy images of the motile unicellular alga *Seminavis robusta*.

1. Introduction

Template matching is a general term for matching an object of interest with known features or shapes to a more complicated scene where the object may be present. Different variants of template matching are widely used in optical character recognition, medical image segmentation, face detection, and action recognition (Gupta et al., 2014; Jain et al., 1998). For a template to be matched in the scene there is normally a large number of possible transformations to be considered such as object scaling, rotation and translations. This means that the computational cost of the matching normally is high and to match numerous objects in different configurations leads to a combinatorial explosion (Korman et al., 2013; Ouyang et al., 2012).

This paper will demonstrate the use of the *template rotation expectation maximization* (TREM) approach for segmentation of individual objects in clusters. By assigning the spatial position of single pixels to terms of a two-dimensional *Gaussian mixture model* (GMM), which includes prior information about how the objects look, we will separate clusters of objects with a known shape. Compared to traditional unconstrained GMMs trained by expectation maximization (EM) (Bishop, 2007), which are used to separate clusters of polymorphonuclear neutrophils (Brandes et al., 2015), we will demonstrate the that TREM is better in resolving clusters of objects with known shape. When compared with other template matching algorithms, which often rely on general purpose optimizers (Bhagya et al., 2014; Ouyang et al., 2012), we will show that TREM is considerably faster without any loss of segmentation performance.

While TREM is generally applicable to different areas of computer vision it is of special interest for automated image analysis of microscopy images, where the occurrence of many identical objects that need to be segmented occurs regularly (Medyukhina et al., 2015). These objects are not seldom tightly clustered to an extent that standard segmentation techniques, such as thresholding or watershed, becomes extremely challenging. For template matching in general, an approximate match often is sufficient (Jain et al., 1998) but is often not useful in microscopy images as we require high resolution, especially when tracking the segmented object. It is also the case that template matching normally breaks down when objects are touching (Zimmer, 2012). For object segmentation in microscopy images EM has been combined with template matching in semi-automated algorithms to extract features

* Corresponding author.
E-mail addresses: carl-magnus.svensson@leibniz-hki.de (C.-M. Svensson), thilo.figge@leibniz-hki.de (M.T. Figge).

http://dx.doi.org/10.1016/j.cviu.2016.08.003
1077-3142/

from high-resolution images of zebrafish (Peravali et al., 2011) and for segmentation of cell nuclei using supervised learning (Chen et al., 2013). For object clusters that efficiently can be skeletonized the combination of EM and template matching the method for untangling *C. elegans* clusters by Raviv et al. (2010) represents an attractive approach. Unsupervised methods that work directly on shapes in binary or grey scale images are constrained to either strongly rod-like (Moss and Hancock, 1997; Zhang et al., 2006) or almost perfectly circular/spherical cells (Wörz et al., 2008). TREM is applicable to both those extremes but also to the whole spectrum of oval and rectangular objects.

For non-biological systems these shape restrictions might seem strongly limiting but in biology there are several bacteria (Enninga et al., 2005; Wang et al., 2010), fungal spores (Munkvold and Neely, 1990) and algae (Chepurnov et al., 2002; Saison et al., 2010; Vanstechelman et al., 2013) that are fulfilling the criteria of shapes that can be well described by a non-circular two dimensional Gaussian function. Because of its probabilistic formulation TREM is robust to small and medium object deformations and moderate affine transformations other than translation and rotation. Here we evaluate two different approaches in fitting a GMM where a fixed template for the shape of the objects is imposed through the definition of a fixed covariance matrix, $\mathbf{\Sigma}_0$.

In this paper we consider the case where the number of objects in clusters, K, is known or approximated using a different method. Naturally, K is known for synthetic data and for microscopy data we manually annotate the number of objects in each cluster. While this pre-requisite implies that task of object enumeration is already solved, TREM attempts to find the most likely configuration of objects within a cluster. This is of importance in object tracking, where faulty segmentation can give false movement patterns of the object (Chen et al., 2006; Yang et al., 2006), or colocalization of objects and functional markers on a sub-cellular level (Rizk et al., 2014). A specific example of the need to correctly identify the position and outline of each object in a cluster is when fluorescence in situ hybridization (FISH) techniques are used to identify abnormal DNA content of cell nuclei (Kallioniemi et al., 1992; Leversha et al., 2009; Pinkel et al., 1988). When using FISH the successful segmentation of each nucleus is a must and this is normally performed using watershed segmentation (Malpica et al., 1997). Supervised template matching has also been suggested as a possible approach when resolving clusters of nuclei for subsequent FISH analysis (Chen et al., 2013). The organism that fulfils the criteria of fixed shape, size, non-rotational symmetric in this study is the algae species *Seminavis robusta*. This is a marine diatom (Chepurnov et al., 2002; Vanstechelman et al., 2013) that is important in the fixing of inorganic carbon (Granum et al., 2005; Vanstechelman et al., 2013) and their motility is studied in connection to environmental variables (Cohn et al., 2003; Cohn and McGuire, 2000; Coquillé et al., 2015). To study the motility it is necessary to first segment the individual alga to then track them and analyse their movement pattern as a function of the environment.

TREM assumes that we already have an appropriate binary representation of the image data. However, no restrictions are placed upon the method used to obtain the binary image. In this paper we will use both fixed threshold for the synthetic data and the more advanced directed acyclic graphical (DAG) continuous max-flow segmentation for the microscopy images (Rajchl et al., 2012; Yuan et al., 2010).

We will here demonstrate the usefulness of TREM by comparing it with a classical GMM. We also consider the case of pure rotation of the template in a hybrid algorithm where classical EM is used to determine the location of objects and general purpose optimization for orientation called *general purpose rotation and expectation maximization translation* (GPR-EMT). Finally we compare TREM to the *direct log-likelihood maximization* (DLLM), where all parameters of the log-likelihood are estimated using a general purpose optimizer. The algorithms are applied to two sets of synthetic data and one microscopy data set. The first synthetic data set is generated to evaluate general performance and runtime of the algorithms. We then move on to a special case of almost parallel objects were we check if the template approach is giving us the expected advantage over traditional GMMs.

2. Methods

2.1. Generative model

In principle, TREM can be applied to higher dimensions than two, for the case of image analysis the case of three dimensions is of interest, but in this paper we will only consider the case of two dimensions for clarity. We assume that the pixels at position $\mathbf{x}_n$ belonging to a foreground object are a realization of the probability function $p(\mathbf{x}_n)$. We formulate the log-likelihood as

$$\mathcal{L}(\theta) = \sum_{n=1}^{N} \log(p(\mathbf{x}_n|\theta)), \tag{1}$$

where N is the total number of foreground pixels in a given cluster. As in other maximum likelihood formulations, we seek the parameters, θ, that maximise the log-likelihood given the data (Bishop, 2007). The generative model of a cluster, where K denotes the total number of objects in the cluster, express the probability that a pixel at position $\mathbf{x}_n$ is part of object k. We assume that this probability is described by

$$p(\mathbf{x}_n|\theta) = \sum_{k=1}^{K} p(\mathbf{x}_n|k,\theta)p(k),$$
$$p(\mathbf{x}_n|k,\theta) = \mathcal{N}(\mathbf{x}_n, \boldsymbol{\mu}_k, \mathbf{\Sigma}_k), \quad p(k) = 1/K.$$

In this model we assume that the prior distribution, $p(k)$, is discretely uniform based on that we are considering near identical objects where the intensities, at least within a cluster, are approximately uniform. The noise model, $p(\mathbf{x}_n|k,\theta)$, is a two dimensional Gaussian governed by the mean, $\boldsymbol{\mu}_k \in \mathbb{R}^2$, and covariance matrix $\mathbf{\Sigma}_k \in \mathbb{R}^{2\times 2}$. For this generative model we have the set of parameters $\theta = \{\boldsymbol{\mu}_k, \mathbf{\Sigma}_k\}$.

2.2. Unconstrained Gaussian mixture model

An established approach to determine the position of cells in a cluster is to fit an unconstrained two dimensional GMM to the cluster (Brandes et al., 2015). Just as in our approach this assumes that the number of objects in the cluster can be well approximated using, for example, the total cluster area. We follow the standard steps of EM (Dempster et al., 1977; Neal and Hinton, 1998) by formulating the variational free energy as

$$\mathbb{F}(\theta, q(k)) = \sum_{n=1}^{N}\left[\sum_{k=1}^{K} q(k)[\log(p(\mathbf{x}_n|k,\theta)) + \log(p(k))]\right] + H(q) \leq \mathcal{L}(\theta). \tag{3}$$

Here, $q(k)$ is an arbitrary probability function and the Shannon entropy of $q(k)$ is defined as $H(q(k)) = -\sum_{n=1}^{N}\sum_{k=1}^{K} q(k)\log(q(k))$. The EM algorithm then alternately increases $\mathbb{F}(\theta, q(k))$ by finding q while fixing the parameters θ (E-step) and then updating the parameters θ while considering $q(k)$ to be constant (M-step). In practice, given a set of parameters θ' we perform the E-step by setting $q(k|\mathbf{x}_n,\theta') = p(k|\mathbf{x}_n,\theta')$, which can be calculated by using the parameters θ' in Eq. (2) and applying Bayes' theorem. During the M step we take the derivative of $\mathbb{F}(\theta, q(k|\mathbf{x}_n,\theta'))$ with regard to the