Computers and Electronics in Agriculture 142 (2017) 326–337

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

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

Original papers

Moment invariants for multi-component shapes with applications to leaf classification

Mohamed Ben Haj Rhouma [a,*], Joviša Žunić [b], Mohammed Chachan Younis [c]

[a] Department of Mathematics, Statistics & Physics, College of Arts and Sciences, University of Qatar, P.O. Box 2713, Doha, Qatar
[b] Mathematical Institute, Serbian Academy of Sciences and Arts, Belgrade, Serbia
[c] Computer Science, University of Exeter, Exeter, UK

A R T I C L E I N F O

Article history:
Received 9 June 2016
Received in revised form 14 April 2017
Accepted 30 August 2017
Available online 25 September 2017

Keywords:
Shape
Multi-component shapes
Moments
Shape invariants
Leaf classification

A B S T R A C T

In this paper we introduce seven new invariants for multi-component shapes, and apply them to the leaf classification problem. One of the new invariants is an area based analogue of the already known boundary based anisotropy measure, defined for the multi-component shapes (Rosin and Žunić, 2011). The other six invariants are completely new. They are derived following the concept of the geometric interpretation (Xu and Li, 2008) of the first two Hu moment invariants (Hu, 1961). All the invariants introduced are computable from geometric moments corresponding to the shape components. This enables an easy and straightforward computation of translation, rotation, and scaling invariants. Also, being area based, the new invariants are robust to noise and mild deformations. Several desirable properties of the new invariants are discussed and evaluated experimentally on a number of synthetic examples. The usefulness of the new multi-component shape invariants, in the shape based object analysis tasks, is demonstrated on a well-known leaf data set.

1. Introduction

Due to the strong and permanently ongoing development in image technology, image based data are becoming easily available in all domains. Different objects are represented by such data, and different computing techniques were employed to analyze them for a wide spectrum of applications: in medicine (Grisan et al., 2008), astrophysics (Lekshmi et al., 2003), geology (Bowman et al., 2001), ecology (Russell et al., 2009), and many more other fields. There are many different techniques and methods to process and analyze images and objects presented on them. In this paper we will be focused on shape based approaches that are very popular in image analysis tasks since they allow for a multitude of numerical characterizations. This is always very suitable for computer supported image data analysis tasks. Many shape properties, herein named *shape descriptors*, are already identified. Perhaps most popular among them are the shape convexity, compactness, linearity, ellipticity, etc. All of them have intuitively clear meaning. This supports a better understanding of the behavior of the computation methods and techniques based on them.

To be employed in object analysis, shape descriptors need methods for their numerical validations. Such defined numerical characteristics of object/shapes are called *shape measures*.[1] By now there are many shape measures that numerically characterize different aspects of shapes. Just to mention a few: convexity (Rahtu et al., 2006; Žunić and Rosin, 2004), compactness (Lee and Sallee, 1970; Žunić et al., 2010), linearity (Gautama et al., 2003; Stojmenović et al., 2008; Žunić and Martinez-Ortiz, 2009), ellipticity (Rosin, 2003; Aktaş and Žunić, 2013), etc. Notice that multiple measures to evaluate a single shape property, are necessary because there is no shape measure that performs well in all applications. A measure can perform well in one application but fails to meet the expectations in another.

In most cases, shape descriptors and their related measures were originally designed for single component objects/shapes. One can of course think of many instances when it could be beneficial to have measures defined for objects/shapes which consist of several components, or can be decomposed into a number of components. Such objects herein will be called *multi-component shapes*. The concept of multi-component shapes, has been introduced in

* Corresponding author.
E-mail addresses: rhouma@qu.edu.qa (M. Ben Haj Rhouma), jovisa_zunic@mi.sanu.ac.rs (J. Žunić), mcy201@exeter.ac.uk (M.C. Younis).

[1] Once a suitable set of shape measures is selected, certain object analysis task can be performed in $\mathbb{R}^n$ space rather than in the object space. In such situations, the objects studied are represented by their corresponding feature vectors, from $\mathbb{R}^n$. The feature vectors have selected shape measures as their components (coordinates).

http://dx.doi.org/10.1016/j.compag.2017.08.029
0168-1699/

Rosin and Žunić (2011) and Žunić and Rosin (2009). The motivation was many-fold. We mention just some of them:

(a) In many situations, several single objects appear and act as a group (e.g. fish shoal (Fig. 1(a)), vehicles on the road, armed forces, etc);
(b) Single shapes could have clear components (see a footprint in Fig. 1(b));
(c) Very often, it is beneficial to consider a single object as a multi-component one, consisting of components defined with respect to some natural criteria (e.g. embryonic tissue displayed in Fig. 1(c), materials micro structures, etc);
(d) The multi-component shape concept can be employed to study the appearance of the same object in a sequence of images (e.g. any 13 consecutive frames, as displayed in Fig. 1(d), represent a "human gait" – as such it can be understood as 13-component shape).

The domain of multi-component shapes is more complex than the domain of the single-component ones. Thus, it is natural to expected some new, not considered before, demands once working with multi-component shapes. Consequently, multi-component shape analysis tools must satisfy some specific properties, not required if working with single-component shapes. For some of these properties, we refer the reader to Rosin and Žunić (2011) and Žunić and Rosin (2009), where a multi-component shape orientation methods have been introduced and analyzed (in sense of specific requests when computing orientation of multi-component shapes). Also, (Rosin and Žunić, 2011) involves a boundary based multi-component shape anisotropy measure. Actually, this is the only shape measure defined for multi-component shapes, so far. Such a measure indicates the degree to which shape components are oriented consistently.

In this paper we continue to develop multi-component shape based tools, applicable in image processing tasks. Our starting point is to develop multi-component shape measures inspired from the well-known Hu moment invariants (Hu, 1961). Even though introduced a long time ago (in 1961), the Hu moment invariants are still a subject of the most recent research and a popular tool in shape based object analysis task. The outcomes of our investigation are:

– Three new multi-component shape measures, motivated with the geometric interpretation (Xu and Li, 2008) of the first Hu moment invariant (Hu, 1961), originally aimed to analyze single-component shapes only.
– Three new multi-component shape measures motivated with the geometric interpretation (Xu and Li, 2008) of the second Hu moment invariant (Hu, 1961) (also, initially developed to deal with single-component shapes).
– An anisotropy measure defined for multi-component shapes. This new measure is adopted from the anisotropy measure developed in Rosin and Žunić (2011) for curves consisting of several parts.

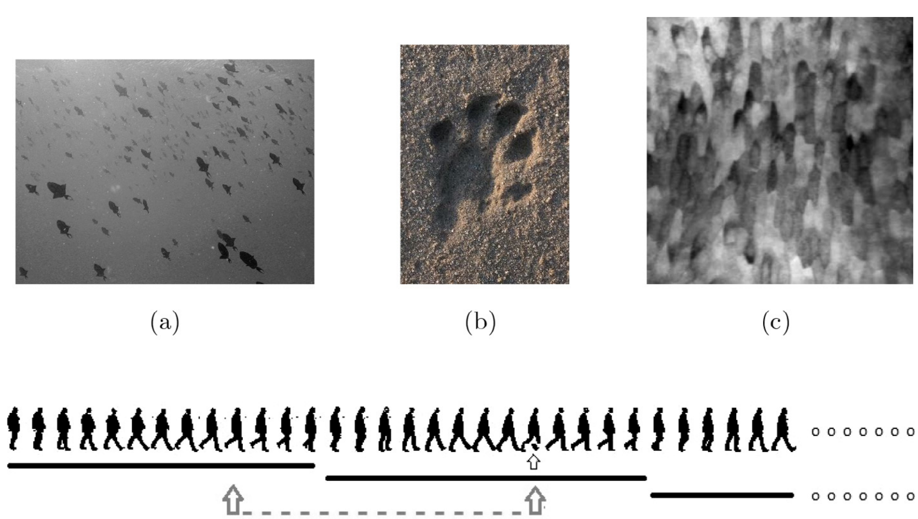
All of the above 7 new measures are translation, rotation and scaling invariant suitable for multi-component shape analysis.

All derivations are made in a continuous space. This makes the methods applicable in all discretization schemes directly (i.e. without any modification needed, depending of the discretization scheme applied). Our experiment are done on digital images, whose model is a finite squared integer grid, where each grid node represents a center of certain pixel in digital image. In such a situation, discretization processes and inherent numerical errors caused by them are at the end of the computation process. This is preferable since the unavoidable numerical errors, at the beginning of the certain process, could lead to a higher cumulative error, at the end of the computation process.

The paper is organized as follows. Necessary definitions and notations are in the next section. Section 3 introduces the new moment based measures/invariants for multi-component shapes. Experiments evaluating the properties of the new measures are in Section 4 where we illustrate the properties of the new features and use them for the leaf classification. Some illustrative synthetic examples are given, as well. Concluding remarks are in Section 5.

2. Preliminaries

In this section we define the basic terms used in this paper and introduce necessary notations. In addition, we give an area based analogue of the anisotropy measure, defined for the multi-component shapes (Rosin and Žunić, 2011), but computed from



(a) (b) (c)

(d)

Fig. 1. (a) A fish shoal. (b) A species footprint. (c) An embryonic tissue with relatively clear cell boundaries. (d) Human gait – considered as a 13-component shape, whose components are the appearances of a walking person in a sequence of 13 consecutive frames.