2009 American Control Conference
Hyatt Regency Riverfront, St. Louis, MO, USA
June 10-12, 2009



# Estimation of the Dynamics of Two-Dimensional Clusters

Daniele Casagrande, M. Sassano and Alessandro Astolfi

*Abstract*— **A method to estimate the dynamics of clusters is presented. An existing static clustering method, which makes use of a Hamiltonian function, is exploited in the dynamic case, when the points to cluster do not remain in fixed positions but move with unknown dynamics. The time is discretised and the static algorithm is applied to each time-instant. While identifying each cluster, the algorithm also performs the computation of the moments of the cluster. By estimating the dynamics of the moments an approximate estimation of the dynamics of the cluster is obtained. An illustrative example shows the performance of the method.**

## I. Introduction

In many scientific and engineering studies the practice of clustering data is very common when there is the need of dividing a set of measurements into subgroups with the aim of possibly deduct common features for measurements belonging to the same group. For instance, in the Hertzsprung-Russell diagram used in astronomy, which classifies the star according to their color and magnitude, the cluster of the *white dwarfs* corresponds to a possible stage of the evolution of a star. Many other applications such as pattern recognition, data analysis, image processing and market research may require (or take advantage from) a clustering of data (for a survey on the applications see, for instance, [1] and [2]).

Usually, the data to be clustered are represented by points in a given space and in most of cases the clustering algorithms perform the classification on the basis of the geometric displacements of the points. In this respect, the clustering algorithm turns out to be easier if the number of cluster is known as an input data. If this number is not known, the algorithm has to perform the so-called *cluster validation* problem [3], [4].

In a recent publication [5] a novel clustering algorithm, exploiting the idea of a *level function* associated to the measurements, was described and shown to perform successfully in some static scenarios without the *a priori* knowledge of the total number of clusters, thus solving also the cluster validation problem. In the particular application considered therein, the measurements points were corresponding to the position of objects in the two-dimensional space and in a particular time-instant. In this paper the result is extended to the dynamic case, namely when the points to be clustered move on the two-dimensional space. The aim is not only to divide the data points into subgroups with similar features, but also to estimate the dynamics of each group.

The first goal is achieved[1] by defining a *clustering function*, to be regarded as a level function, and identifying a cluster as the set of points internal to the same connected region belonging to a level set[2]. The boundary of the connected region is the trajectory of a Hamiltonian system, the Hamiltonian function of which is the clustering function. Once the connected region is identified, the notion of moments of a two-dimensional region (see [7] and the references cited therein) is applied to it in order to estimate important features of the associated cluster such as its size and its centre of mass.

In the dynamic version of the algorithm the displacement of the data points is sampled at time-instants multiple of a sampling time and for each time-instant the static method is applied. This process generates a sequence of values of the quantities represented by the moments, *e.g.* the coordinates of the centre of mass, each of which can be treated as the output signal of an autoregressive (AR) system. Finally, to estimate the parameters of the AR system and to predict the future outputs, standard methods from the system identification theory can be used.

The paper is organized as follows. In Section II the main notation is introduced while the static solution to the clustering problem is presented in details in Section III. Section IV describes how the static method can be exploited to estimate the dynamics of the clusters. The results of a simulation are reported and discussed in Section V. Conclusions and future developments are finally drawn in Section VI.

## II. Notation

The notation used throughout the paper is explained when introduced. Nevertheless, to provide the reader with a quick reference, the most used symbols are here summarised.

- $\mathcal{P}$ denotes the set of data points to cluster,
- $N$ denotes the cardinality of $\mathcal{P}$,
- $\mathcal{D}$ denotes a cluster ($\mathcal{D}_i$ has the same meaning),
- $\mathcal{R}$ denotes a connected region of the level set ($\mathcal{R}_i$, $\mathcal{G}$, $\mathcal{F}$ and $\mathcal{G}_i$ have the same meaning),
- $\mathcal{S}$ denotes the boundary of a connected region of the level set.

This work has been partially supported by the SEAS DTC project "Dynamic Set Recognition".

D. Casagrande is with the Dipartimento di Ingegneria Elettrica, Gestionale e Meccanica, Università degli Studi di Udine, Via delle Scienze 208, 33100 Udine, Italy (Email: daniele.casagrande@uniud.it).

M. Sassano is with the Department of Electrical and Electronic Engineering, Imperial College London, Exhibition Road, London SW7 2AZ, UK (Email: mario.sassano08@imperial.ac.uk).

A. Astolfi is with the Department of Electrical and Electronic Engineering, Imperial College London, Exhibition Road, London SW7 2AZ, UK (Email: a.astolfi@ic.ac.uk) and with the Dipartimento di Informatica, Sistemi e Produzione, Università degli Studi di Roma "Tor Vergata", via del Politecnico, 1 - 00133 Roma, Italy.

[1] For a description of the clustering based on the level set method, see [5].

[2] A similar approach is followed in [6].

978-1-4244-4524-0/09/$25.00 ©2009 AACC

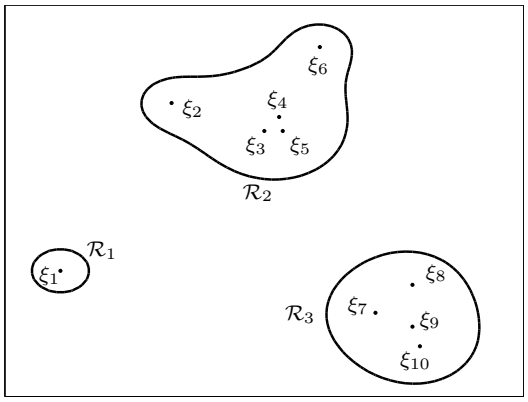


Fig. 1. A possible configuration for $\mathcal{P} = \{\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_{10}\}$. The three regions $\mathcal{R}_1$, $\mathcal{R}_2$ and $\mathcal{R}_3$ (the contour of which are $\mathcal{S}_1$, $\mathcal{S}_2$ and $\mathcal{S}_3$ respectively) identify three clusters: $\mathcal{D}_1 = \{\boldsymbol{\xi}_1\}$, $\mathcal{D}_2 = \{\boldsymbol{\xi}_2, \boldsymbol{\xi}_3, \boldsymbol{\xi}_4, \boldsymbol{\xi}_5, \boldsymbol{\xi}_6\}$ and $\mathcal{D}_3 = \{\boldsymbol{\xi}_7, \boldsymbol{\xi}_8, \boldsymbol{\xi}_9, \boldsymbol{\xi}_{10}\}$.

## III. Overview on the static algorithm

Suppose that the set of points to be clustered is $\mathcal{P} \triangleq \{\boldsymbol{\xi}_1, \ldots \boldsymbol{\xi}_N\} \subset \mathbb{R}^2$, where $N \in \mathbb{N}$. The method presented in [5] solves at the same time both the problem of finding the total number $K$ of clusters and that of partitioning $\mathcal{P}$ into $\mathcal{D}_1, \ldots, \mathcal{D}_K$, such that $\mathcal{P} = \bigcup_{j=1}^{K} \mathcal{D}_j$ and[3] $\mathcal{D}_{j_1} \cap \mathcal{D}_{j_2} = \emptyset$ if $j_1 \neq j_2$. The method can briefly be described as follows.

• A *clustering function* $H : \mathbb{R}^2 \to \mathbb{R}$ is chosen on the basis of which a Hamiltonian system is defined. In particular, one can assign to each point $\xi_i$, $i = 1, \ldots, N$ a function $H_i : \mathbb{R}^2 \to \mathbb{R}$ and define the clustering function as

$$H(\boldsymbol{\xi}) = \sum_{i=1}^{N} H_i(\boldsymbol{\xi}) \,. \tag{1}$$

• A reference level $H_r$ is selected and each connected region $\mathcal{R}_j$ of the level set of $H$ with respect to $H_r$ implicitly identifies the $j$-th cluster. More precisely (see Figure 1) $\mathcal{D}_j \subset \mathcal{R}_j$, $j = 1, \ldots, K$, where the $\mathcal{R}_j$'s are connected and such that

$$\bigcup_{i=1}^{K} \mathcal{R}_i = \{\boldsymbol{\xi} \in \mathbb{R}^2 \,:\, H(\boldsymbol{\xi}) \geqslant H_r\} \,. \tag{2}$$

Note that the value of $K$ depends on the value of $H_r$; hence the *cluster validation* problem mentioned in the Introduction is transformed into the determination of the optimal value of $H_r$.

• To explicitly find the level line $\mathcal{S}_j$ associated to $\mathcal{R}_j$, which solves the clustering problem, the clustering function $H$ is regarded as a Hamiltonian function and the corresponding Hamiltonian system is considered. In fact, denoting with $x$ and $y$ the coordinates of $\boldsymbol{\xi}$, i.e. $(x, y) = \boldsymbol{\xi}$, the level line $\mathcal{S}_j$ corresponds to the trajectory of the system

$$\begin{cases} \dot{x} &= \dfrac{\partial H}{\partial y} f(x, y) \,, & x(0) = \overline{x} \,, \\ \dot{y} &= -\dfrac{\partial H}{\partial x} f(x, y) \,, & y(0) = \overline{y} \,, \end{cases} \tag{3}$$

where $f(x, y)$ is any continuous function and the initial condition $(\bar{x}, \bar{y})$ is such that $(\overline{x}, \overline{y}) \in \mathcal{S}_j$ (and, obviously, $H(\bar{x}, \bar{y}) = H_r$). Easy computations yield

$$\dot{H} = \frac{\partial H}{\partial x}\dot{x} + \frac{\partial H}{\partial y}\dot{y} = 0 \tag{4}$$

which means that $H(x(t), y(t)) \equiv H_r$, for $t \geqslant 0$. Therefore the trajectory of system (3) lies on $\mathcal{S}_j$.

*Remark 1:* The knowledge of an initial condition $\overline{\boldsymbol{\xi}}$ such that $H(\overline{\boldsymbol{\xi}}) = H_r$ is not necessary. As shown in [5], to steer in finite time any initial condition to a level line it is sufficient to modify system (3) as follows:

$$\begin{cases} \dot{x} &= f(x, y) \left( \dfrac{\partial H}{\partial y} - \dfrac{\partial H}{\partial x} (H - H_r)^{1/3} \right) , \\ \dot{y} &= -f(x, y) \left( \dfrac{\partial H}{\partial x} + \dfrac{\partial H}{\partial y} (H - H_r)^{1/3} \right) . \end{cases} \tag{5}$$

Note, in fact, that the additional term vanishes when $H = H_r$. ♢

*Remark 2:* When implementing the method in the practice, solutions to system (3), or (5), are found using standard algorithms for numerical integration. ♢

## IV. The discrete-time dynamic algorithm

The algorithm described in the previous section can be exploited to predict the motion of the clusters in the following way. Suppose that the points to be clustered do not remain in fixed positions but move. Obviously, due to the motion of the points, the set of regions changes in time. It may happen, for instance, that the same point is clustered at time $\tau_{k_1}$ in $\mathcal{R}_{j_1}$ while at time $\tau_{k_2} > \tau_{k_1}$, due to its motion, it is clustered in $\mathcal{R}_{j_2} \neq \mathcal{R}_{j_1}$. In addition, the detection of all the points in all time-instants is not guaranteed; it may happen, for instance, that a point *disappears*, namely that it is detected up to the instant $\tau_{k_1}$ but not from $\tau_{k_1+1}$ onwards; analogously a point may *appear* at time $\tau_{k_2}$, which was not detected for $\tau_k < \tau_{k_2}$. Therefore, in general, the set $\mathcal{P}$ of the points to cluster and its cardinality $N$ are functions of time, *i.e.* $\mathcal{P} = \mathcal{P}(\tau_k)$ and $N = N(\tau_k)$. Moreover, for $\tau_{k_1} \neq \tau_{k_2}$,

$$\bigcup_{i=1}^{N(\tau_{k_1})} \mathcal{R}_i(\tau_{k_1}) \neq \bigcup_{i=1}^{N(\tau_{k_2})} \mathcal{R}_i(\tau_{k_2}) \,.$$

Consider now two consecutive time-instants, say $\tau_{k_1}$ and $\tau_{k_2} = \tau_{k_1} + T$, and a region $\mathcal{G} \in \{\mathcal{R}_1(\tau_{k_2}), \ldots, \mathcal{R}_{N_2}(\tau_{k_2})\}$. The first problem addressed in the paper is the following.

Q1. Is it possible to associate $\mathcal{G}$ to one of the regions $\mathcal{F} \in \{\mathcal{R}_1(\tau_{k_1}), \ldots, \mathcal{R}_{N_1}(\tau_{k_1})\}$ in such a way that $\mathcal{G}$ is likely to be the *evolution* of $\mathcal{F}$?

[3] In some applications, cluster evolution and cluster dynamics are studied assuming that two (or more) clusters which may overlap are kept separated. On the contrary, here we are interested in merging together overlapping clusters; this approach fits best with the final example of merging clouds of water vapor.

We show that this is possible, provided that the velocities of the points admits an upper bound, that the sampling period $T$ is chosen accordingly and that the notion of evolution is properly defined. An interesting consequence of this result is that from the application of the static method to a sequence one can build a sequence of clusters $\mathcal{G}_1$, $\mathcal{G}_2$, ... such that

$$\begin{aligned}
\mathcal{G}_0 &\in \{\mathcal{R}_1(\tau_{k_1}), \ldots, \mathcal{R}_{N(\tau_{k_1})}(\tau_{k_1})\}, \\
\mathcal{G}_1 &\in \{\mathcal{R}_1(\tau_{k_1}+T), \ldots, \mathcal{R}_{N(\tau_{k_1}+T)}(\tau_{k_1}+T)\}, \\
&\vdots \quad\quad \vdots \\
\mathcal{G}_K &\in \{\mathcal{R}_1(\tau_{k_1}+KT), \ldots, \mathcal{R}_{N(\tau_{k_1}+KT)}(\tau_{k_1}+KT)\},
\end{aligned}$$

and $\mathcal{G}_i$ is the *evolution* of $\mathcal{G}_{i-1}$, for $i = 2, \ldots, K$. This leads to the following natural question.

Q2. Given a sequence of *evolving* clusters $\mathcal{G}_1, \mathcal{G}_2, \ldots \mathcal{G}_K$ is it possible to predict how the sequence continues, namely how $\mathcal{G}_{K+1}$ is expected to be?

We show that the least square error method [8] can be used to estimate the dynamical model of the evolution of the clusters, thus answering positively to question Q1 and Q2. However, to provide a solution to the problems posed by Q1 and Q2, we first need to give a quantitative description of the clusters, beyond their implicit definition (3). To this purpose we use the notion of moments [7].

*Definition 1:* For a function $F(x,y) : \mathbb{R}^2 \to \mathbb{R}$ the *moment* of order $p+q$, with $p$ and $q$ integers and such that $p+q \geqslant 0$, is defined as

$$m_{pq}(F) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^p y^q F(x,y)\mathrm{d}x\mathrm{d}y\,. \qquad \Diamond \quad (6)$$

Consider, now, the region $\overline{\mathcal{R}}$ containing the cluster $\overline{\mathcal{D}}$. The *index* function of $\overline{\mathcal{R}}$ is:

$$F_{\overline{\mathcal{R}}}(x,y) \triangleq \begin{cases} 1\,, & \text{if } (x,y) \in \overline{\mathcal{R}}\,, \\ 0\,, & \text{otherwise}\ . \end{cases} \quad (7)$$

From (6) and (7) it is easy to relate the moments of order 0 and 1 to the position and the size of $\overline{\mathcal{R}}$. In fact

- the moment of order 0 of $F_{\overline{\mathcal{R}}}$ is the area of $\mathcal{R}$:

$$m_{00}(F_{\overline{\mathcal{R}}}) = \iint_{\overline{\mathcal{R}}} \mathrm{d}\boldsymbol{\xi}\,,$$

- the position of the center of mass[4] of $\overline{\mathcal{R}}$, can be obtained from the moments of order 0 and 1. Denoting by $x_{\overline{\mathcal{R}}}$ and $y_{\overline{\mathcal{R}}}$ its coordinates, we have

$$x_{\overline{\mathcal{R}}} = \frac{m_{10}(F_{\overline{\mathcal{R}}})}{m_{00}(F_{\overline{\mathcal{R}}})} \qquad \text{and} \qquad y_{\overline{\mathcal{R}}} = \frac{m_{01}(F_{\overline{\mathcal{R}}})}{m_{00}(F_{\overline{\mathcal{R}}})}\,.$$

Thus the size (area) and the position (displacement of the centre of mass) of the cluster can be quantified by using the moments of order 0 and 1. A further step is the quantitative description of the *shape* of the cluster, what can be done, in the first approximation, with the help of the notion of *image ellipse* and *principal axes* [7].

[4]This, in general, does not coincide with the center of mass of $\overline{\mathcal{D}}$.

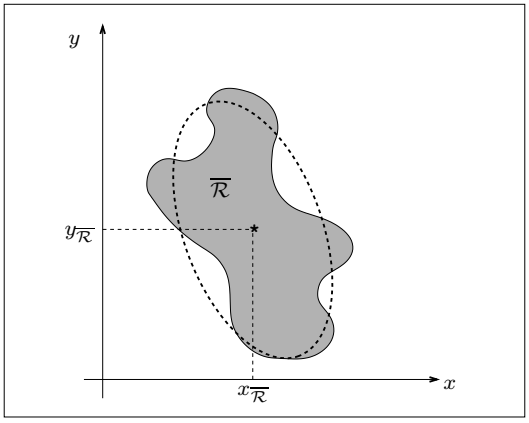


Fig. 2. A connected region $\overline{\mathcal{R}} \subset \mathbb{R}^2$ (shaded region) can be approximated by an ellipse (dashed line) having the same semiaxes of the image ellipse of $\overline{\mathcal{R}}$ and the same centre of mass of $\overline{\mathcal{R}}$ (asterisk) and oriented according to the principal axes of $\overline{\mathcal{R}}$.

*Definition 2:* The *image ellipse* of a region $\overline{\mathcal{R}}$ is the elliptical disk having the same moments of order 0 and 2 of $\overline{\mathcal{R}}$ and such that its major and minor axes lie along the $x$ and, respectively, $y$ axes of a Cartesian system of coordinates. $\Diamond$

*Definition 3:* The *principal axes* of a region $\overline{\mathcal{R}}$ are the pair of straight lines along which there are the minimum and maximum inertia moments of $\overline{\mathcal{R}}$. $\Diamond$

In the following $\alpha(\overline{\mathcal{R}})$ and $\beta(\overline{\mathcal{R}})$ denote the major and, respectively, the minor semi-axis of the image ellipse; $a_m(\overline{\mathcal{R}})$ and $a_M(\overline{\mathcal{R}})$ denote the principal axes of $\overline{\mathcal{R}}$, $a_m(\overline{\mathcal{R}})$ being associated to the minimum inertia moment of $\overline{\mathcal{R}}$.

The meaning of Definitions 2 and 3 is that the best approximation of the shape of a connected region $\overline{\mathcal{R}}$, up to the moments of order 2, is given by an ellipse (see Figure 2) which has the same centre of mass of $\overline{\mathcal{R}}$ and such that its major semi-axis is equal to $\alpha$ and lies along $a_M(\overline{\mathcal{R}})$.

The geometric quantities introduced with Definitions 2 and 3 can be obtained from the moments of order 2 of $\overline{\mathcal{R}}$ computed with respect to its centre of mass, which are called the *central moments*.

*Definition 4:* For a function $F(x,y) : \mathbb{R}^2 \to \mathbb{R}$, the *central moment* of order $p+q$, with $p+q > 2$, is defined as:

$$\mu_{pq}(F) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_T^p y_T^q F(x_T, y_T)\mathrm{d}x\mathrm{d}y\,,$$

where $x_T = x - x_{\overline{\mathcal{R}}}$ and $y_T = y - y_{\overline{\mathcal{R}}}$. $\Diamond$

The moments of order 2 of $F_{\overline{\mathcal{R}}}$ are related to the dimensions of the image ellipse by the following equalities:

$$\begin{aligned}
\alpha(\overline{\mathcal{R}}) &= \left(\frac{2\left(\mu_{20}(F_{\overline{\mathcal{R}}}) + \mu_{02}(F_{\overline{\mathcal{R}}}) + \Delta(F_{\overline{\mathcal{R}}})\right)}{\mu_{00}(F_{\overline{\mathcal{R}}})}\right)^{1/2}, \\
\beta(\overline{\mathcal{R}}) &= \left(\frac{2\left(\mu_{20}(F_{\overline{\mathcal{R}}}) + \mu_{02}(F_{\overline{\mathcal{R}}}) - \Delta(F_{\overline{\mathcal{R}}})\right)}{\mu_{00}(F_{\overline{\mathcal{R}}})}\right)^{1/2},
\end{aligned}$$

where $\Delta(F_{\overline{\mathcal{R}}}) = \sqrt{(\mu_{20}(F_{\overline{\mathcal{R}}}) - \mu_{02}(F_{\overline{\mathcal{R}}}))^2 + 4(\mu_{11}(F_{\overline{\mathcal{R}}}))^2}$. Moreover (see [7]) the orientation of $a_m$ and $a_M$ can also be deducted.

### A. Computation of the moments

One of the main advantages of defining the clustering on the basis of the trajectories of system (3) is that the moments of any order of $\mathcal{F}_{\overline{\mathcal{R}}}$ can be computed by integrating Equation (3). In fact, by using Green's theorem a double integral over a region correspond to a line integral over its boundary.

*Theorem 1 (Green's):* Let $\mathcal{S}$ be a positively oriented, piecewise smooth, simple closed curve in $\mathbb{R}^2$ and let $\mathcal{R}$ be the region bounded by $\mathcal{S}$. If $L$ and $M$ are functions of $x$ and $y$ defined on an open region containing $\mathcal{R}$ and with continuous partial derivatives there, then

$$\iint_{\mathcal{R}} \left( \frac{\partial M}{\partial x} - \frac{\partial L}{\partial y} \right) \mathrm{d}x\mathrm{d}y = \int_{\mathcal{S}} \left( L\mathrm{d}x + M\mathrm{d}y \right) . \qquad \Diamond$$

By applying Theorem 1 to Equation (6) with, for instance, $\frac{\partial M}{\partial x} = x^p y^q$ and $L = 0$ one obtains

$$\iint_{\overline{\mathcal{R}}} x^p y^q \mathrm{d}x\mathrm{d}y = \int_{\overline{\mathcal{S}}} \frac{1}{p+1} x^{p+1} y^q \mathrm{d}y \,, \tag{8}$$

where the term on the right hand side is an integral along a (closed) trajectory $\overline{\mathcal{S}}$ corresponding to one of the level lines of the clustering function $H$. From Equation (8) the dynamics of the value of the moments on the time-scale of the Hamiltonian system (3) can be obtained. In particular, the differential equations

$$\dot{m}_{00} = x\dot{y} \,, \tag{9}$$

$$\dot{m}_{10} = \frac{1}{2} x^2 \dot{y} \,, \tag{10}$$

$$\dot{m}_{01} = xy\dot{y} \,, \tag{11}$$

$$\dot{m}_{20} = \frac{1}{3} x^3 \dot{y} \,, \tag{12}$$

$$\dot{m}_{11} = \frac{1}{2} x^2 y \dot{y} \,, \tag{13}$$

$$\dot{m}_{02} = x y^2 \dot{y} \,, \tag{14}$$

with $m_{ij}(0) = 0$, are such that $m_{ij}(\hat{t}) = m_{ij}(F_{\overline{\mathcal{R}}})$, where $\hat{t}$ is the period of the orbit of system (3) along $\overline{\mathcal{S}}$.

*Remark 3:* The differential equations (9)-(14) do not describe the dynamics of the moments on the time-scale of the motion of the points $\boldsymbol{\xi}$'s but only the variation of their value on the time-scale of the Hamiltonian system (3), the trajectory of which is the level line $\overline{\mathcal{S}}$.

*Remark 4:* The right-hand side term in Equations (9) is a basis for the non-integrable differential forms of order one in the variables $x$ and $y$. Analogously, the right-hand side terms in Equations (10)-(11) and in Equations (12)-(14) are a basis for the space of non-integrable differential forms of order two and three, respectively. The same set of equations, which define all the possible non-holonomic systems of order one, two and three, respectively, is considered (for different scopes) in other works available in the literature (see, for example, [9]).

### B. The sequence of clusters

From their definition, it is clear that moments provide a quantitative description of the dimension, the position, the orientation and the shape of the regions containing the clusters[5]. More precisely, by computing the moments we construct a map $\mathcal{M}$ such that each region $\mathcal{R}$ is mapped into a point $\mathcal{M}(\mathcal{R}) = \mathbf{m}_{\mathcal{R}} \triangleq (m_{00}, m_{10}, m_{01}, \mu_{20}, \mu_{11}, \mu_{02}) \in \mathbb{R}^6$. This can be used to construct a sequence of clusters, each of which is the evolution of the previous one. In fact, if the sampling period $T$ is chosen small with respect to the velocities of the measurement points, reasonably a cluster $\mathcal{D}$ and the region $\mathcal{R}$ surrounding it experience small changes between a sampling time-instant and the following one. As a consequence, the images of the two regions, through $\mathcal{M}$, are not too far one from the other. This behaviour can be described more precisely in the following way. Consider a region $\mathcal{F} \in \{\mathcal{R}_1(\tau_{k_1}), \ldots, \mathcal{R}_{N_1}(\tau_{k_1})\}$ and an ellipsoid in $\mathbb{R}^6$ having the centre of mass in $\mathbf{m}_{\mathcal{F}}$ and the semi-axis of which are the coordinates of a vector $\mathbf{d} = (d_1, \ldots, d_6) \in \mathbb{R}^6$, namely the set

$$E_{\mathbf{d}}(\mathcal{F}) = \left\{ (z_1, \ldots, z_6) \in \mathbb{R}^6 \,:\, \sum_{i=1}^{6} \frac{(z_i - m_{\mathcal{F},i})^2}{d_i^2} \leqslant 1 \right\} ,$$

where $m_{\mathcal{F},i}$ denotes the $i$-th component of $\mathbf{m}_{\mathcal{F}}$.

*Definition 5:* A cluster $\mathcal{G} \in \{\mathcal{R}_1(\tau_{k_2}), \ldots, \mathcal{R}_{N_2}(\tau_{k_2})\}$ is the one-step $E_d$-*evolution* of $\mathcal{F}$, what is denoted by $\mathcal{G} \sim_{E_{\mathbf{d}}} \mathcal{F}$, if $\mathcal{G} \in E_{\mathbf{d}}(\mathcal{F})$. $\Diamond$

With the help of Definition 5 the generic $K$-steps evolution from $\mathcal{G}_0 \in \{\mathcal{R}_1(\tau_{k_0}), \ldots, \mathcal{R}_{N_0}(\tau_{k_0})\}$ can be defined as the sequence $\{\mathcal{G}_k\}_{k \in \{1,\ldots,K\}}$ such that, for $i = 1, \ldots, K$, $\mathcal{G}_i \in \{\mathcal{R}_1(\tau_{k_0} + iT), \ldots, \mathcal{R}_{N_i}(\tau_{k_0} + iT)\}$ and $\mathcal{G}_i \sim_{E_{\mathbf{d}}} \mathcal{G}_{i-1}$.

*Remark 5:* Definition 5 allows the possibility that two or more regions are the $E_d$-evolution of $\mathcal{F}$. Such an occurrence, which prevents from having a uniquely determined $K$-step evolution, can be limited if the $d_i$'s are small; in this case, in fact it is likely that only one region $\mathcal{G} \in \{\mathcal{R}_1(\tau_{k_2}), \ldots, \mathcal{R}_{N_2}(\tau_{k_2})\}$ belongs to $E_{\mathbf{d}}(\mathcal{F})$. However, even when the $d_i$'s are very small, two or more different regions might belong to $E_{\mathbf{d}}$. In this case, if there are no other criteria, the regions are equivalent and one of them can be picked randomly to continue the sequence. Another criterion necessary to construct a meaningful sequence is that of requiring that the value of the $d_i$'s is not too small, otherwise no $\mathcal{G} \in \{\mathcal{R}_1(\tau_{k_2}), \ldots, \mathcal{R}_{N_2}(\tau_{k_2})\}$ belongs to $E_{\mathbf{d}}(\mathcal{F})$ and the sequence is made just by $\mathcal{F}$.

*Remark 6:* Definition 5 is given in the case of six moments but in many practical cases, as the one considered in Section V, it is sufficient to limit the analysis to the projection of $E_{\mathbf{d}}$ onto $\mathbb{R}^2$ given by $d_2$ and $d_3$. In this case $\mathcal{G}_i$ is the evolution of $\mathcal{G}_{i-1}$ if the centre of mass of $\mathcal{G}_i$ belongs to an ellipse (in $\mathbb{R}^2$) centred in the centre of mass of $\mathcal{G}_{i-1}$.

### C. Estimation of the dynamics

Each sequence $\{\mathcal{G}_k\}_{k \in \mathbb{N}}$ generates a sequence of values for each moment up to order 2, namely

$$\nu(0), \nu(1), \ldots, \nu(k), \nu(k+1), \ldots \,, \tag{15}$$

[5]Other kind of metrics to describe clusters can be found in [10].

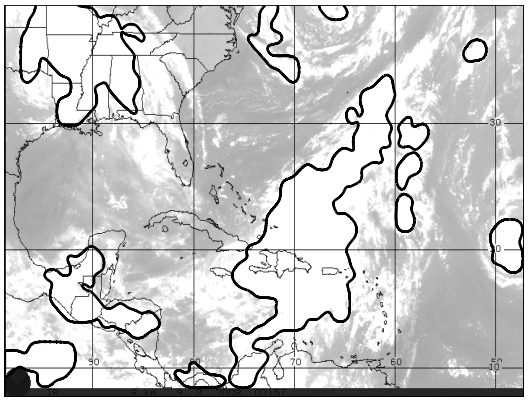

Fig. 3. An infrared image taken from the satellite and representing the presence of water (white clouds). The closed lines represent the regions, namely the result of the clustering method on the sampled measurements. The cloud to be tracked is the first region on the right.

where $\nu$ can be $m_{00}$, $m_{10}$, $m_{01}$, $\mu_{20}$, $\mu_{11}$ or $\mu_{02}$. Regarding this sequence as the output of an autoregressive process,

$$\nu(t) + b_1\nu(t-1) + \ldots + b_n\nu(t-n) = 0\,,$$

standard techniques from system identification [8] can be used to estimate both the order of the difference equation $n$ and the vector of parameters $\mathbf{b} = (b_1, \ldots, b_n)^\top$ which gives the best fit of the sequence $\{\nu(k)\}_{k\in\mathbb{N}}$.

## V. Simulations

To test the performance of the method presented in the previous sections, a real scenario has been considered: the prediction of the motion of a hurricane based on the data provided by images taken from the satellite. The sampling time $T$ is half an hour, in comparison with which the computational time is negligible. Each image is a *greyscale* image and has $720 \times 480$ pixels. To ease the computational load a spatial sampling of each image has been performed reducing the size to $72 \times 48$ pixels. Finally, the *white* level associated to each pixel is compared with a threshold and the measurements set $\mathcal{P}$ is made of all the pixels having a white level larger than the threshold. Figure 3 shows the first image[6] of a stream where the level lines corresponding to each cluster are detected. The cloud the motion of which we want to track is the first region on the right. To provide the reader with an idea of how a sequence of regions corresponding to the clusters can be obtained from a sequence of images, the result of the application of the static method to nine images equally spaced in time is reported in Figure 4.

The clustering technique described in Section IV has been applied to a a sequence of 288 images (six days), using the first 48 (one day) to estimate the model and the last 240 to validate it. In particular, we report the results concerning the estimation of the dynamics of $x_{\overline{\mathcal{R}}}$.

This values have been applied to predict the motion of the cloud in the subsequent five days with four different values

[6]This image and the ones appearing in the following have been taken from the website `http://www.goes-arch.noaa.gov`.

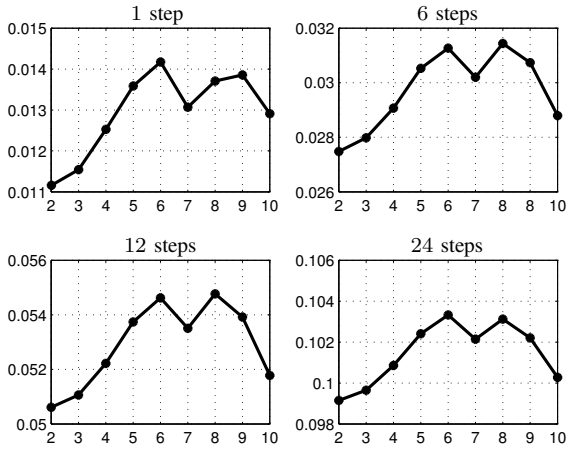


Fig. 5. Four plots corresponding to four different prediction horizons. The average absolute relative error is plotted for all difference equations (from order 2 to order 10). A difference equation of order 2 is the one minimising the error.

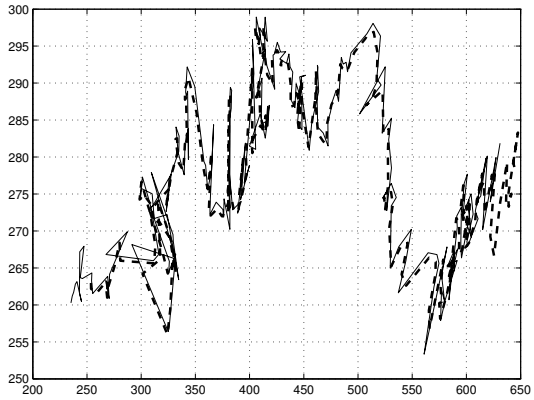

Fig. 6. The trajectory of the centre of mass estimated over a 12 step horizon (dashed bold line) together with the true trajectory (solid line).

of the prediction horizon, namely half an hour (1 step), three hours (6 steps), six hours (12 steps) and twelve hours (24 steps). Results are reported in Figure 5. It can be noted that the best approximation is a difference equation of order two.

By using the model of order two both for the dynamics of $x_{\overline{\mathcal{R}}}$ and for $y_{\overline{\mathcal{R}}}$ the trajectory of the centre of mass has been estimated. Results corresponding to a prediction horizon of length 12 are reported in Figure 6 where the true trajectory is also plotted. On Figure 7 the same trajectory has been plotted over the image corresponding to the last position of the trajectory.

## VI. Conclusions

The idea presented in this paper is a promising first step in the study of the dynamics of two-dimensional regions and clusters of points. The application of the developed method to the problem of predicting the motion of a hurricane has provided interesting results. Different applications and a generalization of the method to the case of $n$-dimensional clusters will be the subjects for further investigation. Other

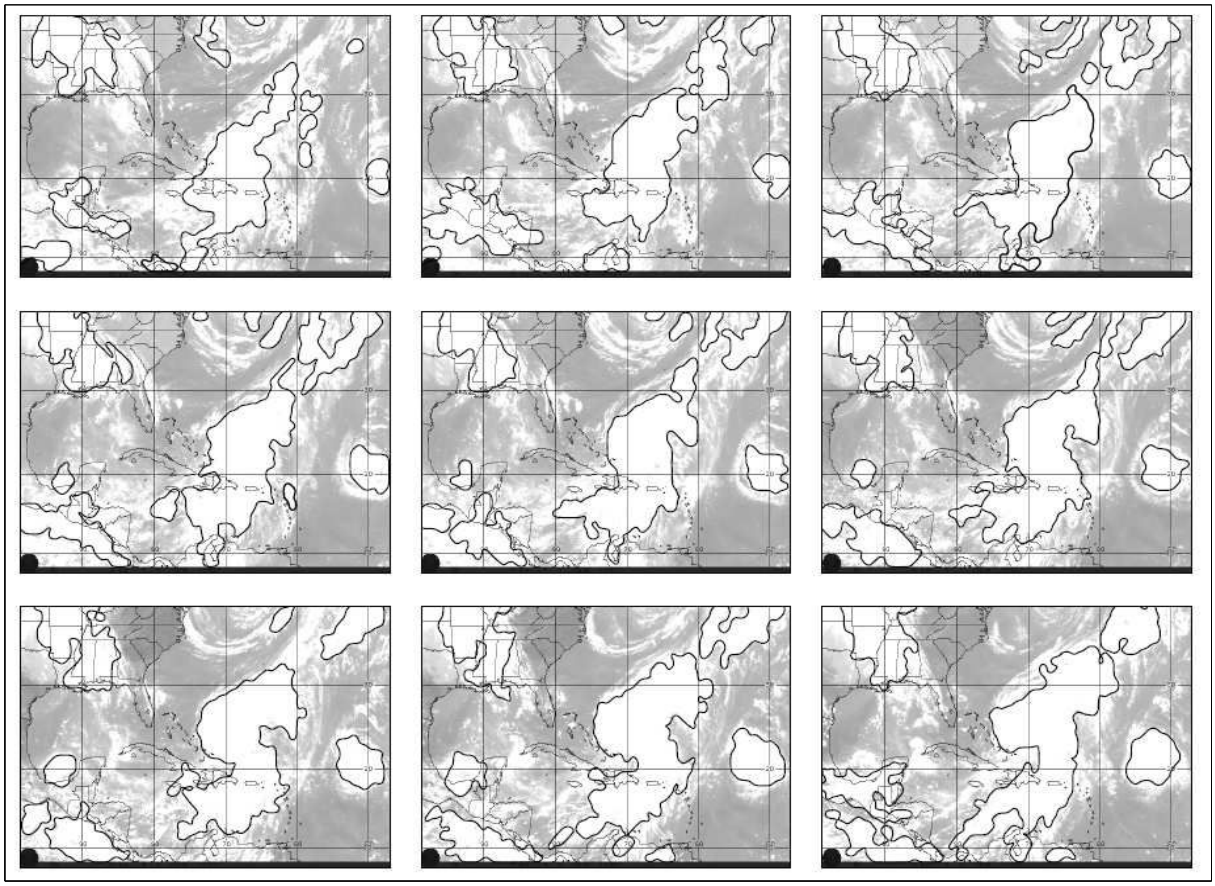

Fig. 4. A set of images and the corresponding clusters. The chronological order is from the left to the right and from the top to the bottom. The length of the time-interval between one image and the following one is two hours.

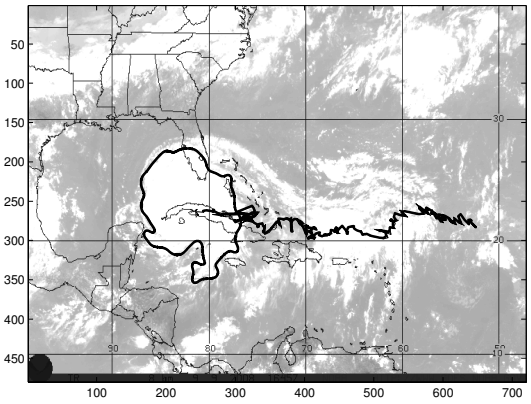


Fig. 7. The same trajectory of Figure 6 plotted over the image corresponding to the last position of the centre of mass.

interesting improvements to pursue may concern different models for estimating the discrete dynamics (instead of the autoregressive one used herein) or the extension of the method to the situations in which the regions for which the moments are computed are not simply connected.

## ACKNOWLEDGEMENTS

The authors would like to thank Neil Cade from SELEX Sensors and Airborne Systems Limited for many useful discussions that contributed to this work.

## REFERENCES

[1] R.O. Duda and P.E. Hart, *Pattern Classification and Scene Analysis*. John Wiley and sons, 1973.

[2] S. Theodoridis and K. Koutroumbas, *Pattern Recognition*. Academic press, 1999.

[3] N.R. Pal and J.C. Bezdek, "On cluster validity for the fuzzy $c$-means model," *IEEE Transactions on Fuzzy Systems*, vol. 3, no. 3, pp. 370–379, 1995.

[4] X.L. Xie and G. Beni, "A validity measure for fuzzy clustering," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, no. 8, pp. 841–847, 1991.

[5] D. Casagrande and A. Astolfi, "A Hamiltonian-based algorithm for measurements clustering," *Proc. of the 47th IEEE Control and Decision Conference*, 2008.

[6] G. V. Nosovskiy, D. Liu, and O. Sourina, "Automatic clustering and boundary detection algorithm based on adaptive influence function," *Pattern Recognition*, vol. 41, no. 9, pp. 2757–2776, 2008.

[7] R. Prokop and A. Reeves, "A survey of moment-based techniques for unocculted object representation and recognition," *CVGIP: Graphical Models and Image Processing*, vol. 54, no. 5, pp. 438–460, 1992.

[8] L. Ljung, *System Identification. Theory for the user*. Prentice Hall, 1999.

[9] R.W. Brockett and L. Dai, "Non-holonomic kinematics and the role of elliptic functions in constructive controllability," in *Nonholonomic Motion Planning*, Z. Li and J. F. Canny, Eds. Boston, MA: Kluwer Academic Publishers, 1993, pp. 1–22.

[10] R.O. Duda, P.E. Hart, and D.G. Stork, *Pattern Classification*. John Wiley and sons, 2001.