Facultad de Ciencias,
Ciudad Universitaria,
UNAM, México, D.F.

August 28, 2016

IEEE Transactions on Cognitive and Developmental Systems
Special Issue on Sensorimotor Contingencies for Cognitive Robotics

Dear Dr. Ricardo Téllez,

Please find attached our submission:

"A Multi-Modal Model for Prediction and Classification of Object Deformation during Robotic Manipulation"
Veronica E. Arriola-Rios (UNAM, Mexico), Jeremy L. Wyatt (University of Birmingham, UK)

Reviewer(s)' Comments to Author:

Reviewer: 1

Comments to the Author

This article aims at tracking three-dimensional deformable objects under robot motion and predicting the force causing the deformation.

It is important to make a precision here: the article aims at predicting the force and deformation on the visual plane of the three-dimensional object. The tracking of the shape is used only during training as part of the evaluation function.

For these purposes, the authors propose two models, including an extension of the mass-spring model and a regression model, where the parameters of the mass-spring model are learned with an evolutionary searching algorithm on visual data and force data, and the parameters of the regression model are learned with least-square methods on stress-strain graph. These models are tested on a sponge and a plasticine, and evaluated based on the balanced F-score. The results are presented both quantitatively and qualitatively.

However, there are several structural problems, as listed below.

⟹1: in the abstract, the authors claim that five features are simultaneously enabled by the proposed single learning model framework, but the article actually presents a collection of different models and methods to solve different problems, thus it is inappropriate to claim them as a single unified model framework. Besides, it seems that only robot pushing is modeled, not general "robot actions".

2: section II is not included in the final paragraph of section I, and this section should be included in section I or section IV.

In addition, this section is really difficult to follow. For example, in line 27 of page 1, "the specific approach taken here separates the predictive model out into components in two ways". It will be easier to read if the authors clearly indicate the second way in line 44. And the third paragraph, on model training, should come before the second graph, which is on model testing.

3: section IV-C, "Integration Scheme", seems to be part of algorithm 1 or section IV-B, thus it should be introduced.

It is neither part of algorithm 1, nor of section IV-B. It is part of a numerical algorithm for solving differential equations.
TODO: Add general algorithm which makes use of all the pieces explained in section IV at the beginning of the section.

4, section IV-D is not explained: what is the relation between "collision detection" and the proposed model?

5, section IV-E, "Geometric Constraints", should be integrated into section IV-A, where the energy function is built on different geometric constraints.

TODO: Move section IV-E after IV-A but must emphasize that the geometric constraints in IV-A are enforced through energy preservation, while those in IV-E have immediate effects on the triangulations without directly affecting the equations.

6, section IV-F, "Many Step Prediction", explains how to recursively train and test the model. It should be included in section V, where the model training is given.

There is a missunderstanding here. Section IV-F is independent of any training. As a physics based model, it depends only on proper entries and parameters, which could be provided by a human expert after manual calibration or by a computer algorithm. The last paragraph of this section was modified accordingly and the algorithm introduced for comment 3 should help here.

7, section V-A explains how to search for the optimal parameter with evolutionary algorithm based on the ground truth; section V-B explains how to get the ground truth; section V-C explains how to test the learned parameter. It seems that the correct order is B-A-C.

Changed order as suggested.

Besides, there are several technical ambiguities in the manuscript, as listed below.

$\implies$1, in line 37 of page 1, "sensorimotor contingencies" needs to be defined or referenced.

$\implies$2, in line 35 of page 1, what does "sequencing the models by mode" mean?

$\implies$3, in line 58 of page 1, the authors claim "In each case the models can be run in two modes, prediction and classification". But the classification mode is not explained later, but a filtering mode is included instead. Besides, please clearly illustrate the mentioned

modes and processes in Figure 2.

What does he mean by "filtering mode instead"?

4, in section IV-A, the author needs to explain how they extend the model proposed in [16].

The general idea had been explained in the first paragraph of section IV-A, the details came later. Since it was not clear enough the wording was modified.

5, in section IV-A, the author needs to briefly introduce the mass-spring model.

Added an introduction at the beginning of section IV.

6, in equation (1), please explain what "particle i" is. Is it vertex i?

Since for comment 5 the introduction explained that a mass particle is located on vertex i, this has been covered already.

7, in line 27 of page 4, please explain what "mass particle" is.

The term was introduced on the second paragraph of section IV.

8, in line 33 of page 4, please explain why the equations are oscillatory.

Added explanation that it happens when/because we use Cs that model springs.

9, please explain how equation (3) is derived.

It is not within the scope of the article to explain that derivation, since it belongs to Teschner's model, however the cite to the article we got it from was repeated here.

10, in line 54 of page 4, please add reference to "Teschner normalised the value dividing by D0 to make the elasticity constants scale independent."

Added.

11, in line 13, please explain "Since Teschner and Morris didn't find it helpful to add damping to the preservation of areas, it is avoided here".

Changed phrase for the explanation that Teschner gave.

12, in line 50 of page 4, please explain "the angles between adjacent edges". As adjacent edges form one angle, do the authors mean adjacent angles? Similar ambiguity exists in line 2 of page 5.

Corrected.

13, in Algorithm 1, please explain how "an integration step" leads to new length of the edge.

Reworded sentence to add more detail.

14, in line 25 of page 4, please explain what "Euler of Verlet" means. Euler or Verlet?

It was "or". Corrected.

15, in equation (17), please explain if $x(t + h)$ is a position, how does $h * \frac{F(t)}{m}$ generate a position vector, since $h * \frac{F(t)}{m}$ seems to be the production of h and acceleration? And please explain "we made the velocities proportional to the forces". Which variable is the velocity? What does "m" mean? The mass?

Explained in more detail and changed notations to make it consistent with previous paragraphs.

16, in section IV-D, please explain what "collision detection" means, the collision between robot and objects, or the collision between objects?

Improved the description across the whole subsection.

17, is algorithm 2 designed by the authors?

No, that is why we mentioned in the text that it is a common algorithm. It can be found across the web in many sites. Should we mention something about that in the caption?

18, in section V-A, please explain why evolutionary algorithm is used. In algorithm 4, please explain what $\alpha, \beta, \gamma, \zeta$ represent.

Section V-B? Added a justification at the beginning of the section. The meaning of $\alpha, \beta, \gamma, \zeta$ was covered on lines 7,9,10,11 but a shorter description was added when they are firstly introduced.

19, in line 60 of page 7, please explain how the object is segmented from the environment, which "simple colour segmentation is used", and how to define "edges that not likely to belong to the object of interest".

Changed paragraph.

20, in line 23 of page 7, please explain what "linear snake" indicates. Is it from active contour model? Also, please add reference.

Reference added. Relationship with active contour explained.

21, in section V-C, please note the measurement has a name, "F-score", in machine learning community. Also, please define $\mu(f)$ based on "the mean over all the frames so far".

Added the name "F-score" and an equation to define $\mu(f)$

22, in line 30 of page 8, please explain how the stress-strain diagram is obtained, and how the training data set is collected.

Added information.

23, in the experiments, please compare with other existing methods.

24, in Table II, please explain the meaning of the values. Are they mean values and standard deviations? If so, what does "best, worst, average" mean?

25, in Figure 12, please clarify the legends. For example, why are there two lines with

the same marker?

26, in line 56 of page 10, please explain "integration step". Is it the "h" in equation 17?

27, in section VII-E, please explain how the model is used in object classification. For example, what does "even for novel interactions" mean? Object classification typically tries to identify an object in a new environment. Besides, how is "Global" criteria evaluated in object classification?

28, in Table III, what does "test 1, test 2" mean? Please explain how these numbers are computed. Besides, object classification is typically evaluated with precision-recall curve.

29, please include more related works, since there are tons of articles on deformable object tracking.

We don't focus on tracking, we want to predict several frames ahead, not follow, that is why closely related works are few.

Associate Editor's Comments to Author:

Associate Editor Comments to the Author: (There are no comments. Please check to see if comments were included as a file attachment with this e-mail or as an attachment in your Author Center.)

I look forward to hearing from you.

Yours sincerely,

Dra. Verónica Esther Arriola Ríos
Profesora Asociada C de T.C.
Departamento de Matemáticas, Cub
119.
Facultad de Ciencias, UNAM
v.arriola@ciencias.unam.mx
+(52)55 5622 5426