Rapport d'évaluation du mémoire de thèse / Evaluation report of the PhD thesis

Doctorant **PhD student**	Nom prénom / Full name	David Pagnon
	Ecole Doctorale / Doctoral School	Mathématiques, Sciences et technologies de l'information, Informatique
	Titre thèse / PhD Title	Design and evaluation of a biomechanically consistent method for markerless kinematic analysis of sports motion

Rapporteur **Reviewer**	Nom prénom / Full name	Dr Steffi Colyer
	Etablissement / Institution	University of Bath
	Statut, fonction / Status, position	Lecturer (Associate Professo) in Biomechanics

Qualité du mémoire : structuration, rédaction & illustrations / Thesis quality, style & illustrations
(A titre indicatif/For information : Exceptionnel = top 5%, Très bon/very good = top 25 %)

⇨ Insatisfaisant / Unsatisfactory [ ] Satisfaisant / Satisfactory [ ] Bon / Good [ ]
Très bon / Very good [ ] Exceptionnel [x]

Commentaires/comments :
This is a well constructed thesis that covers a lot of ground and advances the knowledge at the intersection between computer vision and biomechanics in a unique manner. I find the studies are on-the-whole well thought-out and there is a logical flow across chapters. The thesis is presented very clearly, is easy to follow and the writing style is excellent. The illustrations throughout are exceptional and I find the visual abstracts particularly useful.

Contexte/ collaborations, background : état de l'art / state of the art :

⇨ Insatisfaisant / Unsatisfactory [ ] Satisfaisant / Satisfactory [ ] Bon / Good [ ]
Très bon / Very good [x] Exceptionnel [ ]

Commentaires/comments :
The evaluation of robustness and accuracy are extremely impressive, and these validations are thorough. The novelty of the thesis comes from the use of synthetic environemnt and the fact that Pose2Sim is totally open source nature. Whilst I applaud the candidate for trying to push the boundaries further and apply a lower grade system in challenging settings (BMX racing), however, I felt that this chapter was very incomplete and didn't really add much to the thesis. I felt that the discussions around the boxing analyses could have been been more complete, with reference to some 'smallest worthwhile detectable changes' based on the literature. For example, in boxing, how small are the differences you wish to detect, because they have been shown to be important from a performance perspective in previous literature? This would add greatly to the comparisons to other measurement systems that are already presented.

Qualité scientifique : méthodologie, expérimentations, validation
Scientific quality, methodology, experiments, validation

⇨ Insatisfaisant / Unsatisfactory [ ] Satisfaisant / Satisfactory [ ] Bon / Good [ ]
Très bon / Very good [ ] Exceptionnel [x]

Commentaires/comments :
The candidate has adopted a very novel and interesting approach to this problem. Specifically, the use of the virtual environment to test the robustness of the system is impressive. I have some specific questions about

advantages and disadvantages of the criterion measure, and the way data were processed in OpenSim, but overall the evaluation of the system is thorough and rigorous.

Apports personnels : originalité, valorisation, perspectives
Personal contributions : originality, exploitation and application of results, prospects.

⇨ Insatisfaisant / Unsatisfactory [ ] Satisfaisant / Satisfactory [ ] Bon / Good [ ]
Très bon / Very good [x] Exceptionnel [ ]

Commentaires/comments :

Whilst some groups have conducted similar validation work on markerless technologies, this is the first open-source markerless software platform accessible to biomechanists and non computer scientists, and the first to adopt a synthetic approach to assess system robustness. Thus, the originality of the thesis is excellent. As above, I do feel that the results could have been applied further to sport, drawing upon some values from the wider literature rather than solely comparing to existing motion capture systems. I feel this would provide extra context and evidence for the utility of the Pose2Sim programme for applied sport biomechanics analyses.

Conclusions du rapporteur / Reviewer's conclusions
Commentaires/comments :
Overall, I find the thesis and accompanying released software to be extremely impressive of a doctoral student and therefore give a favourable evaluation for the thesis. The contribution to the field is exceptional and, whilst I have some specific questions, I look forward to discussing the research further at the defence.

Avis du rapporteur / Reviewer's opinion :
Défavorable à la soutenance / Unfavorable to the defence [ ] Favorable [x]

Date 06/02/2023 Signature 

Visa du directeur de l'école doctorale :

Rapport détaillé, commentaires libres, questionnements
Detailed report, free comments, questions

Chapter 4
This is a very well written and thought-out chapter. The use of the synthetic environment to challenge the system in different ways is a very interesting and unique one, and this is where the novelty of the work can be found. I would like to see a little more discussion around the magnitude of the errors introduced by varying the set-up, in relation to known meaningful effects in the biomechanical studies of interest. To me, this would provide more evidence for the conclusions of the chapter.

Specific questions or comments (Chapter 4)

- Is the assumption that a person will not be variable step to step valid? Are you aware of coordination variability research and what the expected variability of e.g. gait would be?
- Regarding the extra processing step required in the situation where the occluded point had high confidence, how often did this happen? How robust is it? Are 'normal users' able to do this?
- Is there / is there a needed for movement specific models with particular joint constraints in built?

- It would be good to see more discussion of the size of the ‘errors’ introduced alongside some ‘smallest detectable differences’ from the literature.
- If a good calibration is ‘not that important’, are there any shortcuts you could make here?
- Figure 4.8. What do the SPM results add? Are the significant differences actually meaningful?
- Did you test any other gait event detection algorithms?
- 10% higher reprojection error seems like a lot. Why do you consider this to be trivial?
- Please comment on the limitation of markerless – the amount of time required to retrain to e.g. segment the spine further, whereas you can just put another couple of markers on.

Chapter 5

This chapter includes a very thorough and rigorous assessment of the accuracy of Pose2Sim. I have some specific questions below about the processing steps, however, my main questions relate to the use of synthetic data sets to assess real world accuracy, and I do feel that more discussion is needed around the differences between Pose2Sim and other similar research that assesses the accuracy of markerless motion capture. I believe that the novelty of the current study could be much clearer in the discussion of this chapter, as there are some key differences to previous related literature, currently not fully debated.

Specific questions or comments (Chapter 5)

- Since this was written there have been some further publications of this nature (e.g. Needham et al. 2022). How does their approach differ to yours / what are the strengths and weaknesses?
- Have you really evaluated the ability to detect pose in the real world if green screens and synthetic images have been used?
- Very low cut-off frequency for kinematics. How was this decided?
- How does the CMC work?
- Why did you use joint centres to drive marker-based model?
- How do your results compare vs. Needham et al.?
- How can we get better at dealing with occlusions?
- Are further validations in different clothing required to fully evaluated Pose2Sim?
- How could the ‘systematic’ offset (which actually might not be systematic at extreme ROMs) be dealt with?

Chapter 6

It is great to see a more practical system being tested in a truly applied setting and it is impressive to see how the candidate dealt with many of the challenges posed by collecting this kind of data in a less controlled environment. I have some specific questions about some of the decisions around the protocol and I would like to see a little more discussion around the errors in relation to the minimum detectable effect, specifically for boxing.

Specific questions or comments (Chapter 6)

- Were coaches consulted on the KPIs?
- Is the sampling frequency too low for such a fast movement?
- Is the model the same as previous chapter? Planar vs. 3D motion
- How large are these differences relative to known meaningful differences (e.g. between skilled and novice boxers)?
- How is the multi-person work going?
- Do you actually need neural nets to get GRFs?

Chapter 7

As above, I am not sure that this forms a whole chapter of the thesis as the analysis (whilst I applaud the effort) is very incomplete and there aren’t really conclusions that can be drawn.

Specific questions or comments (Chapter 7)

- Were coaches consulted on the KPIs?

Broader questions related to the thesis

- What are the disadvantages and advantages of sparse-point pose estimation vs. e.g. silhouette / volumetric / SMPL methods? Why did you decide to adopt the sparse point approach?
- Related to validation and ‘silver standard’ criterion measures. In an ideal world, what methods would you use for gold standard validation?
- Regarding the use of synthetic images. How do you know / could you test whether the scenes are realistic enough? What are the limitations of this approach?
- To date, what movements has Pose2Sim been applied to? Many 3D rotational?]