Entropy-based metrics for joint image alignment

Report Name	Outline Project Specification
Author (User Id)	Laura Collins (lac32)
Supervisor (User Id)	Neil MacParthalain (ncm)
Module	CS39440
Degree Scheme	GH7P (Artificial Intelligence And Robotics)
Date	February 1, 2016
Revision	0.1
Status	Draft

1 Project description

The Project builds upon the Congealing code developed by Erik Learned-Miller from the University of Massachusetts [2]. In the paper Learned-Miller utilises the Congealing algorithm to align both MNIST handwriting data and MRI Scans by reducing the variability between each pixel within a stack of images. For example, the algorithm will analyse the properties of the upper-leftmost pixel in image 1, then in image 2 and through each image in the set. This variation between pixels is known as Entropy, and thus the objective of Congealing is to minimise the Entropy across a set of images.

The aim of this project will be to implement the Congealing algorithm using not only standard Entropy, as in Learned-Miller's application, but using other metrics, as discussed in Section 2.1. As for the data to feed into the application, my image set of choice will be Mammography scans, however due to their large file size initial experimentation will be run using the MNIST handwriting data set.

Aligning mammography scans would help build an average of the image set. With this is in mind it is possible to build an average for the 4 breast tissue composition categories outlined by BI-RADS (Breast Imaging-Reporting and Data System) - with 1 being fatty tissue (lowest risk) and 4 being dense tissue (highest risk) [3]. Approaching this alignment of images using different metrics would output different a range of output results, and further investigation could be undertaken to ascertain the best alignment metric.

Finally, the resulting software tool could be used by Radiologists in assessing the tissue variation between the 4 categories and could also be used to build a classifier model.

I decided to use an adapted-XP approach for my software methodology.

2 Proposed tasks

2.1 Research into alternative metrics

The original Congealing algorithm implemented by Learned-Miller uses standard Entropy and thus I would have to research into different kinds of Entropy-based metrics. This led me on to the paper by **citation needed** about different Fuzzy Entropy implementations. Each one is implemented with slight variation, and thus each gives a different output.

After reading about different metrics used in Congealing, I discovered Cox et al's Least Squares implementation [1], however I decided to stay aligned with the Major Project title, and implemented only entropy-based metrics.

2.2 Graphical User Interface

2.3 Dataset

3 Project deliverables

3.1 Technical Deliverable

3.2 Final Report

3.3 Possible Publication

Annotated Bibliography

[1] M. Cox, S. Sridharan, S. Lucey, and J. Cohn, "Least squares congealing for unsupervised alignment of images," in *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*. IEEE, June 2008, pp. 1–8. [Online]. Available: http://dx.doi.org/10.1109/cvpr.2008.4587573

> Futher builds upon Learned-Miller's Congealing approach with faster convergence and no pre-defined step size. Results are presented on both the MNIST handwriting data set (as in the original Learned-Miller paper) and the MultiPIE face database. This approach is extremely computationally costly, and this would have to be assessed in terms of computing mammography scans.

[2] E. G. Learned-Miller, "Data driven image models through continuous joint alignment," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 28, no. 2, pp. 236–250, Feb. 2006. [Online]. Available: http://dx.doi.org/10.1109/tpami.2006.34

> Something here.

[3] A. C. Society, "Understanding your mammogram report BI-RADS categories," http://www.cancer.org/treatment/understandingyourdiagnosis/examsandtestdescriptions/mammogramsandotherbreastimagingprocedures/mammograms-and-other-breast-imaging-procedures-mammo-report, 2015, accessed January 2016.

> Website outlining Mammography scan result guidelines and the 4 tissue density composition categories. Particularly important for understanding the differences between the 4 sets of images, and how best to name them.