ECAI 2016
G.A. Kaminka et al. (Eds.)

doi:10.3233/978-1-61499-672-9-1688

Shape Invariant Formulation for Change Point Models in Multiple Dimensions

Jesse Glass[1] and **Zoran Obradovic**[1]

Abstract.

Artificial intelligence has achieved superhuman performance in a variety of tasks. Unfortunately, this is often done without interpretable methods. In medicine, it is not sufficient to have an algorithm with maximum accuracy. The methods must be evaluated by experts. Our motivating task is to detect features of the eye using unlabeled data. We detect features using a Bayesian changepoint model. Changepoint detection can perform object recognition when applied to two dimensional feature spaces such as images. We present a formula for detecting any shape with the changepoint model. We then extend it to multiple features in order to capture the color of an object. The work presented here has the ability to incorporate prior information, the ability to handle images of varying granularity, can provide confidence estimates on object features in an image. It will serve as the foundation for a classification method with interpretable results.

1 INTRODUCTION, BACKGROUND, & RELATED WORK

The task of predicting the severity of diabetic retinopathy using images of the retina was a recent Kaggle competition. A significant majority of the top 100 performers used Deep Learning approaches. This approach has several drawbacks. When applying Convolutional Neural Networks, the pixel sizes of the images must be the same. This requires that many of the images are downgraded, which results in information loss. Our approach can learn on varying granularities which potentially enables us to transfer knowledge learned on more detailed images to the less detailed ones.

Another drawback with Deep Learning Models, is a lack of interpretability. Our model provides the location, color, and size of an object with probability estimates for any combination. Detecting objects such as the basal ganglion helps with the prediction task of labeling veins because vascular growth stems out from the basal ganglion. Similarly detecting the macula helps during diagnosis because stage 2 retinopathy is most often diagnosed by determining whether a hemorrhage is blocking light from entering the macula. Importantly, the accuracy of the detection of a macula and hemorrhage can be easily verified by physicians.

The Gamma-Poisson conjugate prior has a closed form Bayesian posterior for the changepoint problem. The Poisson distribution is an obvious choice for image pixel value prediction since a Poisson distribution outputs a natural number and images are coded with integers from 0 to 255 on three color channels.

Bayesian models also give us a unique opportunity to incorporate prior information. Then evaluate/update that prior information. The images in this data set have a great variety some are dark, some have glare, some have artifacts. A well designed model will not require preprocessing to handle these.

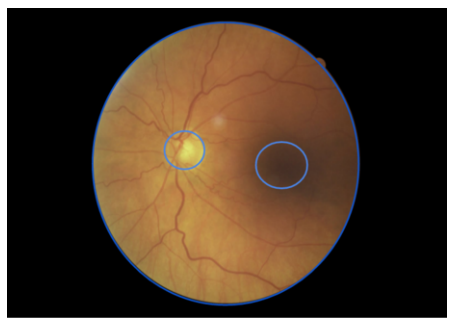

Figure 1. Retina (whole), Optic Nerve (light), Macula (dark)

Bayesian methods can often seem too specific for adaptation to multiple purposes. [1] derived an exact solution for multiple features for changepoints along a single dimension. [2] used heuristics to model a single feature, with the potential to model stacked signals. Bayesian models also have a much better ability to learn confidence. Heuristics can work but often fail miserably in outlier cases. So, we developed a fully Bayesian extension of both of the discussed models.

2 METHOD

Given the following Gamma priors on our average values and Poisson priors on our observations we will develop a Gibbs sampling algorithm for multiple features on arbitrary shapes. We allow different objects within the image (segments) and the different color channels to have different priors (a,b). The notation is: S, shape; f, feature (RGB layer); ij, position in image. The generative model is as follows:

$$\lambda_i^{(f)} \sim Gamma(a,b)$$

$$x_{ij}^{(f)} \sim \begin{cases} Poisson(x_{ij}^{(f)}|\lambda_1^{(f)}) & i,j \notin S \\ Poisson(x_{ij}^{(f)}|\lambda_2^{(f)}) & i,j \in S \end{cases}$$

The problem of inferring the posterior over the latent variables S, λ_1, λ_2 can be solved via Bayes theorem. The full joint distribution

$$p(\vec{\lambda}_1, \vec{\lambda}_2, S|\boldsymbol{X}) \propto p(\boldsymbol{X}|\vec{\lambda}_1)p(\boldsymbol{X}|\vec{\lambda}_2)p(\vec{\lambda}_1)p(\vec{\lambda}_2)p(S)$$

$$\propto \prod_{f=1}^{M}(\prod_{ij\notin S} p(x_{ij}^{(f)}|\lambda_1^{(f)}) \prod_{ij\in S} p(x_{ij}^{(f)}|\lambda_2^{(f)})p(\lambda_1^{(f)})p(\lambda_2^{(f)}))p(S).$$

The probability of the shape depends on the number of features required for the shape. We define a square by its start and end points.

[1] Temple University, Center for Data Analytics and Biomedical Informatics, email: tud25892@temple.edu, zoran.obradovic@temple.edu

We define a circle by is center, (x,y), and radius, r.

$$p(S) = p(x)p(y)p(r)$$

We see that each neighboring x_{ij} is independent of its neighbors given that it is drawn from a known underlying gamma process. This enables us to plug known prior distributions into the above formula.

Obtain the posterior conditionals for each latent variable by collecting the terms in the full joint distribution that include the latent variable of interest. Derive the posterior conditionals for each of the variables S, λ_1, λ_2 by selecting relevant components from the full joint distribution. When calculating the conditional probability we can marginalize constants out of $p(\lambda_1|\cdot)$.

$$p(\lambda_1^{(f)}|S, \vec{\lambda}_1^{-f}, \vec{\lambda}_2, \boldsymbol{X}) \propto Gamma(a + \sum_{ij \in S}(x_{ij}^f), \sum_{ij \in S}(1) + b)$$

$$p(\lambda_2^{(f)}|S, \vec{\lambda}_2^{-f}, \vec{\lambda}_1, \boldsymbol{X}) \propto Gamma(a + \sum_{ij \notin S}(x_{ij}^f), \sum_{ij \notin S}(1) + b)$$

We can see that each $\lambda^{(f)}$ is independent of one another. Calculating the Multinomial Distribution of $p(S|\vec{\lambda_1}, \vec{\lambda}_2, \boldsymbol{X})$ is a bit more complex, and must be done separately for each component of S. We defined the properties of a circle as two center points and a radius. They all have the similar formulations. We present the multinomial formula for the probability of the radius below.

$$\begin{aligned} log(p(r|\vec{\lambda_1}, \vec{\lambda}_2, \boldsymbol{X})) =^+ & \sum_{f=1}^{M}(log(\lambda_1^{(f)}) \sum_{ij \in S} x_{ij}^{(f)}) + \\ & - M \cdot \sum_{ij \in S}(1) \cdot \sum_{f=1}^{M} \lambda_1^{(f)} + \\ & + \sum_{f=1}^{M}(log(\lambda_2^{(f)}) \sum_{ij \notin S} x_{ij}^{(f)}) - M \cdot \sum_{ij \notin S}(1) \cdot \sum_{f=1}^{M} \lambda_2^{(f)}. \end{aligned}$$

After calculating each possible solution, we take the exponent and then normalize.

3 EXPERIMENTS & RESULTS

A set of synthetic experiments were devised in increasing complexity to examine this approaches ability to handle the data. First, we display a change point algorithm run for one and two change points on 1-D data. Then, the models for those are extended to the two dimensional case - the square. Then, we change the shape to a circle. Last, three features are incorporated endowing objects in the image with a color. For all of the synthetic experiments, the latent values of the objects were learned without mistakes in 100% of trials.

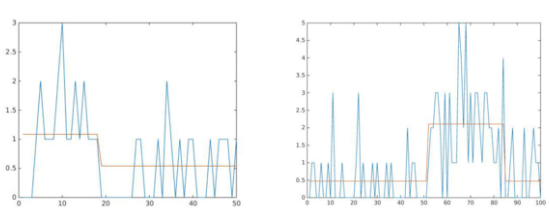

Figure 2. Single Change Point (left) & Segment Shift (right)

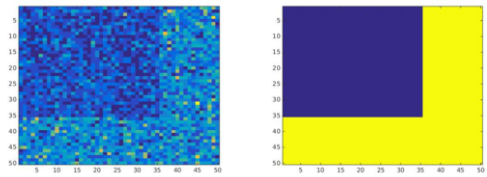

Figure 3. Two Dimensional Change Point

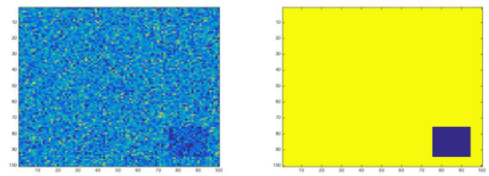

Figure 4. Two Dimensional Segment Shift (a.k.a. Square Detection)

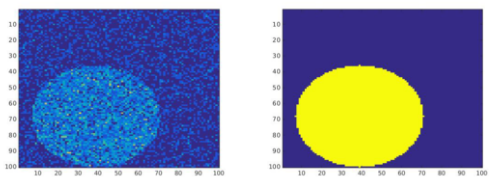

Figure 5. Two Dimensional Circle Detection

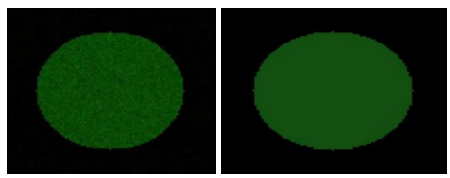

Figure 6. Color Circle Detection

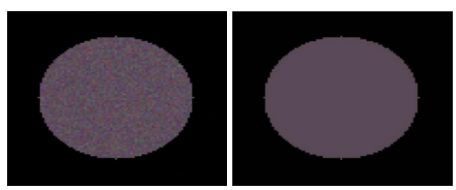

Figure 7. Color Circle Detection

4 CONCLUSION

[1] and [2] are both one dimensional solutions. [2] attempts to handle overlapped signals using heuristics while [1] develops an exact solution for a similar problem but does not handle stacked signals. The next step in this work is to handle stacked signals. Once we can identify the key circular features of the eye, we can attempt an approach for the vasculature and detecting hemorrhages.

ACKNOWLEDGEMENTS

This research was supported in part by DARPA grant FA9550-12-1-0406 negotiated by AFOSR, NSF BIGDATA grant 14476570, ONR grant N00014-15-1-2729 and SNSF Joint Research project (SCOPES), ID: IZ73Z0_152415.

REFERENCES

[1] Paul Fearnhead, 'Exact and efficient bayesian inference for multiple changepoint problems', in *Statistical Computing*, pp. 203–213, (2006).

[2] W.R. Lai, M.D. Johnson, R. Kucherlapati, and P.J. Park, 'Comparative analysis of algorithms for identifying amplifications and deletions in array cgh data', in *Bioinformatics*, pp. 3763–3770, (2005).