Online Statistical Inference for Large-Scale Binary Images

Moo K. Chung, Ying Ji Chuang, Houri K. Vorperian
University of Wisconsin-Madison, USA
email:mkchung@wisc.edu

Introduction

An *online algorithm* is one that processes its inputted data in a sequential manner (Karp, 1992). Instead of processing the entire set of imaging data from the start, an online algorithm processes one image at a time. That way, we can bypass the memory requirement, reduce numerical instability and increase computational efficiency. Since medical image processing is often done semi-automatically, the resulting images may be available at the same time. Further, modern medical imaging datasets are too large to fit into a computer's memory. Thus, there is a need to develop an iterative analysis framework where the final statistical maps are updated sequentially each time a new image is added to the analysis.

Online algorithm for *t*-test

Given images $x_1, \cdots, x_m$, an *online algorithm* for computing the sample mean image μ_m is given by

$$\mu_m = \frac{1}{m}\sum_{i=1}^{m} x_i = \mu_{m-1} + \frac{1}{m}(x_m - \mu_{m-1}).$$

An online algorithm for computing the sample variance map σ_m^2 is algebraically involved (Chan et al., 1983; Knuth, 1981). It can be shown that

$$\begin{aligned}\sigma_m^2 &= \frac{1}{m-1}\sum_{i=1}^{m}(x_i - \mu_m)^2 \\ &= \frac{m-2}{m-1}\sigma_{m-1}^2 + \frac{1}{m}(x_m - \mu_{m-1})^2\end{aligned}$$

for $m \geq 2$. The algorithm starts with the initial value $\sigma_1^2 = 0$.

Given measurements $x_1, \cdots, x_m \sim N(\mu^1, (\sigma^1)^2)$ in one group and $y_1, \cdots, y_n \sim N(\mu^1, (\sigma^2)^2)$ in the other group, the two-sample t-statistic for testing $\mu^1 = \mu^2$ at each voxel is given by

$$T_{m,n} = \frac{\mu_m^1 - \mu_n^2 - (\mu^1 - \mu^2)}{\sqrt{(\sigma^1)_m^2/m + (\sigma^2)_n^2/n}},$$

where μ_m^1, μ_n^2, $(\sigma^1)_m^2, (\sigma^2)_m^2$ are sample means and variances in each group estimated using the online algorithm. $T_{m,n}$ is then sequentially computed as

$$T_{1,0} \to T_{2,0} \to \cdots \to T_{m,0} \to T_{m,1} \to \cdots \to T_{m,n}$$

in $m + n$ steps.

Online algorithm for regression

$(m-1)$-th step. Given data vector $\mathbf{y}_{m-1} = (y_1, \cdots, y_{m-1})'$ and design matrix Z_{m-1}, consider linear model

$$\mathbf{y}_{m-1} = Z_{m-1}\boldsymbol{\lambda}_{m-1}$$

with unknown parameters $\boldsymbol{\lambda}_{m-1} = (\lambda_1, \lambda_2, \cdots, \lambda_k)'$. The least squares estimation (LSE) of $\boldsymbol{\lambda}_{m-1}$ is then given as

$$\boldsymbol{\lambda}_{m-1} = W_{m-1}^{-1} Z_m' \mathbf{y}_{m-1},$$

where $W_{m-1} = Z_{m-1}'Z_{m-1}$ is a $k \times k$ matrix.

m-th step. When new data y_m is introduced to the model, it is updated as

$$\begin{pmatrix}\mathbf{y}_{m-1} \\ y_m\end{pmatrix} = \begin{pmatrix}Z_{m-1} \\ z_m\end{pmatrix}\boldsymbol{\lambda}_m,$$

where z_m is a row vector of size $1 \times k$. Subsequently, we have

$$W_{m-1}'\boldsymbol{\lambda}_{m-1} + z_m' y_m = (W_{m-1} + z_m' z_m)\boldsymbol{\lambda}_m.$$

Using Woodbury formula (Deng, 2011), we can show that the estimated parameters $\boldsymbol{\lambda}_{m-1}$ are updated as

$$\begin{aligned}\boldsymbol{\lambda}_m = (I - W_{m-1}^{-1} z_m' y_m - c_m W_{m-1}^{-1} z_m' W_{m-1}')\boldsymbol{\lambda}_{m-1} \\ -c_m W_{m-1}^{-1} z_m' z_m' y_m,\end{aligned}$$

where $c_m = 1/(1 + z_m W_{m-1} z_m')$ is scalar and I is the identity matrix of size $k \times k$. The algorithm needs to start from

$$\boldsymbol{\lambda}_k \to \boldsymbol{\lambda}_{k+1} \to \cdots \to \boldsymbol{\lambda}_m.$$

Our algorithm does *not* require the factorization or direct inversion of matrices and thus more efficient compared to real-time fMRI analysis (Bagarinao et al., 2006), where the Cholesky factorization was used to invert the covariance matrix.

Online algorithm for *F*-test

Let y_i be the i-th image, $\mathbf{x}_i = (x_{i1}, \cdots, x_{ip})'$ to be the variables of interest and $\mathbf{z}_i = (z_{i1}, \cdots, z_{ik})'$ to be nuisance covariates. We assume there are $m-1$ images to start with. Consider a general linear model

$$\mathbf{y}_{m-1} = Z_{m-1}\boldsymbol{\lambda}_{m-1} + X_{m-1}\boldsymbol{\beta}_{m-1},$$

where $Z_{m-1} = (z_{ij})$ is $(m-1) \times k$ design matrix, $X_{m-1} = (x_{ij})$ is $(m-1) \times p$ design matrix. $\boldsymbol{\lambda}_{m-1} = (\lambda_1, \cdots, \lambda_k)'$ and $\boldsymbol{\beta}_{m-1} = (\beta_1, \cdots, \beta_p)'$ are unknown parameter vectors to be estimated.

Reduced model. The reduced model corresponding to null hypothesis $\boldsymbol{\beta} = 0$ is

$$\mathbf{y}_{m-1} = Z_{m-1}\boldsymbol{\lambda}_{m-1}^0.$$

The goodness-of-fit of the null model is measured by the sum of the squared errors (SSE):

$$\mathrm{SSE}_{m-1}^0 = (\mathbf{y}_{m-1} - Z_{m-1}\boldsymbol{\lambda}_{m-1}^0)'(\mathbf{y}_{m-1} - Z_{m-1}\boldsymbol{\lambda}_{m-1}^0),$$

where $\boldsymbol{\lambda}_{m-1}^0$ is estimated sequentially using online algorithm for regression:

$$\mathrm{SSE}_k^0 \to \mathrm{SSE}_{k+1}^0 \to \cdots \to \mathrm{SSE}_m^0.$$

Full model. The fit of full model is similarly measured by

$$\mathrm{SSE}_{m-1}^1 = (\mathbf{y}_{m-1} - \mathbb{Z}_{m-1}\boldsymbol{\gamma}_{m-1}^1)'(\mathbf{y}_{m-1} - \mathbb{Z}_{m-1}\boldsymbol{\gamma}_{m-1}^1),$$

where $\mathbb{Z}_{m-1} = [Z_{m-1} X_{m-1}]$ and

$$\boldsymbol{\gamma}_{m-1}^1 = \begin{pmatrix}\boldsymbol{\lambda}_{m-1}^1 \\ \boldsymbol{\beta}_{m-1}^1\end{pmatrix}.$$

SSE for the full model is also estimated sequentially:

$$\mathrm{SSE}_{k+p}^1 \to \mathrm{SSE}_{k+1}^1 \to \cdots \to \mathrm{SSE}_m^1.$$

Under null hypotheis $\boldsymbol{\beta} = 0$, the test statistic at the m-th iteration f_m is given by

$$f_m = \frac{(\mathrm{SSE}_0 - \mathrm{SSE}_1)/p}{\mathrm{SSE}_0/(m-p-k)} \sim F_{p,m-p-k},$$

which is the F-statistic with p and $m-p-k$ degrees of freedom.

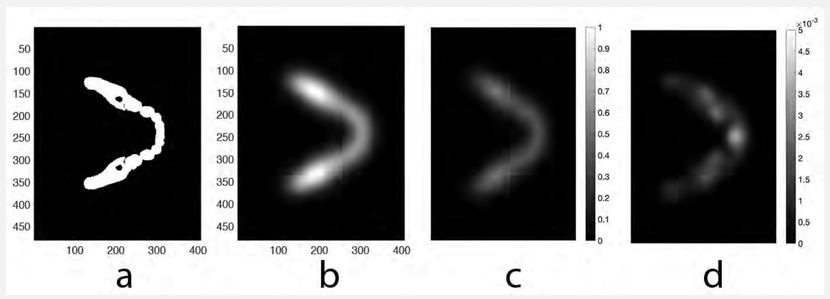

Fig. 1. (a) A representative mandible binary segmentation that are affine registered to the template. (b) Gaussian kernel smoothing of segmentation with bandwidth $\sigma = 20$. Smoothing can easily patch topological artifacts such as cavities and handles. The sample mean (c) and variance (d) of the smoothed maps computed using the online algorithms.

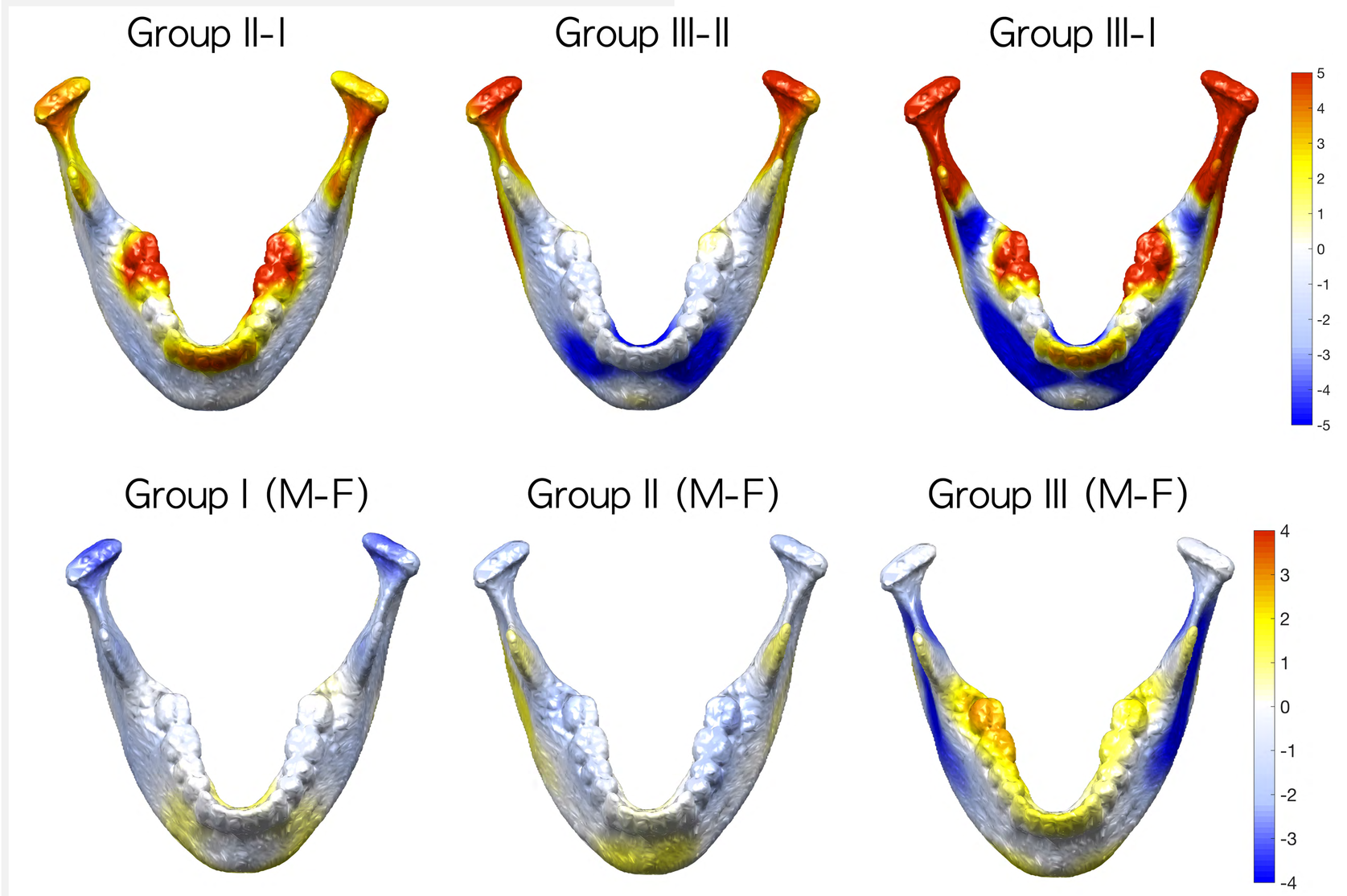

Fig. 2. Top: t-stat. maps showing mandible growth. The elongation of mandible is shown between Groups II and III, and I and III. The condyle regions show prominent growth in Group III- I comparison. At the same time, the elongation is shown as negative growth (dark blue). Bottom: t-stat. maps (male - female) showing sex differences in each age group. There were no significant sex differences in groups I and II. However, pubertal and post-pubertal sex differences are evident in group III that starts at age 13.

Application

Subjects. The dataset consisted of 290 typically developing individuals. The age distribution of the subjects is 9.66 ± 6.34 years. The minimum age was 0.17 years and maximum age was 19.92 years. A total of 160 male and 130 female subjects were divided into 3 groups. Group I (age below 7) contained 130 subjects. Group II (between 7 and 13) contained 48 subjects. Group III (between 13 and 20) contained 112 subjects.

Image preprocessing. The mandibles in CT were semi-automatically segmented using an in-house processing pipeline that involves image intensity thresholding (Kelly et al., 2017). The segmented binary images were then affine registered to the mandible labeled as `F226-15-04-002-M`, which served as the template (Fig. 1). We smoothed the binary images with Gaussian kernel with bandwidth $\sigma = 20$ voxels (Fig. 1). Since the CT image resolution is 0.35mm, 20 voxel wide bandwidth is equivalent to 7mm. The bandwidth was chosen to reflect the size of missing teeth and cavities. Any smaller filter size will not mask large missing teeth and cavites. The average of all 290 smoothed binary images was used as the final template and distributed as a potential prior map for more advanced shape modeling: `http://www.stat.wisc.edu/~mchung/VBM`.

Since the statistic maps are correlated over voxels, we corrected multiple comparisons using the random field theory (Worsley et al., 1998). Using the proposed online algorithm, statistics are computed sequentially by adding one image at a time and tested for age and sex effects.

Age effects. We performed the t-test to assess age effects between the groups (Fig. 2-top). Voxels above or below ± 4.41, 4.37 and 4.43 were considered significant in the t-statistics between age groups I and II, II and III, and I and III respectively at the 0.05 level. The dark red regions show positive growth (bone deposition) and dark blue regions show negative growth (bone resorption). The findings are consistent with previous studies based on 2D surface deformation (Chung et al., 2015) and landmarks (Kelly et al., 2017).

Sex effects. Within each group, we tested the significance of sexual dimorphism by performing the two-sample t-test between males and females (Fig. 2-bottom). Any region above or below $\pm$ 4.37, 4.89 and 4.50 (for group I, II and III respectively) were considered significant at 0.05 level. In group I and II, there is no gender differences. In group III, the statistical significance is localized in the regions between Condyle and Gonion in the both sides. Such findings are consistent with general findings on sexual dimorphism that become evident during puberty.

References

E. Bagarinao, T. Nakai, and Y. Tanaka. Real-time functional MRI: development and emerging applications. *Magnetic Resonance in Medical Sciences*, 5:157–165, 2006.

T.F. Chan, G.H. Golub, and R.J. LeVeque. Algorithms for computing the sample variance: Analysis and recommendations. *The American Statistician*, 37:242–247, 1983.

M.K. Chung, A. Qiu, S. Seo, and H.K. Vorperian. Unified heat kernel regression for diffusion, kernel smoothing and wavelets on manifolds and its application to mandible growth modeling in CT images. *Medical Image Analysis*, 22:63–76, 2015.

C.Y. Deng. A generalization of the Sherman–Morrison–Woodbury formula. *Applied Mathematics Letters*, 24:1561–1564, 2011.

R.M. Karp. On-line algorithms versus off-line algorithms: How much is it worth to know the future? In *IFIP Congress*, volume 12, pages 416–429, 1992.

M.P. Kelly, H.K. Vorperian, Y. Wang, K.K. Tillman, H.M. Werner, M.K. Chung, and L.R. Gentry. Characterizing mandibular growth using three-dimensional imaging techniques and anatomic landmarks. *Archives of Oral Biology*, 2017.

D. Knuth. The art of computing, vol. ii: Seminumerical algorithms, 1981.

K.J. Worsley, J. Cao, T. Paus, M. Petrides, and A.C. Evans. Applications of random field theory to functional connectivity. *Human Brain Mapping*, 6:364–7, 1998.