Variational Inference for Blind Image Deconvolution

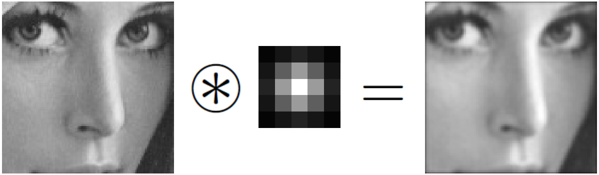
Antonie Brožová

24.6.2021

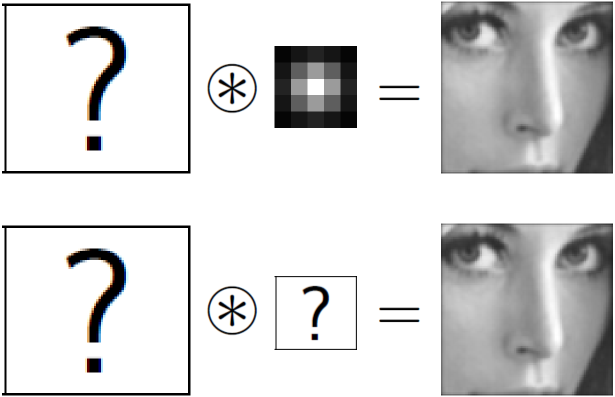
Blurred image model

Blurred image is convolution of sharp image and point spread function (PSF) (assuming space-invariant PSF)

$$\boldsymbol{D} = \boldsymbol{X} \circledast \boldsymbol{K} + \boldsymbol{n}$$



Blind image deconvolution



Bayesian approach

Convolution in matrix form + gaussian noise

$$\boldsymbol{d} = \boldsymbol{X}\boldsymbol{k} + \boldsymbol{n}, \qquad \boldsymbol{n} \sim \mathcal{N}\left(\mathbf{0}, \omega\mathbb{I}\right),$$

Priors must be well-chosen to avoid trivial solution

$$p(\boldsymbol{x}, \boldsymbol{k}|\boldsymbol{d}) \propto p(\boldsymbol{d}|\boldsymbol{x}, \boldsymbol{k})p(\boldsymbol{x}, \boldsymbol{k}),$$

Variational Bayes

Approximation of posterior distribution

$$p(\boldsymbol{\theta}|\boldsymbol{d}) \sim q(\boldsymbol{\theta}|\boldsymbol{d}),$$
$$q(\boldsymbol{\theta}|\boldsymbol{d}) = \prod_i q(\theta_i|\boldsymbol{d})$$
$$q(\boldsymbol{\theta_i}|\boldsymbol{d}) \propto \exp\left[\mathbb{E}_{q(\boldsymbol{\theta}_{\setminus i}|\boldsymbol{d})}\left[\ln p(\boldsymbol{\theta}, \boldsymbol{d})\right]\right],$$

Kullback-Leibler divergence of q from p

$$KL\left(q(\boldsymbol{\theta}|\boldsymbol{d}) \parallel p(\boldsymbol{\theta}|\boldsymbol{d})\right) = \mathbb{E}_{q(\boldsymbol{\theta}|\boldsymbol{d})}\left[\ln \frac{q(\boldsymbol{\theta}|\boldsymbol{d})}{p(\boldsymbol{\theta}|\boldsymbol{d})}\right]$$

IVB algorithm

- Conjugate distributions lead to a system of linear equations
- Iteratively minimizes Kullback-Leibler divergence
- Fast, but requires system of linear equations

ELBO - Evidence Lower Bound

$$KL\left(q(\boldsymbol{\theta}|\boldsymbol{d}) \parallel p(\boldsymbol{\theta}|\boldsymbol{d})\right) = \mathbb{E}_{q(\boldsymbol{\theta}|\boldsymbol{d})}\left[\ln \frac{q(\boldsymbol{\theta}|\boldsymbol{d})}{p(\boldsymbol{\theta}, \boldsymbol{d})}\right] + p(\boldsymbol{d}) =$$
$$= -\mathcal{L} + p(\boldsymbol{d})$$
$$\mathcal{L} = p(\boldsymbol{d}) - KL\left(q(\boldsymbol{\theta}|\boldsymbol{d}) \parallel p(\boldsymbol{\theta}|\boldsymbol{d})\right)$$

- Max ELBO = Min KL
- Gradient descent - slow, but system of linear equations is not needed

Reparametrization trick

- $\mathcal{L}$ requires gradients of expected values with respect to $q(\boldsymbol{\theta}|\boldsymbol{d})$ to be known
- If $q(\theta|d) \sim \mathcal{N}\left(\mu, \sigma^2\right)$, then θ can be reparametrized as $\theta = \sigma\epsilon + \mu, \epsilon \sim \mathcal{N}(0, 1)$
- $\nabla_\mu \mathbb{E}_{q(\theta|d)}\left[f(\theta, d)\right] \approx \frac{1}{L}\sum_{l=1}^{L} \nabla_\mu \left(f(\sigma\epsilon^{(l)} + \mu, d)\right)$

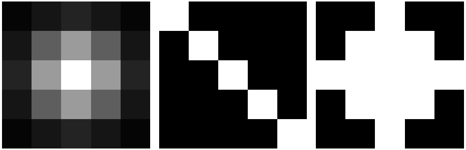
Vadam

- Vadam - Variational Adam optimiser proposed by Khan et al. (2018)
- Derived for distributions from exponential family to maximize ELBO
- The update of posterior parameters is derived from update of natural parameters
- Computes gradient in sample drawn from posterior distribution - the gradient descent is stochastic

Comparison of two algorithms

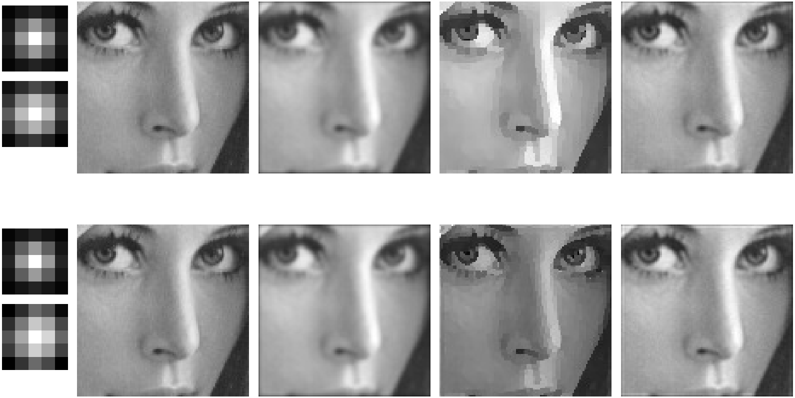
1. Only IVB steps
2. IVB steps for all parameters except for mean and covariance of PSF $\boldsymbol{\mu}_k$ a $\boldsymbol{\Sigma}_k$, which are estimated using gradient descent

Sharp Image and PSF

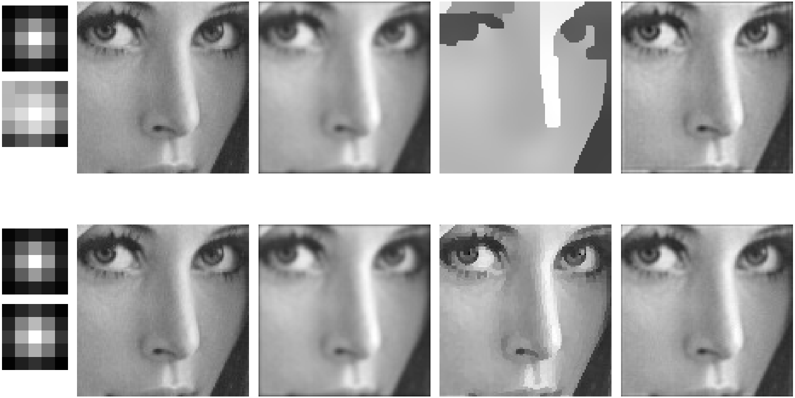


- Posterior distribution $p(\boldsymbol{k}|\boldsymbol{d}) = \mathcal{N}(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$
- Image is assumed to be piece-wise constant

IVB and ELBO for diagonal covariance matrix



IVB and ELBO for full covariance matrix



Higher number of samples for reparametrization trick

- The algorithm takes only one sample of ϵ to estimate $\boldsymbol{\mu}_k$ a $\boldsymbol{\Sigma}_k$
- One sample is enough only in the case of diagonal covariance matrix
- Taking 25 samples gives better results, but the reconstruction still isn't as accurate as in the case of full covariance matrix in IVB algorithm
- Higher number of sample would increase accuracy of reconstruction, but it is too costly

Vadam

- Covariance matrix can be approximated using gradient with respect to mean in case of normally distributed random variable
- Reaches the same accuracy as IVB algorithm with full covariance matrix, even though it uses only diagonal one

PSNR of the mean of image posterior

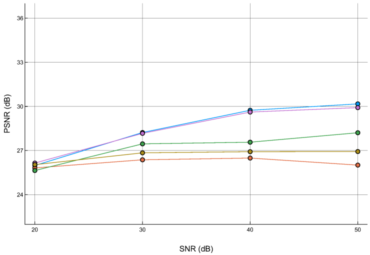


Figure: Blue line shows results of IVB alg. with full covariance, red line of IVB alg. with diagonal covariance, green line of ELBO with diagonal covariance, purple of ELBO with Vadam and brown PSNR of blurred image.

PSNR of non-blind reconstruction

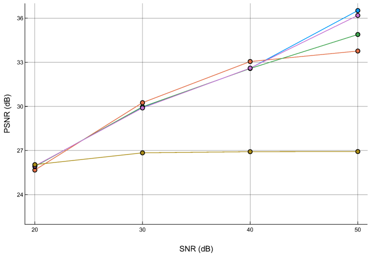


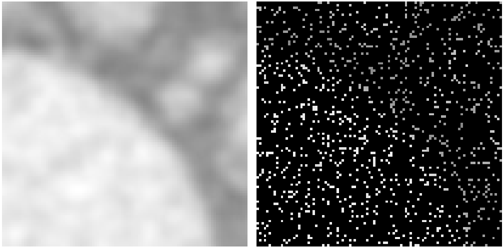
Figure: Blue line shows results of IVB alg. with full covariance, red line of IVB alg. with diagonal covariance, green line of ELBO with diagonal covariance, purple of ELBO with Vadam and brown PSNR of blurred image.

Conclusion

- Steps of IVB algorithm can be replaced by maximization of ELBO using Vadam optimiser without loss of accuracy, which is more flexible than IVB algorithm

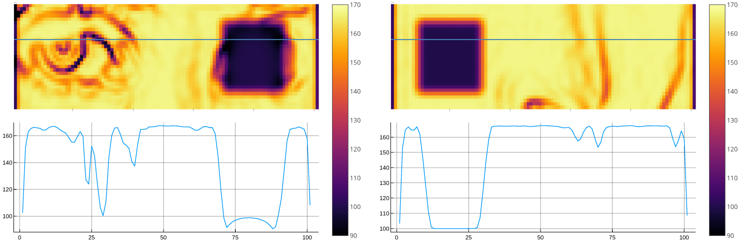
Incomplete image

- Pixels are scanned one by one in environmental scanning microscopy
- Scientists usually aren't interested in the whole image, only in a certain area
- Presented algorithms estimate not only images and PSFs, but also variance in the image

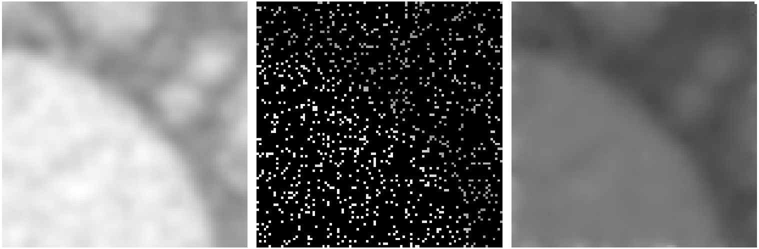


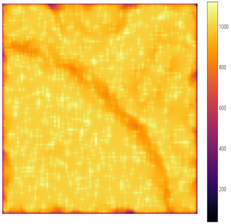
Posterior precision in missing areas

- Two areas in image were deleted - one with high variance and one with almost constant values
- Lower figure shows values of posterior precision in pixels marked with blue line in image above



Precision estimate on image from electron scanning microscope





Conclusion

- Estimation of parameters with ELBO and Vadam shows the same accuracy as estimate with no stochasticity
- The algorithm has potential to find more interesting areas in image
- Future work: analyse the algorithm on larger images with more realistic blurs

References:

- Microscope image from Ing. Vilém Neděla, Ph.D. from Institute of Scientific Instruments CAS in Brno.
- Lena [online]. Available from: `https://www.cosy.sbg.ac.at/~pmeerw/Watermarking/lena.html`, cited 2020-10-12.
- KINGMA, D. a M. WELLING, Auto-Encoding Variational Bayes. ArXiv:1312.6114v10 [stat.ML], 2014.
- KHAN, Mohammad Emtiyaz a et al., Fast and Scalable Bayesian Deep Learning by Weight-Perturbation in Adam. ArXiv:1806.04854v3 [stat.ML], 2018.