Korea Advanced Institute of Science and Technology

School of Electrical Engineering

EE636 Digital Video Processing Fall 2018

Student's Name: Dinh Vu

Student's ID: 20184187

Programming Assignment 2

(a) MAP without occlusion and line fields

The main problem in MAP is the optimization, so it demands a method to make sure the convergence. The convergence of the procedure is strongly related to the simulated annealing schedule. Geman [1] proposed the following temperature schedule:

$$T = \frac{\tau}{\ln(k+1)}, \qquad k = 1,2,3,\ldots$$

Where τ is a constant, set to equal to 0.2 and k is the iteration cycle, the number of iteration is set to equal to 8.

The Metropolis algorithm is used to create new candidate randomly and check whether motion vectors are updated or not. After each iteration, the potential function is calculated and compared with the previous value.

$$\Delta p = p_{new} - p$$

The motion vectors are updated if the below condition is true:

$$(\Delta p \leq 0) \text{ or } \left\{\left[\exp\left(-\frac{\Delta p}{T}\right) > \text{rand}\right] \text{and } \left[\exp\left(-\frac{\Delta p}{T}\right) < 1\right]\right\}$$

❖ The prior model

$$p(\mathbf{d}_1, \mathbf{d}_2 | \mathbf{g}_{k-1}) = \frac{1}{Q_d} \exp\{-U_d(\mathbf{d}_1, \mathbf{d}_2 | \mathbf{g}_{k-1})\}$$

Where Q_d is the partition function, and

$$U_d(\mathbf{d}_1, \mathbf{d}_2 | \mathbf{g}_{k-1}) = \sum_{c \in C_d} V_d^c(\mathbf{d}_1, \mathbf{d}_2 | \mathbf{g}_{k-1})$$

$$V_d^c\left(\mathbf{d}(\mathbf{x}_i), \mathbf{d}(\mathbf{x}_j)\right) = \left\|\mathbf{d}(\mathbf{x}_i) - \mathbf{d}(\mathbf{x}_j)\right\|^2$$

❖ The likelihood model

$$p(\mathbf{g}_k|\mathbf{d}_1,\mathbf{d}_2,\mathbf{g}_{k-1}) = (2\pi\sigma^2)^{-\frac{1}{2N}}\exp\left\{-\sum_{\mathbf{x}\in\Lambda}\frac{\left[\mathbf{g}_k(\mathbf{x})-\mathbf{g}_{k-1}\big(\mathbf{x}-\mathbf{d}(\mathbf{x})\big)\right]^2}{2\sigma^2}\right\}$$

$$= \exp\{-U_g(\mathbf{g}_k|\mathbf{d}_1,\mathbf{d}_2,\mathbf{g}_{k-1})\}$$

Where:

$$U_g(\mathbf{g}_k|\mathbf{d}_1,\mathbf{d}_2,\mathbf{g}_{k-1}) = \frac{\log(2\pi\sigma^2)}{2N}+\frac{1}{2\sigma^2}\sum_{\mathbf{x}\in\Lambda}\left[\mathbf{g}_k(\mathbf{x})-\mathbf{g}_{k-1}\big(\mathbf{x}-\mathbf{d}(\mathbf{x})\big)\right]^2$$

$$\Rightarrow U_g(\mathbf{g}_k|\mathbf{d}_1,\mathbf{d}_2,\mathbf{g}_{k-1})\sim\frac{1}{2\sigma^2}\sum_{\mathbf{x}\in\Lambda}\left[\mathbf{g}_k(\mathbf{x})-\mathbf{g}_{k-1}\big(\mathbf{x}-\mathbf{d}(\mathbf{x})\big)\right]^2$$

The optimization problem:

$$\min_{\mathbf{d}_1,\mathbf{d}_2} U_g(\mathbf{g}_k|\mathbf{d}_1,\mathbf{d}_2,\mathbf{g}_{k-1})+\lambda_d U_d(\mathbf{d}_1,\mathbf{d}_2|\mathbf{g}_{k-1})$$

Where $\lambda_d = 20$

(b) MAP with occlusion field

❖ The likelihood model

$$p(\mathbf{g}_k|\mathbf{d}_1,\mathbf{d}_2,\mathbf{o},\mathbf{g}_{k-1}) = (2\pi\sigma^2)^{-\frac{1}{2N}}\exp\left\{-\sum_{\mathbf{x}\in\Lambda}\frac{[1-o(\mathbf{x})]\left[\mathbf{g}_k(\mathbf{x})-\mathbf{g}_{k-1}\big(\mathbf{x}-\mathbf{d}(\mathbf{x})\big)\right]^2}{2\sigma^2}\right\}$$

$$= \exp\{-U_g(\mathbf{g}_k|\mathbf{d}_1,\mathbf{d}_2,\mathbf{o},\mathbf{g}_{k-1})\}$$

Where:

$$U_g(\mathbf{g}_k|\mathbf{d}_1,\mathbf{d}_2,\mathbf{o},\mathbf{g}_{k-1}) = \frac{\log(2\pi\sigma^2)}{2N}+\frac{1}{2\sigma^2}\sum_{\mathbf{x}\in\Lambda}[1-o(\mathbf{x})]\left[\mathbf{g}_k(\mathbf{x})-\mathbf{g}_{k-1}\big(\mathbf{x}-\mathbf{d}(\mathbf{x})\big)\right]^2$$

$$\Rightarrow U_g(\mathbf{g}_k|\mathbf{d}_1,\mathbf{d}_2,\mathbf{o},\mathbf{g}_{k-1})\sim\frac{1}{2\sigma^2}\sum_{\mathbf{x}\in\Lambda}[1-o(\mathbf{x})]\left[\mathbf{g}_k(\mathbf{x})-\mathbf{g}_{k-1}\big(\mathbf{x}-\mathbf{d}(\mathbf{x})\big)\right]^2$$

❖ The occlusion field model

$$p(\mathbf{o}|\mathbf{g}_{k-1}) = \frac{1}{Q_o}\exp\left\{-\frac{1}{\beta_o}U_o(\mathbf{o}|\mathbf{g}_{k-1})\right\}$$

Where:

$$U_o(\mathbf{o}|\mathbf{g}_{k-1}) = \sum_{c\in C_o} V_o^c(\mathbf{o}|\mathbf{g}_{k-1})$$

$V_o^c\big(o(\mathbf{x})\big) = o(\mathbf{x})T_o$ where $T_o = 500$

The optimization problem:

$$\min_{\mathbf{d}_1,\mathbf{d}_2} U_g(\mathbf{g}_k|\mathbf{d}_1,\mathbf{d}_2,\mathbf{o},\mathbf{g}_{k-1}) + \lambda_d U_d(\mathbf{d}_1,\mathbf{d}_2|\mathbf{g}_{k-1}) + \lambda_o U_o(\mathbf{o}|\mathbf{g}_{k-1})$$

Where $\lambda_d = 20$ and $\lambda_o = 10$

(c) MAP with line field

❖ The motion field

$$p(\mathbf{d}_1,\mathbf{d}_2|\mathbf{l},\mathbf{g}_{k-1}) = \frac{1}{Q_d}\exp\left\{-\frac{1}{\beta_d}U_d(\mathbf{d}_1,\mathbf{d}_2|\mathbf{l},\mathbf{g}_{k-1})\right\}$$

Where:

$$U_d(\mathbf{d}_1,\mathbf{d}_2|\mathbf{l},\mathbf{g}_{k-1}) = \sum_{c\in C_d} V_d^c(\mathbf{d}_1,\mathbf{d}_2|\mathbf{l},\mathbf{g}_{k-1})$$

$$V_d^c\big(\mathbf{d}(\mathbf{x}_i),\mathbf{d}(\mathbf{x}_j)\big|\mathbf{l}\big) = \left\|\mathbf{d}(\mathbf{x}_i) - \mathbf{d}(\mathbf{x}_j)\right\|^2\left(1 - l(\mathbf{x}_i,\mathbf{x}_j)\right)$$

❖ The line field model

$$p(\mathbf{l}|\mathbf{g}_{k-1}) = \frac{1}{Q_l}\exp\left\{-\frac{1}{\beta_l}U_l(\mathbf{l}|\mathbf{g}_{k-1})\right\}$$

Where:

$$U_l(\mathbf{l}|\mathbf{g}_{k-1}) = \sum_{c\in C_l} V_l^c(\mathbf{l}|\mathbf{g}_{k-1})$$

$$V_l^c(\mathbf{l}|\mathbf{g}_{k-1}) = V_l^{c_1}(\mathbf{l}|\mathbf{g}_{k-1}) + V_l^{c_4}(\mathbf{l}|\mathbf{g}_{k-1})$$

The optimization problem:

$$\min_{\mathbf{d}_1,\mathbf{d}_2} U_g(\mathbf{g}_k|\mathbf{d}_1,\mathbf{d}_2,\mathbf{g}_{k-1}) + \lambda_d U_d(\mathbf{d}_1,\mathbf{d}_2|\mathbf{g}_{k-1}) + \lambda_l U_l(\mathbf{l}|\mathbf{g}_{k-1}) \quad (\mathrm{c.}\,1)$$

Where $\lambda_l = 100$

(d) MAP with occlusion and line fields

The optimization problem:

$$\min_{\mathbf{d}_1,\mathbf{d}_2} U_g(\mathbf{g}_k|\mathbf{d}_1,\mathbf{d}_2,\mathbf{g}_{k-1}) + \lambda_d U_d(\mathbf{d}_1,\mathbf{d}_2|\mathbf{g}_{k-1}) + \lambda_l U_l(\mathbf{l}|\mathbf{g}_{k-1}) + \lambda_o U_o(\mathbf{o}|\mathbf{g}_{k-1}) \quad (\mathrm{d.}\,1)$$

Where:

$$U_g(\mathbf{g}_k|\mathbf{d}_1,\mathbf{d}_2,\mathbf{o},\mathbf{g}_{k-1}) \sim \frac{1}{2\sigma^2}\sum_{\mathbf{x}\in\Lambda}[1 - o(\mathbf{x})]\big[\mathbf{g}_k(\mathbf{x}) - \mathbf{g}_{k-1}\big(\mathbf{x} - \mathbf{d}(\mathbf{x})\big)\big]^2$$

$$U_d(\mathbf{d}_1, \mathbf{d}_2 | \mathbf{l}, \mathbf{g}_{k-1}) = \sum_i \sum_j \left\| \mathbf{d}(\mathbf{x}_i) - \mathbf{d}(\mathbf{x}_j) \right\|^2 \left(1 - l(\mathbf{x}_i, \mathbf{x}_j)\right)$$

The experimental results

The number of iteration is 16, $\sigma = 3.3$ and the 10$^{\text{th}}$, 11$^{\text{th}}$ frame are chosen. Figure 1 shows the reconstructed images of each method and compare with the 11$^{\text{th}}$ frame. Figure 2 shows PSNR of each method. Figure 3 displays the difference between the reconstructed images and the 11$^{\text{th}}$ frame.

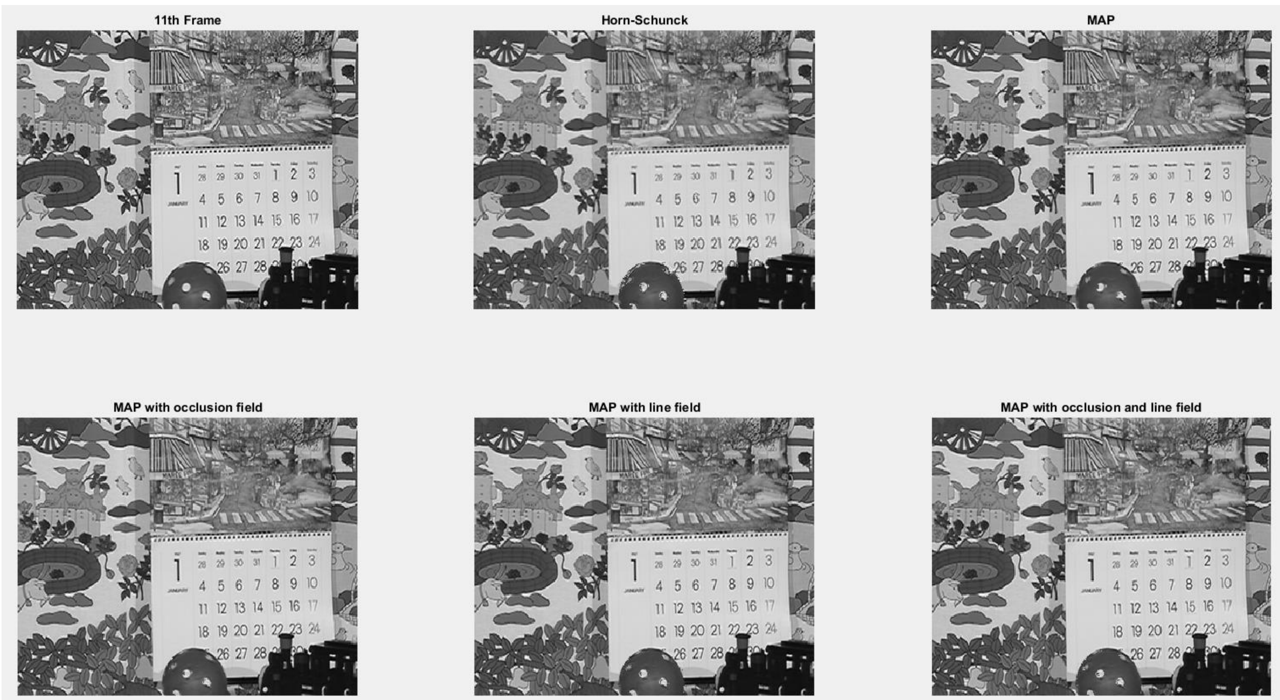

Figure 1. The reconstruction frames

```
Command Window
Horn-Schunck algorithm: PSNR = 27.2546 dB
MAP without occlusion and line field: PSNR = 32.4780 dB
MAP with occlusion field: PSNR = 32.5802 dB
MAP with line field: PSNR = 32.7694 dB
MAP with occlusion and line field: PSNR = 32.6723 dB
```

Figure 2. The PSNR of each method

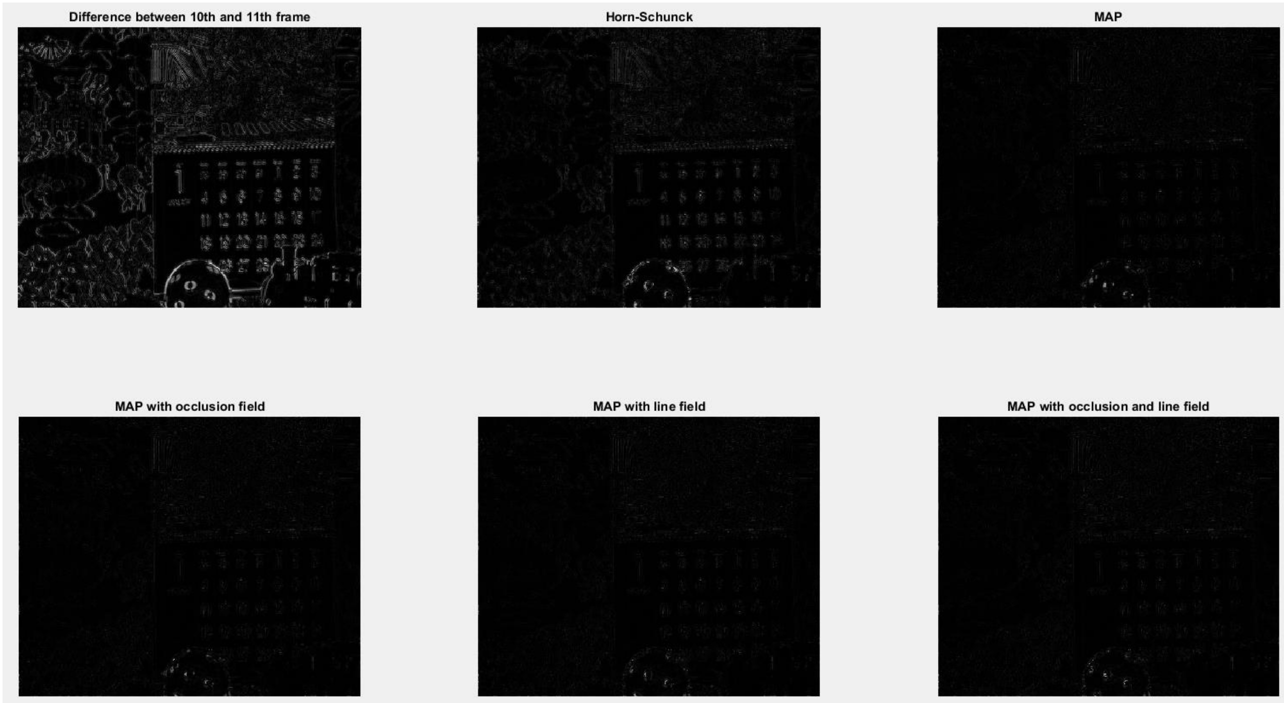

Figure 3. The difference between the reconstruction images and 10th frame

The MAP with or without occlusion and line field look sharped than Horn-Schunck algorithm, for examples, the boundary of the ball and the numbers in the calendar. The MAP without occlusion and line field, comparing with three other MAP, has the higher difference with 11th frame.

In PSNR, strangely, the PSNR of MAP with occlusion and line field is lightly lower than MAP with line field. The reason may be that (d.1) has more optimization variables than (c.1), in other work, (d.1) is more nonconvex than (c.1). Therefore, with the same number of iterations, it is more difficult to optimize (d.1) to global optimum point than (c.1).

(e) Comment on the prior probabilities

The prior probabilities for motion field, occlusion field do not take penalty on line-field site, so it may be cause blurring of motion boundaries. Moreover, there are several hundreds of thousands of unknowns for reasonable size of image [2]. For example, in this programming assignment, the size of frame is 288×352, so there are 101376 motion vectors (202752 components), 101376 occlusion labels and 202752 line-field labels for a total of 506880 unknowns.

Reference

[1] Stuart Geman and Donald Geman, "Stochastic Relaxtion, Gibbs Distributions, and the Bayesian Restoration of Images", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Volume PAMI-6, Issue 6, November 1984

[2] A. Murat Tekalp, "Digital Video Processing", Prentice Hall PTR, 1995