Conditional Velocity Score Estimation for Image Restoration Supplementary Material

Ziqiang Shi, Rujie Liu
Fujitsu R&D Center
Beijing, China
shiziqiang@fujitsu.com

1. Observation Velocity Score Approximation

Let $\mathbf{u}_0 = (\mathbf{x}_0, \mathbf{v}_0)^\top$, $\mathbf{u}_t = (\mathbf{x}_t, \mathbf{v}_t)^\top$, they are linked by the probabilistic transition kernel of linear SDEs [2]. In terms of CLD as follows

$$\begin{cases} d\mathbf{x}_t &= M^{-1}\beta \mathbf{v}_t dt, \\ d\mathbf{v}_t &= -\beta \mathbf{x}_t dt + \Gamma\beta M^{-1} \mathrm{v}_t dt + \sqrt{2\Gamma\beta} d\mathbf{w}_t. \end{cases} \tag{1}$$

we have the exact relation [1]

$$p(\mathbf{u}_t|\mathbf{u}_0) = \mathcal{N}(\mathbf{u}_t; \boldsymbol{\mu}_t, \boldsymbol{\Sigma}_t), \tag{2}$$

where

$$\begin{aligned} \boldsymbol{\mu}_t &= \begin{pmatrix} 2\beta t\Gamma^{-1}\mathbf{x}_0 + 4\beta t\Gamma^{-2}\mathbf{v}_0 + \mathbf{x}_0 \\ -\beta t \mathbf{x}_0 - 2\beta t\Gamma^{-1}\mathbf{v}_0 + \mathbf{v}_0 \end{pmatrix} e^{-2\beta t\Gamma^{-1}} \\ &= e^{-2\beta t\Gamma^{-1}} \begin{pmatrix} 2\beta t\Gamma^{-1} + 1 & 4\beta t\Gamma^{-2} \\ -\beta t & -2\beta t\Gamma^{-1} + 1 \end{pmatrix} \begin{pmatrix} \mathbf{x}_0 \\ \mathbf{v}_0 \end{pmatrix} \\ &= \boldsymbol{D}_t \begin{pmatrix} \mathbf{x}_0 \\ \mathbf{v}_0 \end{pmatrix}, \end{aligned} \tag{3}$$

and

$$\begin{aligned} \boldsymbol{\Sigma}_t &= \Sigma_t \otimes \boldsymbol{I}_d, \\ \Sigma_t &= \begin{pmatrix} \Sigma_t^{xx} & \Sigma_t^{xv} \\ \Sigma_t^{xv} & \Sigma_t^{vv} \end{pmatrix} e^{-4\beta t\Gamma^{-1}}, \\ \Sigma_t^{xx} &= \Sigma_0^{xx} + e^{4\beta t\Gamma^{-1}} - 1 + 4\beta t\Gamma^{-1}\left(\Sigma_0^{xx} - 1\right) \\ &\quad + 4\beta^2 t^2 \Gamma^{-2}\left(\Sigma_0^{xx} - 2\right) + 16\beta^2 t^2\Gamma^{-4}\Sigma_0^{vv}, \\ \Sigma_t^{xv} &= -\beta^2 t \Sigma_0^{xx} + 4\beta t\Gamma^{-2}\Sigma_0^{vv} \\ &\quad - 2\beta^2 t^2\Gamma^{-1}\left(\Sigma_0^{xx} - 2\right) - 8\beta^2 t^2 \Gamma^{-3}\Sigma_0^{vv}, \\ \Sigma_t^{vv} &= \tfrac{\Gamma^2}{4}\left(e^{4\beta t\Gamma^{-1}} - 1\right) + \beta t\Gamma + \beta^2 t^2\left(\Sigma_0^{xx} - 2\right) \\ &\quad + \Sigma_0^{vv}\left(1 + 4\beta t^2\Gamma^{-2} - 4\beta t\Gamma^{-1}\right). \end{aligned} \tag{4}$$

Then we can determine the probability of observing $\mathbf{y}$ given $\mathbf{u}_t$ by the following proposition.

Proposition 1.1. *(Determine the probability of observing $\mathbf{y}$ given $\mathbf{u}_t$) For simplicity of notation, $\mathbb{E}_{\mathbf{u}_0 \sim p(\mathbf{u}_0|\mathbf{u}_t)}$ is denoted as $\mathbb{E}$. Under the condition of $\mathbf{u}_t$ at time t, the probability of observing $\mathbf{y}$ can be derived from*

$$\mathbb{E}\left[p(\mathbf{y}|\mathbf{u}_0)\right] = p(\mathbf{y}|\hat{\mathbf{u}}_0) + c\mathbb{E}\left[\left\|\mathcal{H}(\mathbf{x}_0 - \hat{\mathbf{x}}_0)\right\|^2 \left\|\mathbf{x}_0 - \hat{\mathbf{x}}_0\right\|^2\right], \tag{5}$$

where

$$c = \frac{1}{\sigma^2}\int_0^1 \frac{1}{2\pi^{d/2}\sigma^{d/2}} \exp\left(-\frac{s^2}{2}\right)(1-s)\left(\frac{s^2}{\sigma^2} - 1\right) ds. \tag{6}$$

Proof. Assuming $\mathbf{u}_0$ is fixed, known from the image restoration problem formulation, $h(s) = p(\mathbf{y}|s\mathbf{u}_0 + (1 - s)\hat{\mathbf{u}}_0)$ is a quadratically differentiable function of s. So from the second-order Taylor formula, expand $h(s)$ at 0, we have

$$h(1) = h(0) + h'(0) + \int_0^1 h''(s)(1-s)ds, \tag{7}$$

which is

$$\begin{aligned} & p(\mathbf{y}|\mathbf{u}_0) & (8) \\ &= p(\mathbf{y}|\hat{\mathbf{u}}_0) + p'(\mathbf{y}|s\mathbf{u}_0 + (1-s)\hat{\mathbf{u}}_0)(\mathbf{u}_0 - \hat{\mathbf{u}}_0)|_{s=0} & (9) \\ &+ \int_0^1 (\mathbf{u}_0 - \hat{\mathbf{u}}_0)^\top \frac{d^2 p(\mathbf{y}|s\mathbf{u}_0 + (1-s)\hat{\mathbf{u}}_0)}{ds^2} & (10) \\ &\times (\mathbf{u}_0 - \hat{\mathbf{u}}_0)(1-s)ds. & (11) \end{aligned}$$

Taking expectations for $\mathbf{u}_0$ on both sides, we have

$$\begin{aligned} & \mathbb{E}\left[p(\mathbf{y}|\mathbf{u}_0)\right] & (12) \\ &= p(\mathbf{y}|\hat{\mathbf{u}}_0) + \mathbb{E}[\int_0^1 (\mathbf{u}_0 - \hat{\mathbf{u}}_0)^\top & (13) \\ &\times \frac{d^2 p(\mathbf{y}|s\mathbf{u}_0 + (1-s)\hat{\mathbf{u}}_0)}{ds^2}(\mathbf{u}_0 - \hat{\mathbf{u}}_0)(1-s)ds]. & (14) \end{aligned}$$

The calculation of $\frac{d^2 p(\mathbf{y}|s\mathbf{u}_0+(1-s)\hat{\mathbf{u}}_0)}{ds^2}$ is very critical. Assuming $\mathcal{H}$ is linear, it is not hard to obtain

$$\frac{dp(\mathbf{y}|s\mathbf{u}_0+(1-s)\hat{\mathbf{u}}_0)}{ds} \tag{15}$$

$$= p(\mathbf{y}|s\mathbf{u}_0+(1-s)\hat{\mathbf{u}}_0) \tag{16}$$

$$\times \frac{1}{\sigma^2}\left[\mathbf{y}-\mathcal{H}(s\mathbf{x}_0+(1-s)\hat{\mathbf{x}}_0)\right]\mathcal{H}(\mathbf{x}_0-\hat{\mathbf{x}}_0), \tag{17}$$

and

$$\frac{d^2 p(\mathbf{y}|s\mathbf{u}_0+(1-s)\hat{\mathbf{u}}_0)}{ds^2} \tag{18}$$

$$= p(\mathbf{y}|s\mathbf{u}_0+(1-s)\hat{\mathbf{u}}_0)\frac{\|\mathcal{H}(\mathbf{x}_0-\hat{\mathbf{x}}_0)\|^2}{\sigma^2} \tag{19}$$

$$\times\left(\frac{\|\mathbf{y}-\mathcal{H}(s\mathbf{x}_0+(1-s)\hat{\mathbf{x}}_0)\|^2}{\sigma^2}-1\right). \tag{20}$$

So

$$\mathbb{E}\left[p(\mathbf{y}|\mathbf{u}_0)\right] \tag{21}$$

$$= p(\mathbf{y}|\hat{\mathbf{u}}_0)+\mathbb{E}[\frac{\|\mathcal{H}(\mathbf{x}_0-\hat{\mathbf{x}}_0)\|^2\|\mathbf{x}_0-\hat{\mathbf{x}}_0\|^2}{\sigma^2} \tag{22}$$

$$\times\int_0^1 p(s)(1-s)\left(\frac{\|\mathbf{y}-\mathcal{H}(s\mathbf{x}_0+(1-s)\hat{\mathbf{x}}_0)\|^2}{\sigma^2}-1\right)ds]. \tag{23}$$

Thus conclusion of Eq. (5) is proved.

□

2. Posterior Estimation of Ground Truth

Proposition 2.1. *(Estimate the mean of the initial image and velocity from the current moment) Under the condition of $\mathbf{u}_t$ at time t, the posterior mean value of the image at time 0 can be derived from*

$$\left[\mathbf{\Sigma}_t^{-1}\boldsymbol{D}_t\mathbb{E}[\mathbf{u}_0|\mathbf{u}_t]\right]_\Delta = \nabla_{\mathbf{v}_t}\log p(\mathbf{u}_t)+\left[\mathbf{\Sigma}_t^{-1}\mathbf{u}_t\right]_\Delta, \tag{24}$$

where $\Delta \doteq d+1, d+2, \cdots 2d$ indicates to take the second half of this vector.

Proof. First from the transition densities Eq. (2) we have

$$\begin{aligned}
&p(\mathbf{u}_t|\mathbf{u}_0)\\
&=\mathcal{N}(\mathbf{u}_t;\boldsymbol{\mu}_t,\mathbf{\Sigma}_t)\\
&=\frac{1}{(2\pi)^{d/2}|\mathbf{\Sigma}_t|^{1/2}}\exp\left[-\frac{1}{2}(\mathbf{u}_t-\boldsymbol{\mu}_t)^\top\mathbf{\Sigma}_t^{-1}(\mathbf{u}_t-\boldsymbol{\mu}_t)\right]\\
&=\frac{1}{(2\pi)^{d/2}|\mathbf{\Sigma}_t|^{1/2}}\exp\left[-\frac{1}{2}(\mathbf{u}_t-\boldsymbol{D}_t\mathbf{u}_0)^\top\mathbf{\Sigma}_t^{-1}(\mathbf{u}_t-\boldsymbol{D}_t\mathbf{u}_0)\right]\\
&=\frac{1}{(2\pi)^{d/2}|\mathbf{\Sigma}_t|^{1/2}}\exp\left[-\frac{1}{2}\mathbf{u}_t^\top\mathbf{\Sigma}_t^{-1}\mathbf{u}_t\right]\\
&\quad\times\exp\left[\mathbf{u}_0^\top\boldsymbol{D}_t^\top\mathbf{\Sigma}_t^{-1}\mathbf{u}_t-\frac{1}{2}\mathbf{u}_0^T\boldsymbol{D}_t^\top\mathbf{\Sigma}_t^{-1}\boldsymbol{D}_t\mathbf{u}_0\right].
\end{aligned}$$

To simplify notation, let $p_0(\mathbf{u}_t) = \frac{1}{(2\pi)^{d/2}|\mathbf{\Sigma}_t|^{1/2}}\exp\left[-\frac{1}{2}\mathbf{u}_t^\top\mathbf{\Sigma}_t^{-1}\mathbf{u}_t\right]$, $T(\mathbf{u}_t) = \boldsymbol{D}_t^\top\mathbf{\Sigma}_t^{-1}\mathbf{u}_t$, and $\varpi(\mathbf{u}_0)=\frac{1}{2}\mathbf{u}_0^T\boldsymbol{D}_t^\top\mathbf{\Sigma}_t^{-1}\boldsymbol{D}_t\mathbf{u}_0$, then we have

$$p(\mathbf{u}_t|\mathbf{u}_0)=p_0(\mathbf{u}_t)\exp\left[\mathbf{u}_0^\top T(\mathbf{u}_t)-\varpi(\mathbf{u}_0)\right].$$

Integrate over initial data and velocity pair distribution we obtain

$$\begin{aligned}
p(\mathbf{u}_t)&=\int p(\mathbf{u}_t|\mathbf{u}_0)p(\mathbf{u}_0)d\mathbf{u}_0\\
&=p_0(\mathbf{u}_t)\int\exp\left[\mathbf{u}_0^\top T(\mathbf{u}_t)-\varpi(\mathbf{u}_0)\right]p(\mathbf{u}_0)d\mathbf{u}_0.
\end{aligned}$$

In order to get the score, then take the derivative of $\mathbf{v}_t$ on both sides of the previous formula

$$\begin{aligned}
&\nabla_{\mathbf{v}_t}p(\mathbf{u}_t)\\
&=\nabla_{\mathbf{v}_t}p_0(\mathbf{u}_t)\int\exp\left[\mathbf{u}_0^\top T(\mathbf{u}_t)-\varpi(\mathbf{u}_0)\right]p(\mathbf{u}_0)d\mathbf{u}_0\\
&\quad+p_0(\mathbf{u}_t)\int\nabla_{\mathbf{v}_t}\left[\exp\left[\mathbf{u}_0^\top T(\mathbf{u}_t)-\varpi(\mathbf{u}_0)\right]\right]p(\mathbf{u}_0)d\mathbf{u}_0\\
&=\frac{\nabla_{\mathbf{v}_t}p_0(\mathbf{u}_t)}{p_0(\mathbf{u}_t)}p(\mathbf{u}_t)\\
&\quad+\left[p_0(\mathbf{u}_t)\int\left(\nabla_{\mathbf{u}_t}\left[T(\mathbf{u}_t)\right]\right)^\top\mathbf{u}_0\exp\left[\mathbf{u}_0^\top T(\mathbf{u}_t)-\varpi(\mathbf{u}_0)\right]p(\mathbf{u}_0)d\mathbf{u}_0\right]_\Delta\\
&=\frac{\nabla_{\mathbf{v}_t}p_0(\mathbf{u}_t)}{p_0(\mathbf{u}_t)}p(\mathbf{u}_t)\\
&\quad+\left[\left(\nabla_{\mathbf{u}_t}\left[T(\mathbf{u}_t)\right]\right)^\top\int p_0(\mathbf{u}_t)\mathbf{u}_0\exp\left[\mathbf{u}_0^\top T(\mathbf{u}_t)-\varpi(\mathbf{u}_0)\right]p(\mathbf{u}_0)d\mathbf{u}_0\right]_\Delta\\
&=\frac{\nabla_{\mathbf{v}_t}p_0(\mathbf{u}_t)}{p_0(\mathbf{u}_t)}p(\mathbf{u}_t)\\
&\quad+\left[\left(\nabla_{\mathbf{u}_t}\left[T(\mathbf{u}_t)\right]\right)^\top\int\mathbf{u}_0 p(\mathbf{u}_t|\mathbf{u}_0)p(\mathbf{u}_0)d\mathbf{u}_0\right]_\Delta.
\end{aligned}$$

Divide both sides by $p(\mathbf{u}_t)$, that is to say

$$\begin{aligned}
&\frac{\nabla_{\mathbf{v}_t}p(\mathbf{u}_t)}{p(\mathbf{u}_t)}\\
&=\frac{\nabla_{\mathbf{v}_t}p_0(\mathbf{u}_t)}{p_0(\mathbf{u}_t)}+\left[\left(\nabla_{\mathbf{u}_t}\left[T(\mathbf{u}_t)\right]\right)^\top\int\mathbf{u}_0\frac{p(\mathbf{u}_t|\mathbf{u}_0)p(\mathbf{u}_0)}{p(\mathbf{u}_t)}d\mathbf{u}_0\right]_\Delta.
\end{aligned}$$

Transforming the formula and simplifying the notation, we get

$$\begin{aligned}
&\left[\mathbf{\Sigma}_t^{-1}\boldsymbol{D}_t\mathbb{E}[\mathbf{u}_0|\mathbf{u}_t]\right]_\Delta\\
&=\left[\left[\left(\nabla_{\mathbf{u}_t}\left[T(\mathbf{u}_t)\right]\right)^\top\right]\mathbb{E}[\mathbf{u}_0|\mathbf{u}_t]\right]_\Delta\\
&=\nabla_{\mathbf{v}_t}\log p(\mathbf{u}_t)-\nabla_{\mathbf{v}_t}\log p_0(\mathbf{u}_t)\\
&=\nabla_{\mathbf{v}_t}\log p(\mathbf{u}_t)+\left[\mathbf{\Sigma}_t^{-1}\mathbf{u}_t\right]_\Delta.
\end{aligned}$$

The conclusion of Eq. (24) is proved.

□

References

[1] Tim Dockhorn, Arash Vahdat, and Karsten Kreis. Score-based generative modeling with critically-damped langevin diffusion. *arXiv preprint arXiv:2112.07068*, 2021. 1

[2] Simo Särkkä and Arno Solin. *Applied stochastic differential equations*, volume 10. Cambridge University Press, 2019. 1