实验报告

1 总体框架

LVOC 的总体框架如下图所示。虚线内为训练变分形式部分，给定几组

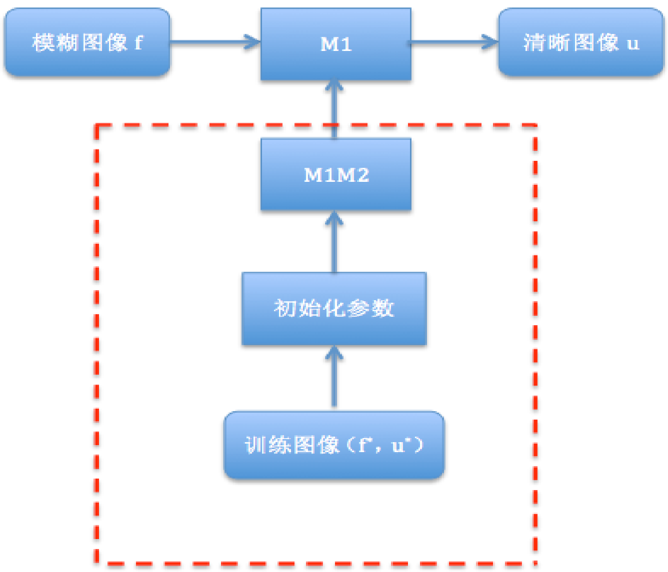


图 1: 框架

训练图像 $\hat{u}, \hat{f}$，其中 $\hat{u}$ 表示模糊图像，$\hat{f}$ 表示 Groudtruth，经过初始化参数后，进入 M_1M_2 进行训练系数，训练后的系数输入到 cgm 模型中，给定真实的模糊图像后，经过 cgm 模型处理后得到清晰图像，原 L-PDE 给出大致思路也是这样（经过训练后的系数用于解决一类问题如去模糊）。

对于某一类问题，训练得到的系数不一定都使用。例如对于同一图像的去模糊问题，系数可能和模糊程度有关，所以希望引入评价指标，对图像的模糊程度进行评价，进而选择同等级的系数。

2 CGM 模型 (M1)

这里对 CGM 模型做了一点修改。

在论文《A Primal-Dual Active-Set Method for Non-negeativity Constrained Total Variation Deblurring Problems》考虑原始问题：

$$\min_{u}\{\frac{1}{2}||Ku-f||^2+\beta|u|_{TV}\} \tag{1}$$

引入对偶变量 $p=\dfrac{\nabla u}{|\nabla u|}$ 后等价与解决下列方程：

$$\begin{aligned} -\beta divp - K^T f + K^T K u &= 0 \\ |\nabla u| p - \nabla u &= 0 \end{aligned} \tag{2}$$

应用牛顿法解上述方程：

$$\begin{bmatrix} -\beta div & K^T K \\ |\nabla u| & -(I-\frac{p(\nabla u)^T}{|\nabla u|})\nabla \end{bmatrix} \begin{bmatrix} \delta p \\ \delta u \end{bmatrix} = \begin{bmatrix} F_1(p,u) \\ F_2(p,u) \end{bmatrix} \tag{3}$$

与原有 CGM 模型不同的是我们假设对（3）迭代求解时的系数 β_k 是随迭代步数变化而变化的，这也与 L-PDE 设定系数 $a(t)$ 随时间变化而变化是一致的，而系数同样是通过模块 M_1M_2 训练而来。

3 CGMOV(M1M2)

3.1 模型

对 CGM 模型加目标函数:

$$\begin{array}{rll} & \min_\beta & \{\frac{1}{2}\sum_{k=1}^K\int_\Omega(u_k(t)-\hat{f})^2 d\Omega+\frac{1}{2}a_1\beta^2\} \\ s.t. & \min_{u_k} & \{\frac{1}{2}||Ku_k-f||^2+\beta|u_k|_{TV}\} \end{array} \tag{4}$$

记 $J(u,\beta)=\{\frac{1}{2}\sum_{k=1}^K\int_\Omega(u_k(t)-\hat{f})^2 d\Omega+\frac{1}{2}a_1\beta^2\}$，$E_1(u,\beta)=\{\frac{1}{2}||Ku_k-f||^2+\beta|u_k|_{TV}\}$，借助于 G-导数对 $\min_\beta J(u,\beta)$ 进行求解:

$$\frac{\partial J(u,\beta)}{\partial\beta}=a_1\beta+\sum_{k=1}^K\int_\Omega(u_k(t)-\hat{f})\frac{\partial u_k(t)}{\partial\beta}d\Omega=0 \tag{5}$$

对于 $E_1(u,\beta)$ 有:

$$\nabla_u E_1(u,\beta)=K^TKu-K^Tf-\beta divp=0 \tag{6}$$

两边同时对上式求 β 导数:

$$K^TK\frac{\partial u}{\partial\beta}-divp=0 \tag{7}$$

所以

$$\frac{\partial u_k(t)}{\partial\beta}=(K^TK)^{-1}divp \tag{8}$$

我们用（5）对 β 进行修正

$$\beta=-\frac{1}{a_1}\sum_{k=1}^K\int_\Omega(u_k(t)-\hat{f})\frac{\partial u_k(t)}{\partial\beta}d\Omega=-\frac{1}{b_1\int_\Omega f^2 d\Omega}\sum_{k=1}^K\int_\Omega(u_k(t)-\hat{f})\frac{\partial u_k(t)}{\partial\beta}d\Omega \tag{9}$$

b_1 应该较大些，保证每步搜索的精确性，$\int_\Omega f^2 d\Omega$ 反应了真实图像对 β 的影响。

3.2 停止准则

在这里使用两个停止准则：一是程序中给出的 KKT 误差；二是 $Index = \int_\Omega (u-f)^2 d\Omega$ 一直减小。第二个准则是在实验中给出的，如下图所示，$Index$ 随迭代步数先减小后增大，所以若 $Index$ 增加，则停止。

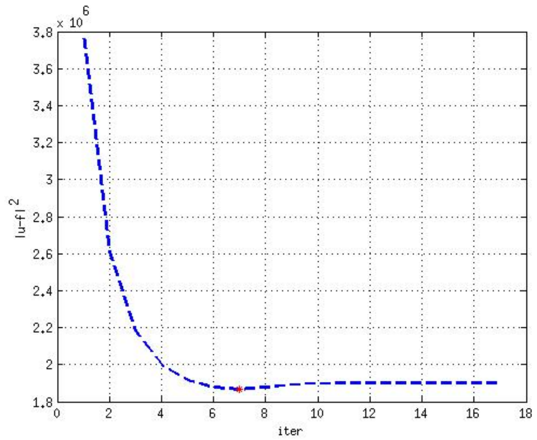


图 2: 迭代步数与 $|u-f|^2$ 的关系

4 实验结果

说明：(a)$\int_\Omega |u-f|^2 d\Omega$ 与 β_0 的关系；(b)CGM 模型目标函数与 β_0 的关系；(c) 运行时间与 β_0 的关系；(d) 总迭代步数与 β_0 的关系。

对比图 3 和图 4（或图 5 和图 6）可以看出，若引入第二个停止准则可以使得 $\int_\Omega |u-f|^2 d\Omega$、时间、迭代步数变小，但 CGM 模型的目标函数值却变大，由此可以看出 CGM 模型的目标函数值标识处理效果有局限性。

另外，系数 β_k 与两个 KKt 条件的误差似乎存在一定关系，这地方我还没想清楚，留在后面探究。

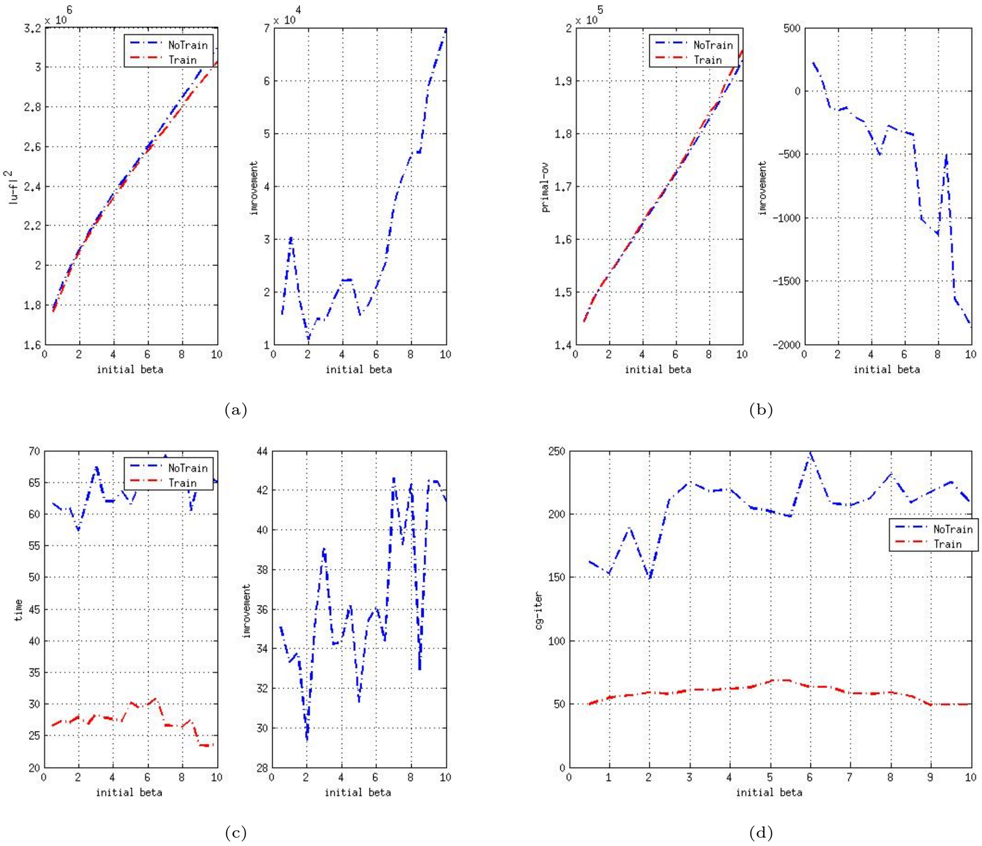


图 3: 两个停止准则，satellite128*128.tiff

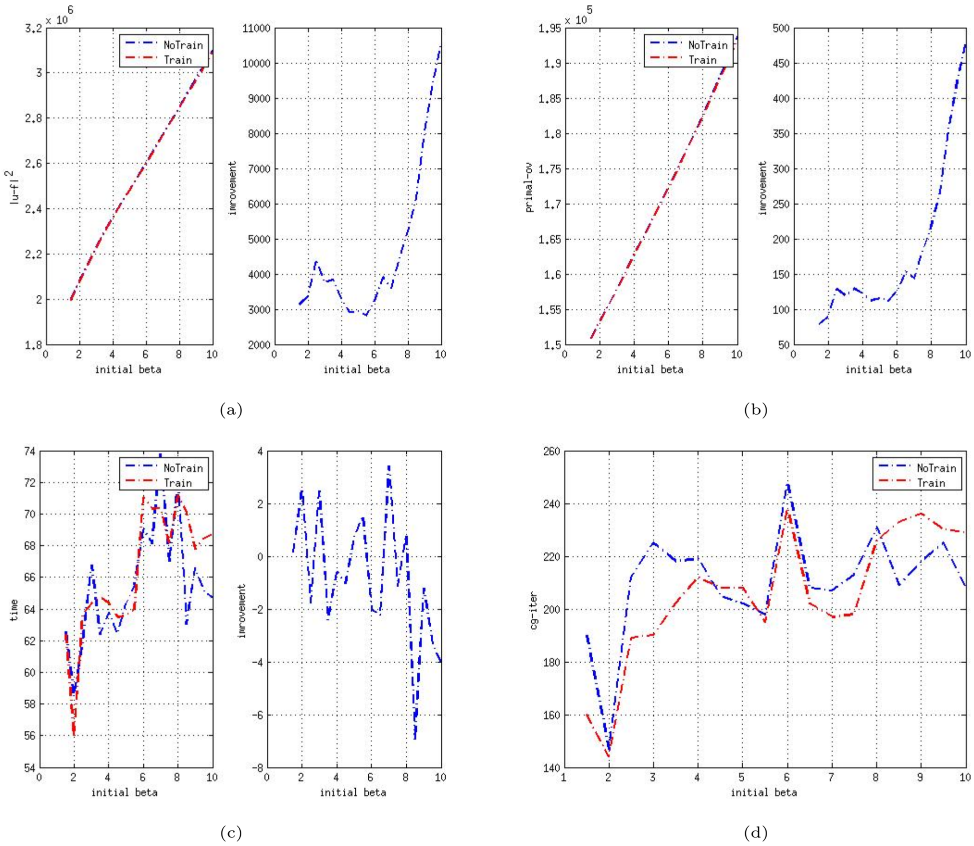


图 4: 停止准则一，satellite128*128.tiff

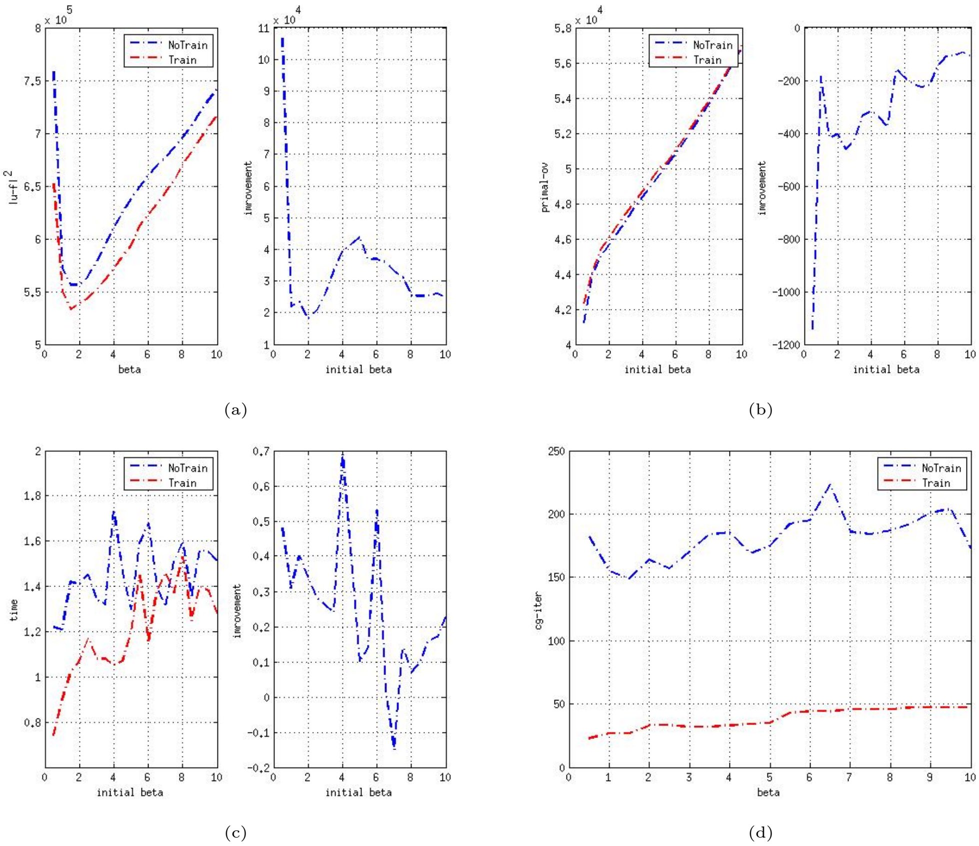
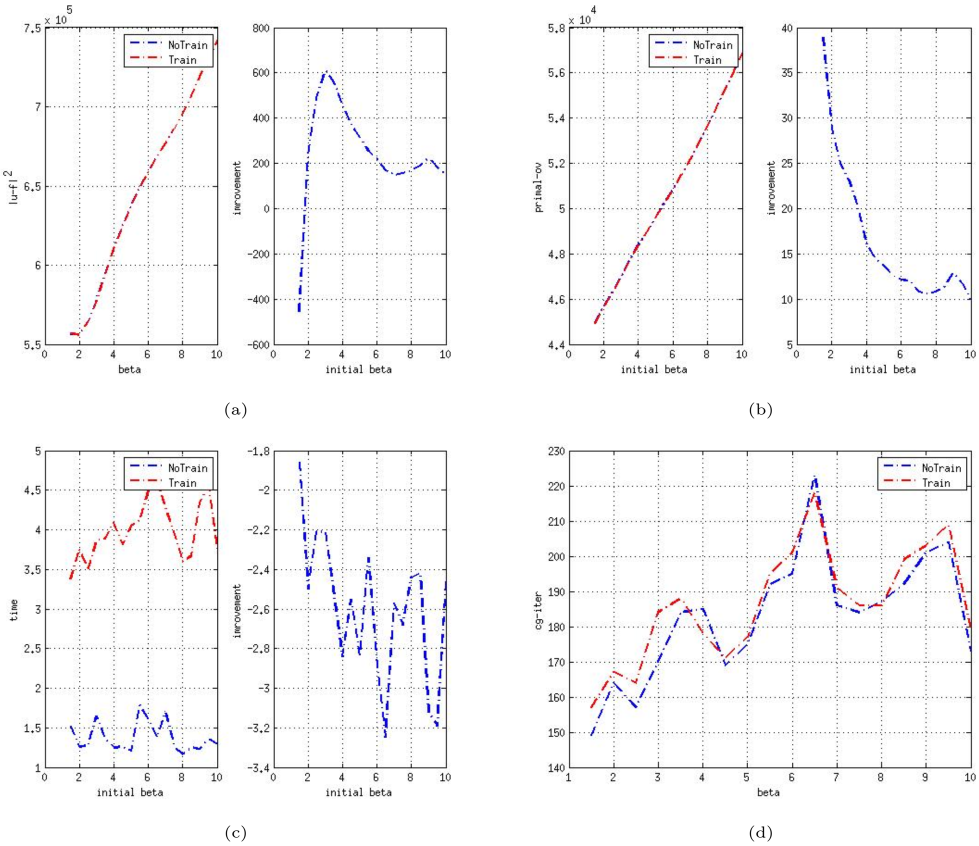


图 5: 两个停止准则，satellite32*32.tiff



(a) (b) (c) (d)

图 6: 停止准则一，satellite32*32.tiff

下一步：由于 KKt 条件达到较好时，J 不一定小，而 J 最小时，kkt 条件不小，所以寻找并证明一个条件下两者都达到最小。

5 修改 M1M2

在 CGM 模型 (1) 中可得到欧拉 -拉格朗日方程:

$$\nabla_u E = -K^T f + K^T K u - \beta div(\frac{\nabla u}{|\nabla u|_\epsilon}) = 0 \tag{10}$$

$$\begin{aligned} \delta\nabla_u E &= \nabla_u E(u+\tau\delta u) - \nabla_u E(u) \\ &= [-K^T f + K^T K(u+\tau\delta u) - \beta div(\tfrac{\nabla u+\tau\nabla\delta u}{|\nabla u+\tau\nabla\delta u|_\epsilon})] \\ &\quad -[-K^T f + K^T K u - \beta div(\tfrac{\nabla u}{|\nabla u|_\epsilon})] \\ &= \tau K^T K\delta u + \beta div(\tfrac{\nabla u}{|\nabla u|} - \tfrac{\nabla(u+\tau\delta u)}{|\nabla(u+\tau\delta u)|}) + o(\tau) \end{aligned} \tag{11}$$

于是,CGMOV 模型 (4) 可转化为:

$$\begin{aligned} \min_\beta \quad & \{\tfrac{1}{2}\textstyle\sum_{k=1}^K \int_\Omega (u_k(t) - \hat{f})^2 d\Omega + \tfrac{1}{2}a_1\beta^2\} \\ s.t. \quad & -K^T f + K^T K u - \beta div(\tfrac{\nabla u}{|\nabla u|_\epsilon}) = 0 \end{aligned} \tag{12}$$

考虑上述方程的拉格朗日方程:

$$\tilde{J} = J + < \phi, -K^T f + K^T K u - \beta div(\frac{\nabla u}{|\nabla u|_\epsilon}) > \tag{13}$$

$$\begin{aligned}
\delta\tilde{J} &= \tilde{J}(u+\tau\delta u)-\tilde{J}(u)\\
&= J(u+\tau\delta u)-J(u)+<\phi,\delta\nabla_u E>+o(\tau)\\
&= 2\tau\int_\Omega(u-f)\delta ud\Omega+<\phi,\delta\nabla_u E>+o(\tau)\\
&= 2\tau\int_\Omega(u-f)\delta ud\Omega+\tau\int_\Omega\phi K^TK\delta ud\Omega\\
&\quad +2\beta\tau\int_\Omega\nabla(\nabla\phi\tfrac{\nabla u\nabla u}{|\nabla u+\tau\nabla\delta u||\nabla u|(|\nabla u|+|\nabla u+\tau\nabla\delta u|)})\delta ud\Omega\\
&\quad -\beta\tau\int_\Omega\nabla(\tfrac{\nabla\phi}{|\nabla u+\tau\nabla\delta u|})\delta ud\Omega+o(\tau)
\end{aligned}\tag{14}$$

所以,

$$\frac{\partial\tilde{J}}{\partial u}=2(u-f)+\phi K^TK=0\tag{15}$$

$$\frac{\partial\tilde{J}}{\partial\beta}=a\beta-\int_\Omega\phi div(\frac{\nabla u}{|\nabla u|})=0\tag{16}$$

若引入对偶向量 p,CGMOV 模型转化为:

$$\begin{array}{ll}
\min_\beta & \{\frac{1}{2}\sum_{k=1}^K\int_\Omega(u_k(t)-\hat{f})^2d\Omega+\frac{1}{2}a_1\beta^2\}\\
s.t. & \begin{cases} |\nabla u|p-\nabla u=0\\ -K^Tf+K^TKu-\beta divp=0\end{cases}
\end{array}\tag{17}$$

上述问题对应的增广拉格朗日方程为:

$$\breve{J}=\int(u-f)^2d\Omega+\frac{1}{2}a\beta+\int\phi_1[|\nabla u|p-\nabla u]d\Omega+\int\phi_2[-K^Tf+K^TKu-\beta divp]d\Omega\tag{18}$$

故,

$$\frac{\partial\breve{J}}{\partial u}=2(u-f)-\nabla(\phi_1pp)+\nabla\phi_1+\phi_2K^TK=0\tag{19}$$

$$\frac{\partial\breve{J}}{\partial p}=\phi_1|\nabla u|+\beta\nabla\phi_2=0\tag{20}$$

$$\frac{\partial \breve{J}}{\partial \beta} = a\beta - \int \phi_2 divp \tag{21}$$

求解 ϕ_1,ϕ_2 所用的方程 (19)(20) 等价于:

$$\left\{\begin{array}{lcl} F_1 & = & 2(u-f) + (I-p^2)\nabla\phi_1 - 2p\nabla p\phi_1 + \phi_2 K^T K = 0 \\ F_2 & = & \phi_1|\nabla u| + \beta\nabla\phi_2 = 0 \end{array}\right. \tag{22}$$

故:

$$\left[\begin{array}{cc} p\nabla p + (I-p^2)\nabla & K^T K \\ |\nabla u| & \beta\nabla \end{array}\right]\left[\begin{array}{c} \delta\phi_1 \\ \delta\phi_2 \end{array}\right] = \left[\begin{array}{c} -F_1 \\ -F_2 \end{array}\right] \tag{23}$$

如果综合考虑两个优化问题, 我们利用欧拉 -拉格朗日方程对 u,p,ϕ,β 同时更新:

$$\left\{\begin{array}{rcl} |\nabla u|p - \nabla u & = & 0 \\ -K^T f + K^T K u - \beta divp & = & 0 \\ 2(u-f) + \phi K^T K & = & 0 \\ a\beta - \int_\Omega \phi divp & = & 0 \end{array}\right. \tag{24}$$

6 附录 1

$$\begin{array}{rcl} <\phi, \delta\nabla_u E> & = & \int \phi[\tau K^T K\delta u + \beta div(\frac{\nabla u}{|\nabla u|} - \frac{\nabla(u+\tau\delta u)}{|\nabla(u+\tau\delta u)|})d\Omega \\ & = & \tau\int\phi K^T K\delta u d\Omega + \beta\int\phi div(\frac{\nabla u}{|\nabla u|} - \frac{\nabla(u+\tau\delta u)}{|\nabla(u+\tau\delta u)|})\Omega \end{array} \tag{25}$$

$$\begin{array}{rcl} \int\phi div(\frac{\nabla u}{|\nabla u|} - \frac{\nabla(u+\tau\delta u)}{|\nabla(u+\tau\delta u)|})\Omega & = & \int_{\partial\Omega}\phi(\frac{\nabla u}{|\nabla u|} - \frac{\nabla(u+\tau\delta u)}{|\nabla(u+\tau\delta u)|})d\Omega \\ & & + \int\nabla\phi(\frac{\nabla(u+\tau\delta u)}{|\nabla(u+\tau\delta u)|} - \frac{\nabla u}{|\nabla u|})d\Omega \\ & = & \int\nabla\phi\nabla u(\frac{1}{|\nabla(u+\tau\delta u)|} - \frac{1}{|\nabla u|})d\Omega + \tau\int\nabla\phi\frac{\nabla\delta u}{|\nabla(u+\tau\delta u)|}d\Omega \end{array} \tag{26}$$

$$\begin{aligned}
\int \nabla\phi\nabla u(\tfrac{1}{|\nabla(u+\tau\delta u)|}-\tfrac{1}{|\nabla u|})d\Omega &= \int \nabla\phi\nabla u\tfrac{|\nabla u|-|\nabla(u+\tau\delta u)|}{|\nabla(u+\tau\delta u)||\nabla u|}d\Omega \\
&= \int \nabla\phi\nabla u\tfrac{(\nabla u)^2-(\nabla u+\tau\nabla\delta u)^2}{|\nabla(u+\tau\delta u)||\nabla u|(|\nabla(u+\tau\delta u)|+|\nabla u|)}d\Omega \\
&= \int \nabla\phi\nabla u\tfrac{-2\tau\nabla\delta u\nabla u}{|\nabla(u+\tau\delta u)||\nabla u|(|\nabla(u+\tau\delta u)|+|\nabla u|)}d\Omega \\
&= -2\tau\int \tfrac{\nabla\phi\nabla u\nabla u}{|\nabla(u+\tau\delta u)||\nabla u|(|\nabla(u+\tau\delta u)|+|\nabla u|)}\nabla\delta u d\Omega \\
&= 2\tau\int \nabla(\tfrac{\nabla\phi\nabla u\nabla u}{|\nabla(u+\tau\delta u)||\nabla u|(|\nabla(u+\tau\delta u)|+|\nabla u|)})\delta u d\Omega
\end{aligned} \tag{27}$$

$$\int \tfrac{\nabla u}{|\nabla(u+\tau\delta u)|}\nabla\delta u d\Omega = -\int \nabla(\tfrac{\nabla u}{|\nabla(u+\tau\delta u)|})\delta u d\Omega \tag{28}$$

$$\delta\breve{J} = \int (u+\tau\delta u-f)^2-(u-f)^2 d\Omega+\delta_2+\delta_3 \tag{29}$$

$$\begin{aligned}
\delta_2 &= \int \phi_1[|\nabla(u+\tau\delta u)|p-\nabla(u+\tau\delta u)]d\Omega-\int\phi_1[|\nabla u|p-\nabla u]d\Omega \\
&= \int \phi_1[|\nabla(u+\tau\delta u)|-|\nabla u|]pd\Omega-\tau\int\phi_1\nabla\delta u d\Omega \\
&= \int \phi_1\tfrac{(\nabla(u+\tau\delta u)^2-(\nabla u)^2}{|\nabla(u+\tau\delta u|+|\nabla u|}pd\Omega+\tau\int\nabla\phi_1\delta u d\Omega \\
&= \int \phi_1\tfrac{2\tau\nabla u\nabla\delta u}{|\nabla(u+\tau\delta u|+|\nabla u|}pd\Omega+\tau\int\nabla\phi_1\delta u d\Omega \\
&= -2\tau\int\nabla(\phi_1\tfrac{\nabla u}{|\nabla(u+\tau\delta u|+|\nabla u|}p)\delta u d\Omega+\tau\int\nabla\phi_1\delta u d\Omega
\end{aligned} \tag{30}$$

$$\begin{aligned}
\delta_3 &= \int \phi_2[-K^Tf+K^TK(u+\delta u)-\beta divp]d\Omega-\int\phi_2[-K^Tf+K^TKu-\beta divp]d\Omega \\
&= \tau\int\phi_2K^TK\delta u d\Omega
\end{aligned} \tag{31}$$

$$\begin{aligned}
-\beta\int\phi_2 div(p+\tau\delta p)d\Omega+\beta\int\phi_2 divpd\Omega &= -\tau\beta\int\phi_2 div\delta pd\Omega \\
&= \tau\beta\int\nabla\phi_2\delta pd\Omega
\end{aligned} \tag{32}$$

$$\delta_{\phi_1} F_1 = [-\nabla(\phi_1 + \tau\delta\phi_1)p^2 - 2(\phi_1 + \tau\delta\phi_1)p\nabla p + \nabla(\phi_1 + \tau\delta\phi_1)] - [-\nabla\phi_1 p^2 - 2\phi_1 p\nabla p + \nabla\phi_1]$$

$$(33)$$