

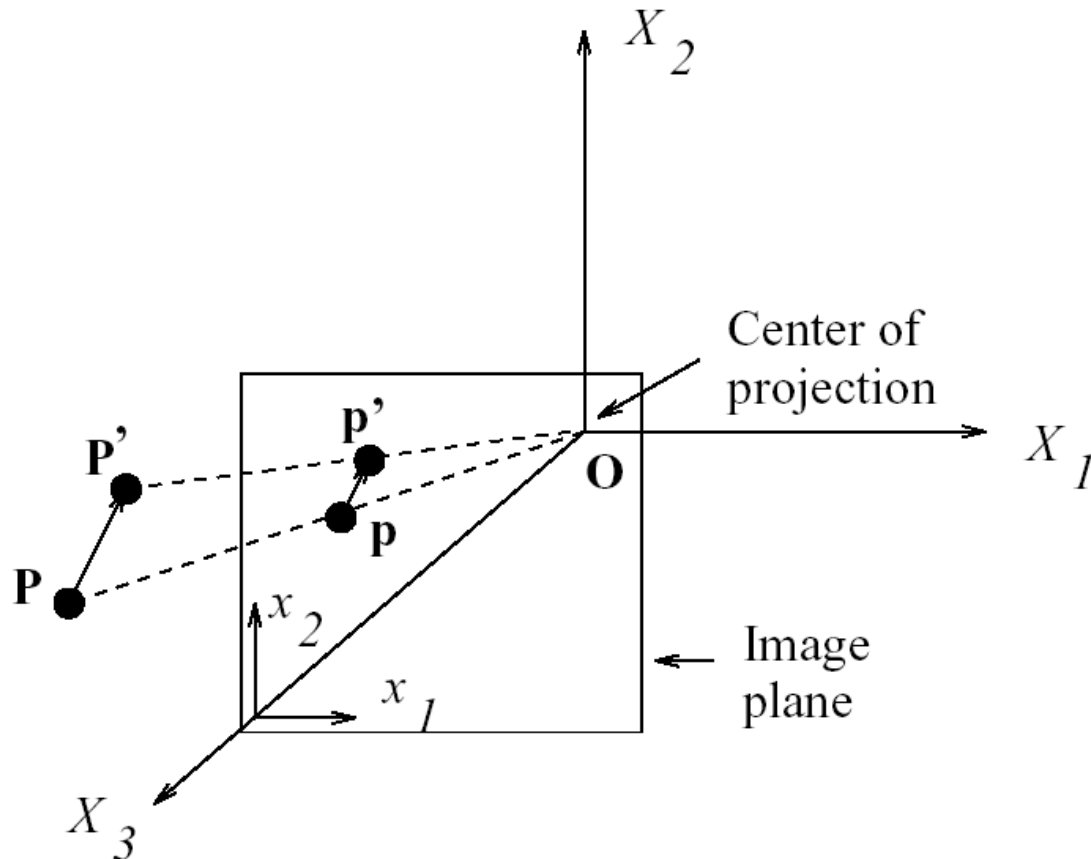
KECE471 Computer Vision

Motion

Chang-Su Kim

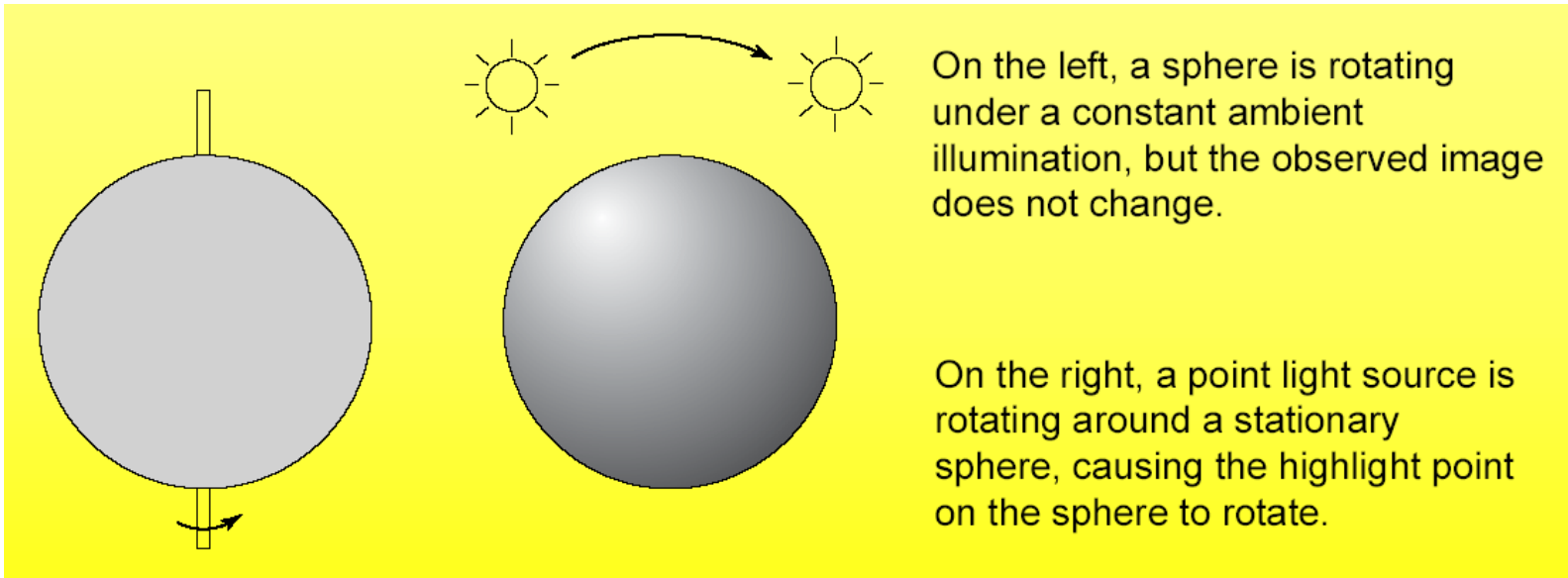
2D Motion vs. Optical Flow

- True 2D motion
 - There is 3D motion between object and camera
 - It is projected onto 2D imaging plane



2D Motion vs. Optical Flow

- Optical flow
 - **observed, perceived, apparent** 2D motion based on changes in pixel luminance
 - It also depends on illumination and object surface texture
 - It may **not** represent true 2D motion



Optical Flow Equation

- Given only video sequence without any other information (such as illumination condition), we cannot estimate true 2D motion.
- The best one can hope to estimate is optical flow
- Constant intensity assumption → optical flow equation

Under "constant intensity assumption":

$$\psi(x + d_x, y + d_y, t + d_t) = \psi(x, y, t)$$

But, using Taylor's expansion :

$$\psi(x + d_x, y + d_y, t + d_t) = \psi(x, y, t) + \frac{\partial \psi}{\partial x} d_x + \frac{\partial \psi}{\partial y} d_y + \frac{\partial \psi}{\partial t} d_t$$

Compare the above two, we have the equation that

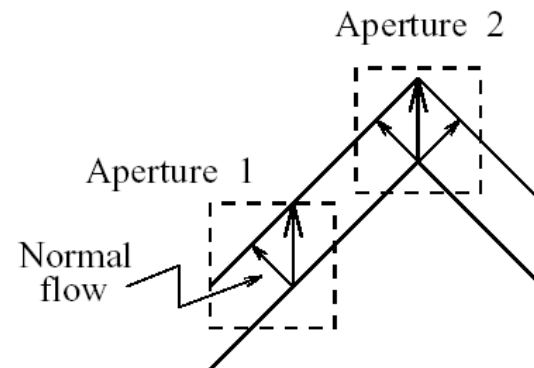
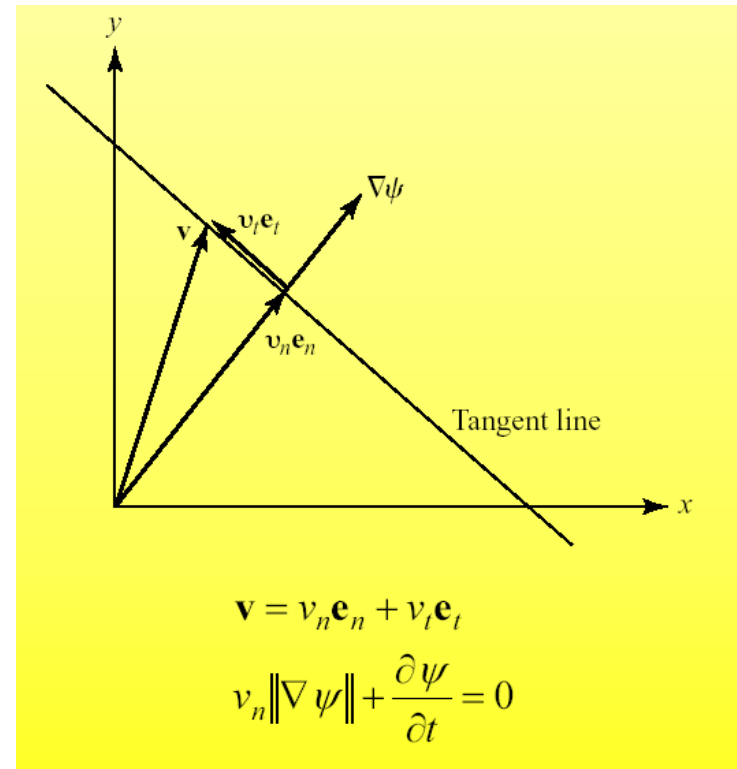
$$\frac{\partial \psi}{\partial x} d_x + \frac{\partial \psi}{\partial y} d_y + \frac{\partial \psi}{\partial t} d_t = 0 \quad \text{or} \quad \frac{\partial \psi}{\partial x} v_x + \frac{\partial \psi}{\partial y} v_y + \frac{\partial \psi}{\partial t} = 0 \quad \text{or} \quad \nabla \psi \cdot \mathbf{v} + \frac{\partial \psi}{\partial t} = 0$$

spatial
gradient

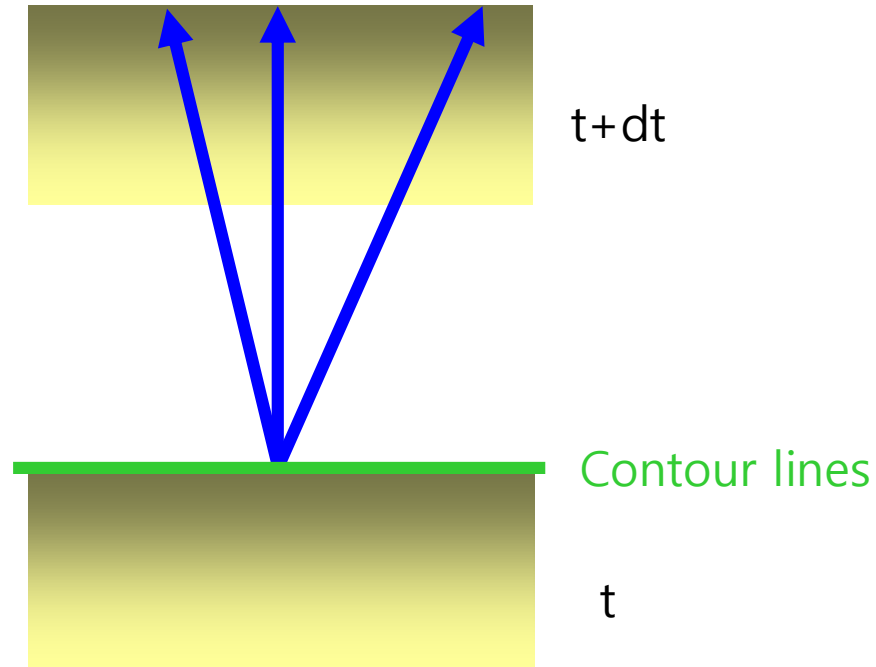
optical
flow

Ambiguities in Optical Flow Estimation

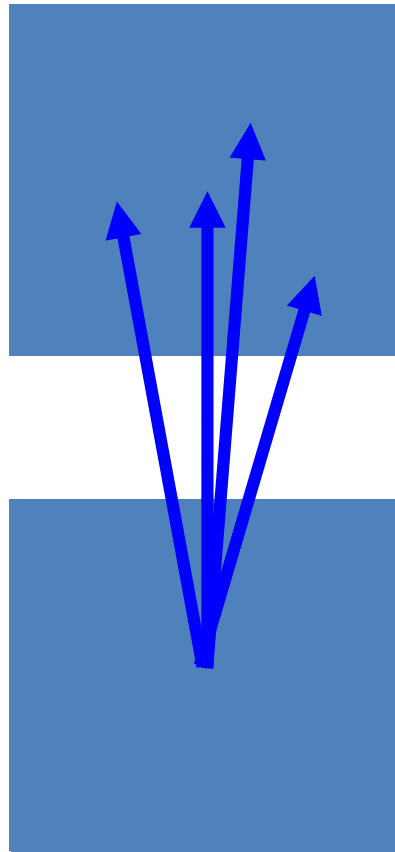
- Optical flow equation constrains the motion vector in the gradient direction v_n only
- The flow vector in the tangent direction v_t is under-determined
 - We can only determine the displacement that is orthogonal to the edges
- In regions with constant brightness $\nabla\psi = 0$, the flow is indeterminate
 - Optical flow estimation is unreliable in regions with flat texture and more reliable near edges



Ambiguities in Optical Flow Estimation



Ambiguities in Optical Flow Estimation



Example

- Consider a video signal $s(x, y, t)$ defined over the entire 3-D space (x, y, t) , where $s(x, y, t)$ is generated from one object, which undergoes translational motion with a uniform constant velocity (v_x, v_y) . Suppose that $s(x, y, 0) = x + y + xy$.
 1. Determine $s(x, y, t)$.
 2. Show that in this case the following optical flow equation holds true.

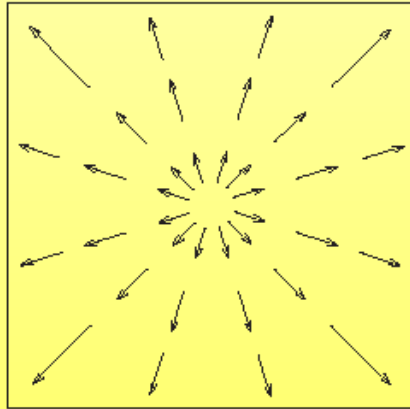
$$v_x \frac{\partial s(x, y, t)}{\partial x} + v_y \frac{\partial s(x, y, t)}{\partial y} + \frac{\partial s(x, y, t)}{\partial t} = 0$$

General Considerations for Motion Estimation

- Two categories of approaches
 - Feature based
 - Correspondence between edges, points, etc
 - Object tracking, 3D reconstruction from 2D
 - Intensity based
 - Optical flow estimation based on constant intensity assumption
 - Motion compensated prediction and filtering
 - Focus in this class
- Three important questions
 - How to **represent** the motion field?
 - Which **cost function** (criterion) to use to estimate motion parameters?
 - Which **optimization** technique?

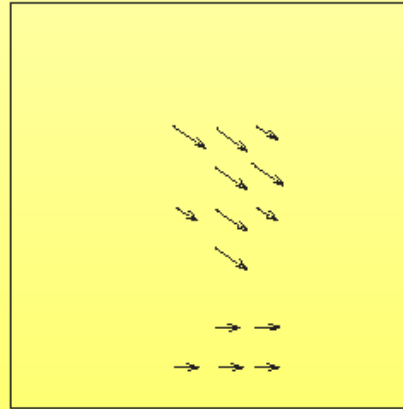
Motion Representation

Global:
Entire motion field is represented by a few global parameters



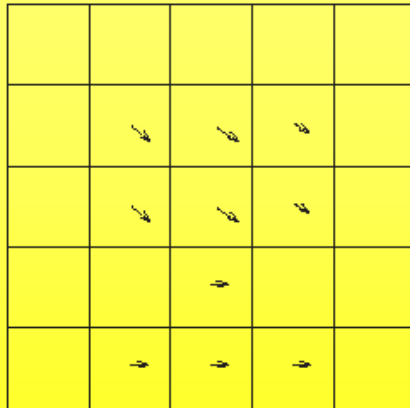
(a)

Pixel-based:
One MV at each pixel, with some smoothness constraint between adjacent MVs.



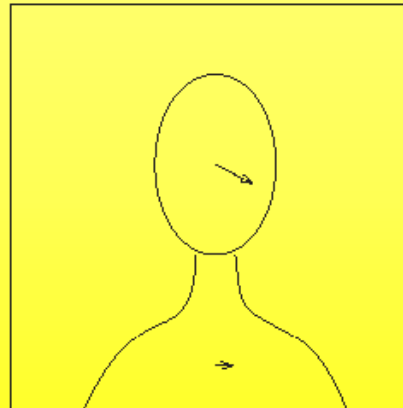
(b)

Block-based:
Entire frame is divided into blocks, and motion in each block is characterized by a few parameters.



(c)

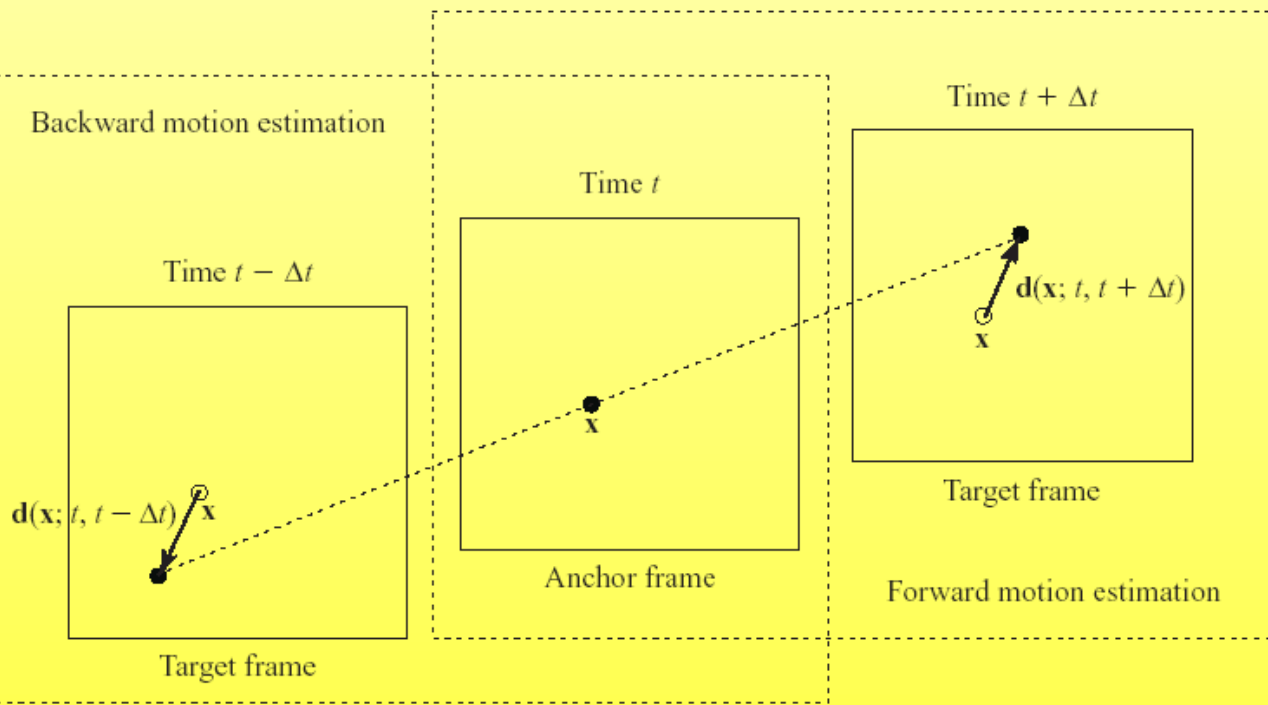
Region-based:
Entire frame is divided into regions, each region corresponding to an object or sub-object with consistent motion, represented by a few parameters.



(d)

Other representation: mesh-based (control grid) (to be discussed later)

Notations



Anchor frame: $\psi_1(\mathbf{x})$

Target frame: $\psi_2(\mathbf{x})$

Motion parameters: \mathbf{a}

Motion vector at a pixel in the anchor frame: $\mathbf{d}(\mathbf{x})$

Motion field: $\mathbf{d}(\mathbf{x}; \mathbf{a}), \mathbf{x} \in \Lambda$

Mapping function:

$$\mathbf{w}(\mathbf{x}; \mathbf{a}) = \mathbf{x} + \mathbf{d}(\mathbf{x}; \mathbf{a}), \mathbf{x} \in \Lambda$$

Motion Estimation Criteria

- Minimize displaced frame difference

$$E_{DFD}(\mathbf{a}) = \sum_{\mathbf{x} \in \Lambda} |\psi_2(\mathbf{w}(\mathbf{x}; \mathbf{a})) - \psi_1(\mathbf{x})|^p$$

- For $p = 2$ (MSE), the necessary condition for minimum is that the derivative is zero. Let $\mathbf{z} = \mathbf{w}(\mathbf{x}; \mathbf{a}) = \mathbf{x} + \mathbf{d}(\mathbf{x}; \mathbf{a})$. Then, the derivative is given by

$$\begin{aligned} \frac{\partial E_{DFD}}{\partial \mathbf{a}} &= 2 \sum_{\mathbf{x} \in \Lambda} (\psi_2(\mathbf{z}) - \psi_1(\mathbf{x})) \frac{\partial \psi_2}{\partial \mathbf{z}} \frac{\partial \mathbf{z}}{\partial \mathbf{a}} \\ &= 2 \sum_{\mathbf{x} \in \Lambda} (\psi_2(\mathbf{z}) - \psi_1(\mathbf{x})) \frac{\partial \psi_2}{\partial \mathbf{z}} \frac{\partial \mathbf{d}}{\partial \mathbf{a}} \end{aligned}$$

Motion Estimation Criteria

- Optical Flow Equation

$$\begin{aligned} \frac{\partial \psi}{\partial x} d_x + \frac{\partial \psi}{\partial y} d_y + \frac{\partial \psi}{\partial t} d_t &= 0 \\ \Rightarrow \frac{\partial \psi}{\partial x} v_x + \frac{\partial \psi}{\partial y} v_y + \frac{\partial \psi}{\partial t} &= 0 \end{aligned}$$

where $\mathbf{d}(\mathbf{x}; \mathbf{a}) = [d_x, d_y]^T$ and

$\mathbf{v}(\mathbf{x}; \mathbf{a}) = [v_x, v_y]^T = [d_x/d_t, d_y/d_t]^T$.

- Thus, obtain the motion vector parameters \mathbf{a} which minimizes the cost function

$$E_{OF}(\mathbf{a}) = \sum_{\mathbf{x} \in \Lambda} \left| \frac{\partial \psi}{\partial \mathbf{x}} \mathbf{v}(\mathbf{x}; \mathbf{a}) + \frac{\partial \psi}{\partial t} \right|^p$$

Lucas-Kanade Method

- Based on optical flow equation
- Assuming all pixels in a small block surrounding a pixel have the same motion vector

Lucas-Kanade Method

■ Optical Flow Equation $\frac{\partial \psi}{\partial x} v_x + \frac{\partial \psi}{\partial y} v_y + \frac{\partial \psi}{\partial t} = 0$

■ For each pixel one equation two unknowns (v_x, v_y)

■ Under-determined system

■ Idea

■ Assume that a block \mathcal{B} centered around the current pixel has the same motion and then minimize

$$E(v_x, v_y) = \sum_{\mathbf{x} \in \mathcal{B}} \left(\frac{\partial \psi}{\partial x}(\mathbf{x}) v_x + \frac{\partial \psi}{\partial y}(\mathbf{x}) v_y + \frac{\partial \psi}{\partial t}(\mathbf{x}) \right)^2$$

■ Then, the optimal vector is set as the vector of the current pixel

Lucas-Kanade Method

■ Solution

■ By setting the partial derivatives with respect to v_x and v_y to zeros

$$\sum_{\mathbf{x} \in \mathcal{B}} \left(\frac{\partial \psi}{\partial x}(\mathbf{x})v_x + \frac{\partial \psi}{\partial y}(\mathbf{x})v_y + \frac{\partial \psi}{\partial t}(\mathbf{x}) \right) \frac{\partial \psi}{\partial x}(\mathbf{x}) = 0,$$

$$\sum_{\mathbf{x} \in \mathcal{B}} \left(\frac{\partial \psi}{\partial x}(\mathbf{x})v_x + \frac{\partial \psi}{\partial y}(\mathbf{x})v_y + \frac{\partial \psi}{\partial t}(\mathbf{x}) \right) \frac{\partial \psi}{\partial y}(\mathbf{x}) = 0,$$

■ we have

$$\begin{bmatrix} v_x \\ v_y \end{bmatrix} = - \begin{bmatrix} \sum_{\mathbf{x} \in \mathcal{B}} \frac{\partial \psi}{\partial x}(\mathbf{x}) \frac{\partial \psi}{\partial x}(\mathbf{x}) & \sum_{\mathbf{x} \in \mathcal{B}} \frac{\partial \psi}{\partial x}(\mathbf{x}) \frac{\partial \psi}{\partial y}(\mathbf{x}) \\ \sum_{\mathbf{x} \in \mathcal{B}} \frac{\partial \psi}{\partial x}(\mathbf{x}) \frac{\partial \psi}{\partial y}(\mathbf{x}) & \sum_{\mathbf{x} \in \mathcal{B}} \frac{\partial \psi}{\partial y}(\mathbf{x}) \frac{\partial \psi}{\partial y}(\mathbf{x}) \end{bmatrix}^{-1} \begin{bmatrix} \sum_{\mathbf{x} \in \mathcal{B}} \frac{\partial \psi}{\partial x}(\mathbf{x}) \frac{\partial \psi}{\partial t}(\mathbf{x}) \\ \sum_{\mathbf{x} \in \mathcal{B}} \frac{\partial \psi}{\partial y}(\mathbf{x}) \frac{\partial \psi}{\partial t}(\mathbf{x}) \end{bmatrix}$$