MA3113: Topics in Mathematical Image Processing I
Some Project Topics

Suh-Yuh Yang (楊肅煜)

Department of Mathematics, National Central University
Jhongli District, Taoyuan City 320317, Taiwan

February 17, 2026

Topic 1: Perona-Malik-type denoising models

- The earliest nonlinear diffusion model proposed in image processing is the anisotropic diffusion (非等方性的擴散) by Perona and Malik [PM-1990].
- Let $f : \overline{\Omega} \subset \mathbb{R}^2 \to \mathbb{R}$ be a given noisy grayscale image. The IBVP of the Perona-Malik equation can be posed as follows:

$$
\begin{aligned}
&\frac{\partial u}{\partial t} - \nabla \cdot \Big(g(|\nabla u|) \nabla u \Big) = 0 \quad \text{for } (t, \boldsymbol{x}) \in (0, T) \times \Omega, \\
&u(0, \boldsymbol{x}) = f(\boldsymbol{x}) \text{ for } \boldsymbol{x} \in \overline{\Omega}, \\
&\nabla u \cdot \boldsymbol{n} = 0 \quad \text{for } t \in [0, T] \text{ and } \boldsymbol{x} \in \partial\Omega.
\end{aligned}
$$

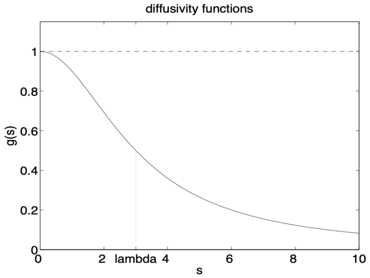
where g is a smooth non-increasing diffusivity function with

$$
\begin{aligned}
&g(0) = 1, \quad g(s) \geq 0, \quad \text{and } \lim_{s \to \infty} g(s) = 0; \\
&|\nabla u| = |(u_x, u_y)^\top| := \sqrt{u_x^2 + u_y^2}.
\end{aligned}
$$

Two diffusivity functions: Perona and Malik suggested two different choices for the diffusivity function g:

$$g(s) = \frac{1}{1+(s/\lambda)^2}, \quad s \geq 0,$$
$$g(s) = e^{-(s/\lambda)^2}, \quad s \geq 0,$$

where $\lambda > 0$ is a given parameter.



Topic 1: Perona-Malik-type denoising models (cont'd)

The regularized Perona-Malik diffusion equation: To alleviate the *staircasing effect*, we consider the following regularized equation:

$$\frac{\partial u}{\partial t} - \nabla \cdot \Big(g(|\nabla u_\sigma|) \nabla u \Big) = 0,$$

where $u_\sigma := G_\sigma * u$, $\sigma > 0$, is a Gaussian-smoothed version of u.

(a) **(b)** **(c)**

(a) original noisy image; (b) Perona-Malik diffusion; (c) regularized Perona-Malik diffusion.

Topic 1: Perona-Malik-type denoising models (cont'd)

References

- **[PM-1990]** P. Perona and J. Malik, Scale-space and edge detection using anisotropic diffusion, *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 12 (1990), pp. 629-639.
- **[E-2012]*** E. Erdem, Nonlinear diffusion PDEs, *Lecture Notes,* Hacettepe University, 2012.

  https://web.cs.hacettepe.edu.tr/~erkut/bil717.s12/

Topic 2: A denoising model with adaptive diffusivity

Let $f : \overline{\Omega} \subset \mathbb{R}^2 \to \mathbb{R}$ be a given noisy grayscale image. Consider the following minimization problem:

$$\min_u \left(\int_\Omega \phi(|\nabla u|)\, d\boldsymbol{x} + \frac{\lambda}{2} \int_\Omega (u - f)^2\, d\boldsymbol{x} \right),$$

where $\lambda > 0$ is a regularization parameter, ϕ is a regularization function, and $|\cdot|$ denotes the usual ℓ^2-norm,

$$|\nabla u| = |(u_x, u_y)^\top| := \sqrt{u_x^2 + u_y^2}.$$

Note: ROF model: $\phi(|\nabla u|) := |\nabla u|$.

Topic 2: A denoising model with adaptive diffusivity (cont'd)

Staircasing effect and edge-preserving: Consider the regularization function $\phi(s) = s^p$:

- $1 < p \leq 2$: eliminating the *staircasing effect* to obtain a more smooth image, e.g., $p = 2$, the Tikhonov model.
- $p = 1$: the ROF total variation model.
- $0 < p < 1$: more effective than ROF ($p = 1$) *for preserving edges, but it results in a non-convex problem.*

Topic 2: A denoising model with adaptive diffusivity (cont'd)

We propose the following convex regularization model:

$$\min_u \int_\Omega \left(\frac{\alpha_{p,q}}{q} |\nabla u|^q + \frac{\lambda}{2}(u - f)^2 \right) d\boldsymbol{x},$$

where

- $q = 1$ or 2, and here we consider $q = 2$.
- *$\alpha_{p,q} = \alpha_{p,q}(|\nabla u^*(x)|) > 0$ is a spatially variable controller that will be defined later.*
- $|\nabla u^*(x)|$ is a quantity that approximates the magnitude of the gradient of the original image at x.
- $0 < p \leq 1$ and $\lambda > 0$ is a constant regularization parameter.

The Euler-Lagrange equation can be derived as follows:

$$\alpha_{p,2}\Delta u + \lambda(f - u) = 0 \quad \text{in } \Omega. \qquad (\star_1)$$

Topic 2: A denoising model with adaptive diffusivity (cont'd)

From another perspective, to keep the edge-preserving property, we consider the lower-order regularization model with $0 < p \leq 1$:

$$\min_u \int_\Omega \frac{1}{p}|\nabla u|^p + \frac{\lambda}{2}(u-f)^2\,dx.$$

The Euler-Lagrange equation is given by

$$\nabla \cdot \left(|\nabla u|^{p-2}\nabla u\right) + \lambda(f-u) = 0 \quad \text{in } \Omega.$$

In a very rough approximation form, we may consider the related equation,

$$|\nabla u|^{p-2}\Delta u + \lambda(f-u) = 0 \quad \text{in } \Omega. \qquad (\star_2)$$

Topic 2: A denoising model with adaptive diffusivity (cont'd)

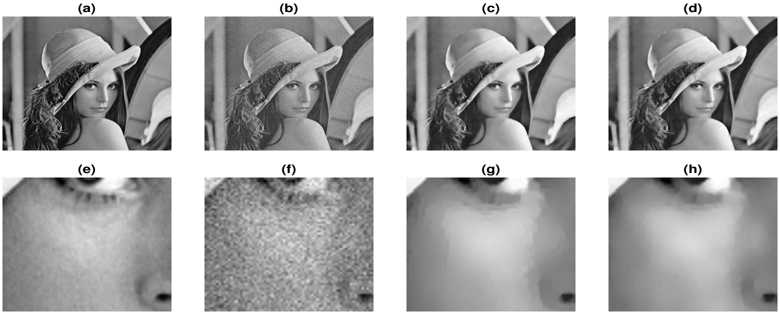
In comparison with $(\star_1)$ and $(\star_2)$ suggests us to define the adaptive diffusivity $\alpha_{p,2}$ as

$$\alpha_{p,2}(|\nabla u^*(x)|) := \begin{cases} |\nabla u^*(x)|^{p-2} & \text{if} \quad |\nabla u^*(x)| \neq 0, \\ \varepsilon^{-1} & \text{if} \quad |\nabla u^*(x)| = 0, \end{cases}$$

where $0 < p \leq 1$ and $\varepsilon > 0$ is a prescribed small number.

Remark: For the image pixels where $|\nabla u^*(x)| \approx 0$, the model for $q = 2$ and $0 < p \leq 1$ with the above adaptive diffusivity has extra smoothing effect since $\alpha_{p,2}$ is large and its EL equation is very close to $\Delta u = 0$ at those pixels.

Topic 2: A denoising model with adaptive diffusivity (cont'd)



Numerical results "Lena", where the noise level is $\sigma = 15$ and the denoising parameter is $\lambda = 0.04$ for both ROF and present algorithm. (a) and (e) are the global and local plots of the original "Lena" image; (b) and (f) correspond to the noisy version; (c) and (g) are the denoising results by ROF; (d) and (h) are the denoising results by present algorithm with $(p, q) = (1, 2)$.

Topic 2: A denoising model with adaptive diffusivity (cont'd)

References

- **[C-2007]** R. Chartrand, Nonconvex regularization for shape preservation, In: *Proceedings of IEEE International Conference on Image Processing (ICIP),* (2007), pp. 293-296.
- **[LZOX-2015]** Y. Lou, T. Zeng, S. Osher, and J. Xin, A weighted difference of anisotropic and isotropic total variation model for image processing, *SIAM Journal on Imaging Sciences,* 8 (2015), pp. 1798-1823.
- **[YL-2015]** J. Yan and W. S. Lu, Image denoising by generalized total variation regularization and least squares fidelity, *Multidimensional Systems and Signal Processing*, 26 (2015), pp. 243-266.
- **[HSY-2018]*** P.-W. Hsieh, P.-C. Shao, and S,-Y, Yang, A regularization model with adaptive diffusivity for variational image denoising, *Signal Processing,* 149 (2018), pp. 214-228.

Topics 3: Sparse representation and dictionary learning

Sparse representation problem: Given a signal vector $\boldsymbol{x} \in \mathbb{R}^m$ and a dictionary matrix $\boldsymbol{D} \in \mathbb{R}^{m\times n}$, we seek a coefficient vector $\boldsymbol{z}^* \in \mathbb{R}^n$ such that

$$\boldsymbol{z}^* = \arg\min_{\boldsymbol{z}} \left\{ \frac{1}{2}\|\boldsymbol{x} - \boldsymbol{D}\boldsymbol{z}\|_2^2 + \lambda \|\boldsymbol{z}\|_1 \right\},$$

where $\lambda > 0$ is a penalty parameter.

Sparse dictionary learning problem: Let $\{\boldsymbol{x}_i\}_{i=1}^N \subset \mathbb{R}^m$ be a given dataset of signals. We seek a dictionary matrix $\boldsymbol{D} = [\boldsymbol{d}_1, \boldsymbol{d}_2, \cdots, \boldsymbol{d}_n]$ $\in \mathbb{R}^{m\times n}$ together with the sparse coefficient vectors $\{\boldsymbol{z}_i\}_{i=1}^N \subset \mathbb{R}^n$ that solve the minimization problem:

$$\min_{\boldsymbol{D},\{\boldsymbol{z}_i\}} \left\{ \frac{1}{2}\sum_{i=1}^N \|\boldsymbol{x}_i - \boldsymbol{D}\boldsymbol{z}_i\|_2^2 + \lambda \sum_{i=1}^N \|\boldsymbol{z}_i\|_1 \right\}$$

$$\text{subject to } \|\boldsymbol{d}_k\|_2 \le 1, \ \forall\, 1 \le k \le n,$$

where $\lambda > 0$ is a penalty parameter.

Topics 3: Compact form of the SDL problem

To simplify the formulation of the SDL problem, we define

$$\boldsymbol{X} = [\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_N] \in \mathbb{R}^{m\times N}, \qquad \boldsymbol{Z} = [\boldsymbol{z}_1, \boldsymbol{z}_2, \cdots, \boldsymbol{z}_N] \in \mathbb{R}^{n\times N}.$$

Then the SDL problem can be posed as follows: *Given a training data matrix* $\boldsymbol{X}$*, find a dictionary matrix* $\boldsymbol{D}$ *and a coefficient matrix* $\boldsymbol{Z}$ *such that*

$$\min_{\boldsymbol{D},\boldsymbol{Z}} \left(\frac{1}{2} \|\boldsymbol{X} - \boldsymbol{D}\boldsymbol{Z}\|_F^2 + \lambda \|\boldsymbol{Z}\|_{1,1} \right)$$

$$\text{subject to } \|\boldsymbol{d}_k\|_2 \leq 1, \ \forall\, 1 \leq k \leq n.$$

In the compact form, $\|\cdot\|_F$ denotes the Frobenius norm defined as follows: for a matrix $\boldsymbol{A} = [\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_N] \in \mathbb{R}^{m\times N}$,

$$\|\boldsymbol{A}\|_F^2 := \sum_{i=1}^{N} \|\boldsymbol{a}_i\|_2^2$$

and $\|\boldsymbol{Z}\|_{1,1}$ is the $L_{1,1}$-norm which is defined as

$$\|\boldsymbol{Z}\|_{1,1} := \sum_{i=1}^{N} \|\boldsymbol{z}_i\|_1.$$

Topic 3: SR and DL for image fusion and mage inpainting

(L) source A; (M) source B; (R) fused image.

corrupted image *inpainted image*

$\lambda = 0.2$ *for DL and* $\lambda = 0.1$ *for SR*

Topic 3: SR and DL for image fusion (cont'd)

- *Matlab codes developed by Brendt Wohlberg at Los Alamos National Laboratory: `http://brendt.wohlberg.net/software/SPORCO/`
- **[SWM-2007]** Y. Sharon, J. Wright, and Y. Ma, Computation and relaxation of conditions for equivalence between ℓ^1 and ℓ^0 minimization, *UIUC Technical Report UILU-ENG-07-2008, 2007.*
- **[LCWW-2016]** Y. Liu, X. Chen, R. K. Ward, and Z. J. Wang, Image fusion with convolutional sparse representation, *IEEE Signal Processing Letters,* 23 (2016), pp. 1882-1886.
- **[MA3111]*** S.-Y. Yang, *Image Inpainting, Lecture Slides,* 2024.

Topic 4: Poisson image editing

The objective is to seamlessly clone a region selected from an image over a background image.

Let u^* be the destination image on $\overline{D}$ and I_0 be the selected image. This problem can be formulated as a variational problem:

$$\min_{u \in \mathcal{C}^2(D),\, u|_{D\setminus\Omega} = u^*|_{D\setminus\Omega}} \int_\Omega |\nabla u - \boldsymbol{v}|^2 d\boldsymbol{x},$$

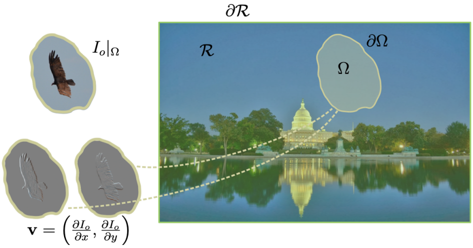
- $D \subset \mathbb{R}^2$ is an open bounded subset of $\mathbb{R}^2$, $\overline{D}$ represents the image domain, and $\Omega \subset D$.
- $\mathcal{C}^2(D)$ is the set of real functions twice differentiable over D.
- $\boldsymbol{v}$ is a differentiable gradient field obtained from the selected image, $\boldsymbol{v} = \nabla I_0$.

Topic 4: Poisson image editing (cont'd)

The solution of the minimization problem must satisfy the Euler-Lagrange equation:

$$\Delta u = \nabla \cdot v \quad \text{in } \Omega, \quad \text{and } u = u^* \quad \text{on} \partial\Omega,$$

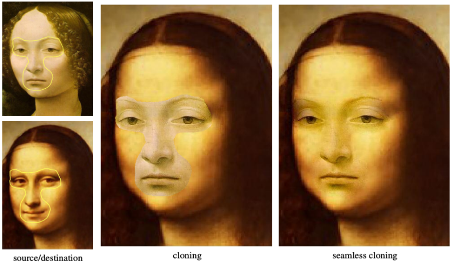
and of course outside $\overline{\Omega}$ the solution is trivial and takes the same values of u^*.



$$D := \mathcal{R}$$

Topic 4: Poisson image editing (cont'd)

- **[PGB-2003]** P. Pérez, M. Gangnet, and A. Blake, Poisson image editing, *SIGGRAPH03: Special Interest Group on Computer Graphics and Interactive Techniques,* San Diego, California, July 27-31, 2003.
- **[DFM-2016]*** J. M. Di Martino, G. Facciolo, E. Meinhardt-Llopis, Poisson image editing, *Image Processing On Line,* 6 (2016), pp. 300-325.



Topic 5: Single image dehazing

Dark channel prior: *most local patches (excluding sky regions) in outdoor haze-free images contain some pixels whose intensity is very low in at least one color channel (R, G, or B). That is,*

$$I^{\mathrm{dark}}(\boldsymbol{x}) \approx 0, \quad \forall\, \boldsymbol{x} \in \overline{\Omega}.$$

The dark channel I^{dark} *of an image* $\boldsymbol{I} = (I_R, I_G, I_B)$ *on* $\overline{\Omega}$ *is defined as:*

$$I^{\mathrm{dark}}(\boldsymbol{x}) = \min_{\boldsymbol{y} \in N(\boldsymbol{x})} \left(\min_{c \in \{R,G,B\}} I_c(\boldsymbol{y}) \right), \quad \forall\, \boldsymbol{x} \in \overline{\Omega}.$$

(L) $\boldsymbol{J}$*; (M) For each* $\boldsymbol{x}$*, minimum of its* (R, G, B) *values; (R)* I^{dark}.

Topic 5: Single image dehazing (cont'd)

The dark channel prior helps estimate the haze concentration and recover a clear image by combining it with *the atmospheric scattering model:*

$$\boldsymbol{I}(\boldsymbol{x}) = t(\boldsymbol{x})\boldsymbol{J}(\boldsymbol{x}) + (1 - t(\boldsymbol{x}))\boldsymbol{A},$$

- $\boldsymbol{I}$ is the observed hazy image,
- $\boldsymbol{J}$ is the haze-free image,
- $\boldsymbol{A}$ is the global atmospheric light (a constant vector),
- $t(\boldsymbol{x})$ is the transmission map, representing the portion of light that reaches the camera without being scattered. If the atmosphere is homogenous, we usually set $t(\boldsymbol{x}) = e^{-\beta d(\boldsymbol{x})}$.

hazy image $\boldsymbol{I}$

dehazed image

Topic 5: Single image dehazing (cont'd)

- **[HST-2011]** K. He, J. Sun, and X. Tang, Single image haze removal using dark channel prior, *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 33 (2011), pp. 2341-2353.
- **[YE-2021]** G. Yang and A. N. Evans, Improved single image dehazing methods for resourceconstrained platforms, *Journal of Real-Time Image Processing,* 18 (2021), pp. 2511-2525.
- **[ZMS-2015]** Q. Zhu, J. Mai, and L. Shao, A fast single image haze removal algorithm using color attenuation prior, *IEEE Transactions on Image Processing,* 24 (2015), pp. 3522-3533.

  Color attenuation prior: *the hazy regions are characterized by high brightness and low saturation in the HSB color space, and the concentration of haze is positively correlated with the difference between brightness and saturation.*

Topic 6: Variational methods for image contrast enhancement

The main purpose of contrast enhancement is to adjust the image intensity to enhance the quality and features of the image for a better human visual perception or machine vision identification.

A low-light image and its enhanced result

References:

- **[HSY-2020]** P.-W. Hsieh, P.-C. Shao, and S.-Y. Yang, Adaptive variational model for contrast enhancement of low-light images, *SIAM Journal on Imaging Sciences,* 13 (2020), pp. 1-28.
- **[MA3111]** S.-Y. Yang, *Image Contrast Enhancement, Lecture Slides,* 2024.

Topic 7: Robust PCA for low-rank textures

- Singular value decomposition (SVD), principal component analysis (PCA), and the Eckart-Young Theorem:

$$\|D - A\| \leq \|D - B\|, \quad \forall\, B \text{ and } \operatorname{rank}(B) = k,$$

  where D is the data matrix with low rank r and A is the matrix constructed by SVD with the first $k \leq r$ singular values.

- Robust principal component analysis (RPCA): If D is corrupted, then we first decompose D into $D = A + E$ and consider

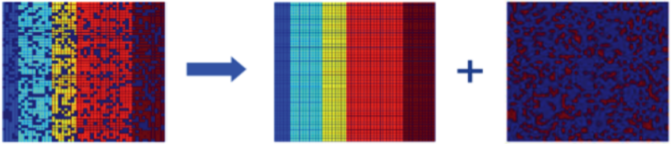
$$\min_A \; \|A\|_* + \lambda \|E\|_1 \quad \text{subject to } D = A + E,$$

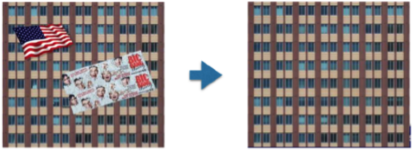
  where $\|\cdot\|_*$ denotes the nuclear norm.

- Augmented Lagrangian multiplier method:

$$\mathcal{L} = \|A\|_* + \lambda \|E\|_1 + \langle Y, D - A - E \rangle + \frac{\mu}{2} \|D - A - E\|_F^2.$$

Topic 7: Robust PCA for low-rank textures (cont'd)





Topic 7: Robust PCA for low-rank textures (cont'd)

References:

- **[BZ-2014]** T. Bouwmans and E. H. Zahzah, Robust PCA via principal component pursuit: a review for a comparative evaluation in video surveillance, *Computer Vision and Image Understanding,* 122 (2014), pp. 22-34.
- **[Lin-2016]** Z. Lin, A review on low-rank models in data analysis, *Big Data and Information Analytics,* 1 (2016), pp. 139-161.
- **[MA3111]** S.-Y. Yang, *Principal Component Pursuit, Lecture Slides,* 2024.

Topic 8: Level-set approach for image segmentation

The Chan-Vese two-phase segmentation model:

$$\min_{c_1,c_2,\mathcal{C}} \left(\mu|\mathcal{C}| + \nu|\Omega_{\text{in}}| + \lambda_1 \int_{\Omega_{\text{in}}} (f(\boldsymbol{x}) - c_1)^2 \, d\boldsymbol{x} + \lambda_2 \int_{\Omega_{\text{out}}} (f(\boldsymbol{x}) - c_2)^2 \, d\boldsymbol{x} \right).$$

- Ω_{in} denotes the region enclosed by contour $\mathcal{C}$ with area $|\Omega_{\text{in}}|$, and $\Omega_{\text{out}} := \Omega \setminus \Omega_{\text{in}}$.
- $\mu > 0$, $\nu > 0$, $\lambda_1 > 0$, and $\lambda_2 > 0$ are tuning parameters (actually, one of them can be fixed as 1).
- Chan-Vese model finds a piecewise constant function u and an edge set $\mathcal{C}$ to minimize the energy functional, where u has only two constant values,

$$u(\boldsymbol{x}) = \begin{cases} c_1, & \boldsymbol{x} \text{ is inside } \mathcal{C}, \\ c_2, & \boldsymbol{x} \text{ is outside } \mathcal{C}. \end{cases}$$

Topic 8: Level-set approach for image segmentation (cont'd)

Introducing the level-set function ϕ, then the minimization can be solved by *an alternating iterative scheme,* i.e., alternatingly updating c_1, c_2 and ϕ.

(1) Fixed ϕ, the optimal values of c_1 and c_2 are the region averages,

$$c_1 = \frac{\int_\Omega f(\boldsymbol{x}) H(\phi(\boldsymbol{x}))\, d\boldsymbol{x}}{\int_\Omega H(\phi(\boldsymbol{x}))\, d\boldsymbol{x}}, \quad c_2 = \frac{\int_\Omega f(\boldsymbol{x}) \big(1 - H(\phi(\boldsymbol{x}))\big)\, d\boldsymbol{x}}{\int_\Omega \big(1 - H(\phi(\boldsymbol{x}))\big)\, d\boldsymbol{x}}.$$

(2) Fixed c_1, c_2, we solve the IBVP to reach a steady-state:

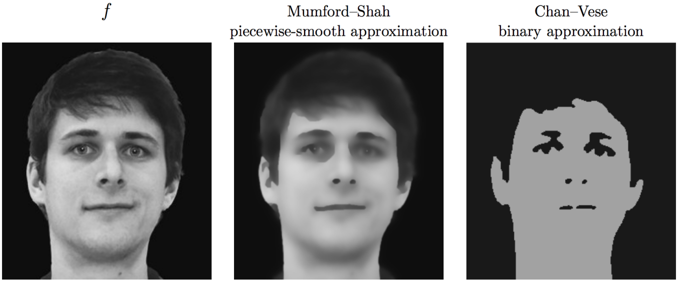
$$\frac{\partial \phi}{\partial t} = \delta_\epsilon(\phi)\left(\mu \nabla \cdot \frac{\nabla \phi}{|\nabla \phi|} - \nu - \lambda_1 (f - c_1)^2 + \lambda_2 (f - c_2)^2\right),$$

$$t > 0, \boldsymbol{x} \in \Omega,$$

$$\phi(0, \boldsymbol{x}) = \phi_0(\boldsymbol{x}), \boldsymbol{x} \in \Omega,$$

$$\frac{\partial \phi}{\partial \boldsymbol{n}} = 0 \text{ on } \partial\Omega, t \geq 0.$$

Topic 8: Level-set approach for image segmentation (cont'd)



- **[G-2012]** P. Getreuer, Chan-Vese segmentation, *Image Processing On Line,* 2 (2012), pp. 214-224.
- **[MA3111]** S.-Y. Yang, Image segmentation, *Lecture Slides,* 2024.

Topic 9: Optical flow estimation

Let $I : (\Omega, T) \subset \mathbb{R}^3 \to \mathbb{R}$ be a given image sequence, of grayscale values in space $z = (x, y) \in \Omega$ and time $t \in T$. The optical flow

$$\boldsymbol{w}(\boldsymbol{z}) = \big(u(\boldsymbol{z}), v(\boldsymbol{z})\big)^\top$$

is defined as the velocity field, such that one frame $I(x, y, t)$ is translated to the next frame $I(x + \delta x, y + \delta y, t + \delta t)$ by the mapping $(\delta x, \delta y, \delta t) := (u\delta t, v\delta t, \delta t)$. *The brightness constancy assumption* can be written as

$$I(x, y, t) = I(x + \delta x, y + \delta y, t + \delta t).$$

Using the first-order Taylor expansion at (x, y, t), we obtain

$$I_x(x, y, t)u + I_y(x, y, t)v + I_t(x, y, t) = 0.$$

That is,

$$\nabla I \cdot \boldsymbol{w} + I_t = 0.$$

Topic 9: Optical flow estimation (cont'd)

In 1981, Horn and Schunck proposed a global smoothing approach by minimizing the following energy functional

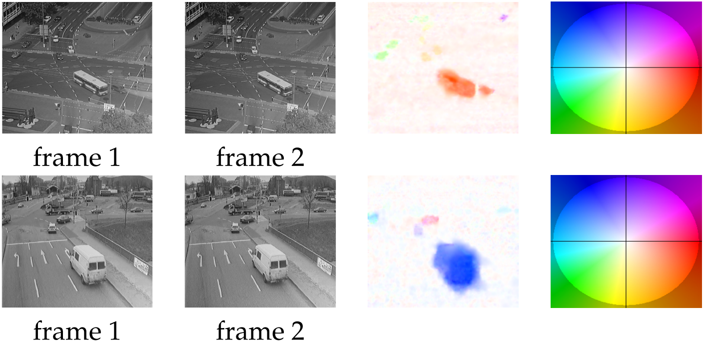
$$\mathcal{E}_{HS}(\boldsymbol{w}) = \int_{\Omega} \alpha(|\nabla u|^2 + |\nabla v|^2) + (\nabla I \cdot \boldsymbol{w} + I_t)^2 dxdy,$$

where $\alpha > 0$ is the regularization parameter. By the calculus of variation, we obtain the Euler-Lagrange equations,

$$\begin{aligned} \alpha\nabla^2 u - I_x(I_x u + I_y v + I_t) &= 0, \\ \alpha\nabla^2 v - I_y(I_x u + I_y v + I_t) &= 0, \end{aligned}$$

and then solve the equations approximately.

Topic 9: Optical flow estimation (cont'd)



References:

- **[HS-1981]** B. K. P. Horn, and B. G. Schunck, Determining optical flow, *Artificial Intelligence,* 17 (1981), pp. 185-203.
- **[GN-1987]** M. A. Gennert, and S. Negahdaripour, Relaxing the brightness constancy assumption in computing optical flow, *A. I. Memo No. 975,* (1987).

Topic 10: Video foreground extraction

- Video foreground extraction is a fundamental task in computer vision, aiming to accurately extract the foreground from a sequence of frames containing *dynamic foreground objects and a static background.*
- Let $\boldsymbol{D} \in \mathbb{R}^{m \times n}$ be the data matrix formed by stacking vectorized grayscale video frames as columns of a video sequence. We assume that $\boldsymbol{D}$ is composed of a rank-1 background component and a sparse foreground component, modeled as

$$\boldsymbol{D} = \boldsymbol{u}\mathbf{1}^\top + \boldsymbol{S},$$

  where $\boldsymbol{u} \in \mathbb{R}^{m \times 1}$ and $\mathbf{1} \in \mathbb{R}^{n \times 1}$ is the all-ones vector.
- Based on this formulation, we consider the following rank-1 principal component pursuit (PCP) problem:

$$\min_{\boldsymbol{u} \in \mathbb{R}^{m \times 1}, \boldsymbol{S} \in \mathbb{R}^{m \times n}} \|\boldsymbol{S}\|_1 \quad \text{subject to } \boldsymbol{D} = \boldsymbol{u}\mathbf{1}^\top + \boldsymbol{S},$$

  where $\|\cdot\|_1$ denotes the ℓ_1 norm.

Topic 10: Video foreground extraction (cont'd)

Define the augmented Lagrangian function by

$$\begin{aligned}\mathcal{L}(\boldsymbol{u},\boldsymbol{S},\boldsymbol{Y}) &:= \|\boldsymbol{S}\|_1 + \langle \boldsymbol{Y}, \boldsymbol{D}-\boldsymbol{u}\mathbf{1}^\top - \boldsymbol{S}\rangle + \frac{\mu}{2}\|\boldsymbol{D}-\boldsymbol{u}\mathbf{1}^\top-\boldsymbol{S}\|_F^2 \\ &= \|\boldsymbol{S}\|_1 + \frac{\mu}{2}\|\boldsymbol{D}-\boldsymbol{u}\mathbf{1}^\top-\boldsymbol{S}+\mu^{-1}\boldsymbol{Y}\|_F^2 - \frac{1}{2\mu}\|\boldsymbol{Y}\|_F^2,\end{aligned}$$

where $\mu > 0$ is a penalty parameter, $\boldsymbol{Y}$ is a Lagrange multiplier. The iterative scheme of the ALM method for the rank-1 PCP problem is as follows:

$$\begin{aligned}\boldsymbol{u}^{(k+1)} &:= \arg\min_{\boldsymbol{u}} \mathcal{L}_\mu(\boldsymbol{u},\boldsymbol{S}^{(k)},\boldsymbol{Y}^{(k)}) \\ &= \arg\min_{\boldsymbol{u}} \left(\|\boldsymbol{S}^{(k)}\|_1 + \frac{\mu}{2}\|\boldsymbol{D}-\boldsymbol{u}\mathbf{1}^\top-\boldsymbol{S}^{(k)}+\mu^{-1}\boldsymbol{Y}^{(k)}\|_F^2 - \frac{1}{2\mu}||\boldsymbol{Y}^{(k)}||_F^2\right), \\ \boldsymbol{S}^{(k+1)} &:= \arg\min_{\boldsymbol{S}} \mathcal{L}_\mu(\boldsymbol{u}^{(k+1)},\boldsymbol{S},\boldsymbol{Y}^{(k)}) \\ &= \arg\min_{\boldsymbol{S}} \left(\|\boldsymbol{S}\|_1 + \frac{\mu}{2}\|\boldsymbol{D}-\boldsymbol{u}^{(k+1)}\mathbf{1}^\top-\boldsymbol{S}+\mu^{-1}\boldsymbol{Y}^{(k)}\|_F^2 - \frac{1}{2\mu}\|\boldsymbol{Y}^{(k)}\|_F^2\right), \\ \boldsymbol{Y}^{(k+1)} &:= \boldsymbol{Y}^{(k)} + \mu(\boldsymbol{D}-\boldsymbol{u}^{(k+1)}\mathbf{1}^\top-\boldsymbol{S}^{(k+1)}).\end{aligned}$$

Topic 10: Video foreground extraction (cont'd)

- **[Wei-2025]** S.-T. Wei, Video foreground extraction via efficient rank-1 principal component pursuit, *Master Thesis,* Department of Mathematics, National Central University, June 2025.
- **[MA3111]** S.-Y. Yang, Principal component pursuit problem, *Lecture Slides,* 2025.