

面向可信视觉任务的不变表征方法研究

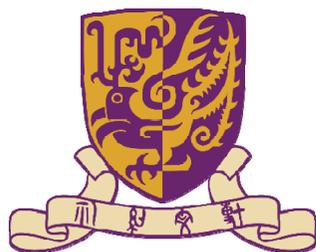
Invariant Representations for Trustworthy Vision Tasks

报告人：祁树仁

南京航空航天大学(毕业)

香港中文大学(入职)

2024年7月3日



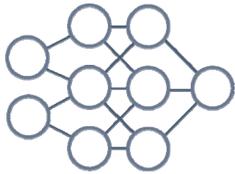
Center
for
Mathematical
Artificial
Intelligence
CMAI

汇报纲要

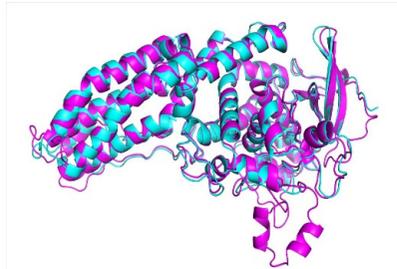
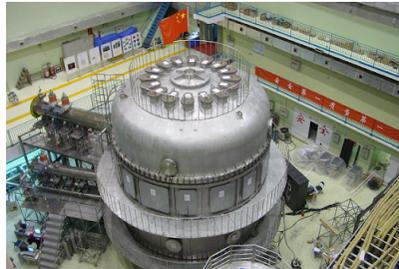
- **研究背景与意义**
- 分析经典不变量
- 完善全局不变量
- 设计局部不变量
- 探索层次不变量
- 研究总结与展望

可信人工智能

- 回顾人工智能近年来的发展，随着深度机器学习方法取得突破性进展，其应用场景正自然拓展至**稳健性**、**可解释性**攸关的可信领域。
- 包括但不限于加速和增强科学研究、生物和身份认证、医疗数据分析、自动驾驶和机器人、以及信息安全和取证。



AI-based Representation,
Solution, Prediction, and Control



自然世界的对称性先验

- 宏观上，人工智能系统是对**自然世界物理系统的一种数字模拟**，因而对**自然先验**的利用水平决定了人工智能系统的稳健性、可解释性水平。
- 在诸多的先验中，存在一类最为基础和通用的先验——**对称性**：系统的对称性是一种变换，使所述系统的某些属性保持**不变**。



F. Klein

1872

Erlangen Program



E. Noether

1918

Noether's Theorem



H. Weyl

1929

The Book of Symmetry



C. N. Yang

1954

Yang-Mills Theory

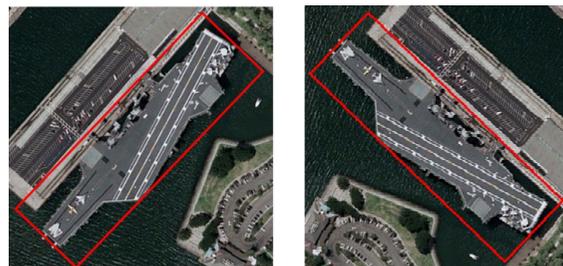


R. L. Mills

数据表征的对称性原则



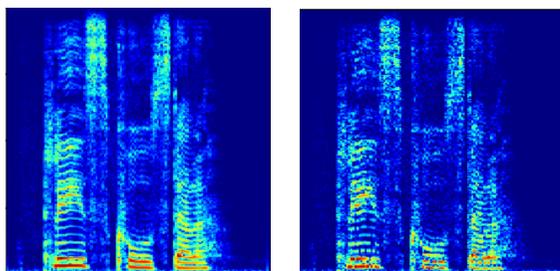
Image Classification
Position



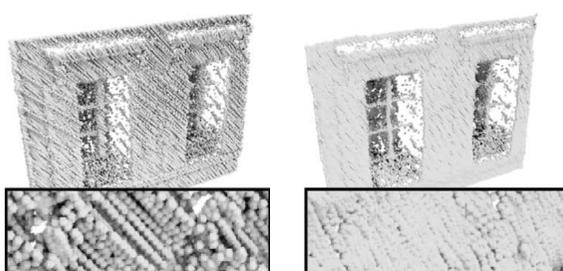
Remote Sensing
Orientation



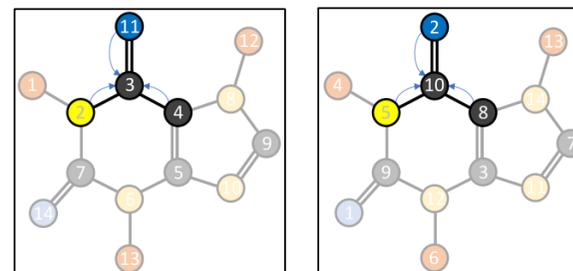
Self-driving Car
Motion blurring



Speech Command
Time warping



Point Cloud Analysis
Noise



Molecular Properties
Permutation

研究现状和难点

原则指引

因此在视觉表征中嵌入任务所需的**不变性或对称性结构**是实现计算机视觉系统稳健性和可解释性的一种重要的原则性保障

路径分析

手工表征

举例 不变量、SIFT、DAISY

优势 知识驱动，可以定制化满足对称性先验

劣势 强假设，缺乏数据自适应性和大规模数据判别性

路径分析

学习表征

举例 CNN

优势 自适应性和判别性，同时满足给定对称性先验（仅平移）

劣势 稳健性和解释性问题，比如缺乏定制化引入先验的手段

VS

宏观难点

不变性和判别性表征分别在手工和学习路径上近乎独立的发展导致在迈向稳健和可解释视觉任务时面临**二者棘手的权衡**

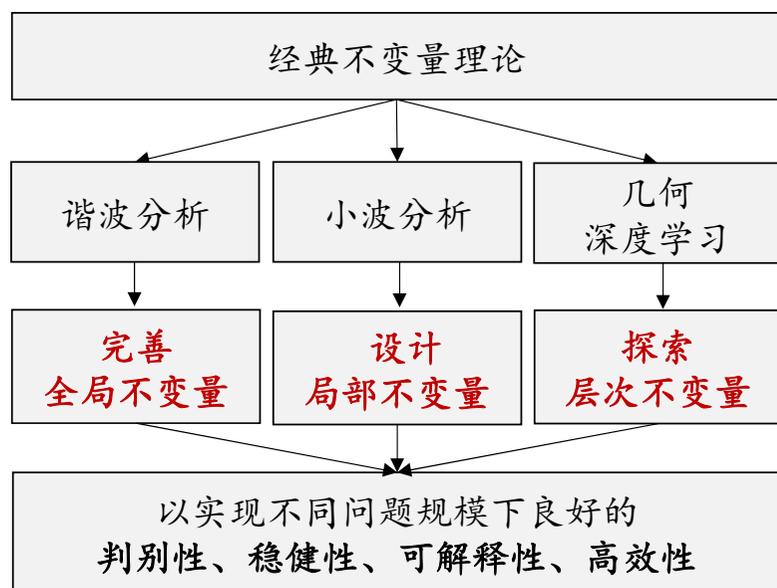
研究内容和创新

立论核心

以可信人工智能为**背景**、
以自然世界的对称性先验为**原则**、
在**理论**和**实践**层面发展了面向视觉任务的不变表征方法

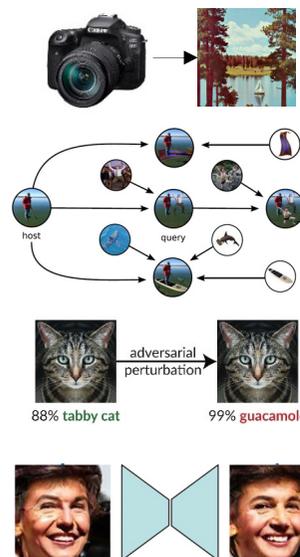
理论层面

不变表征理论研究



实践层面

视觉取证应用研究



内容、硬件指纹的
取证分析

人为篡改的
取证分析

对抗性扰动的
取证分析

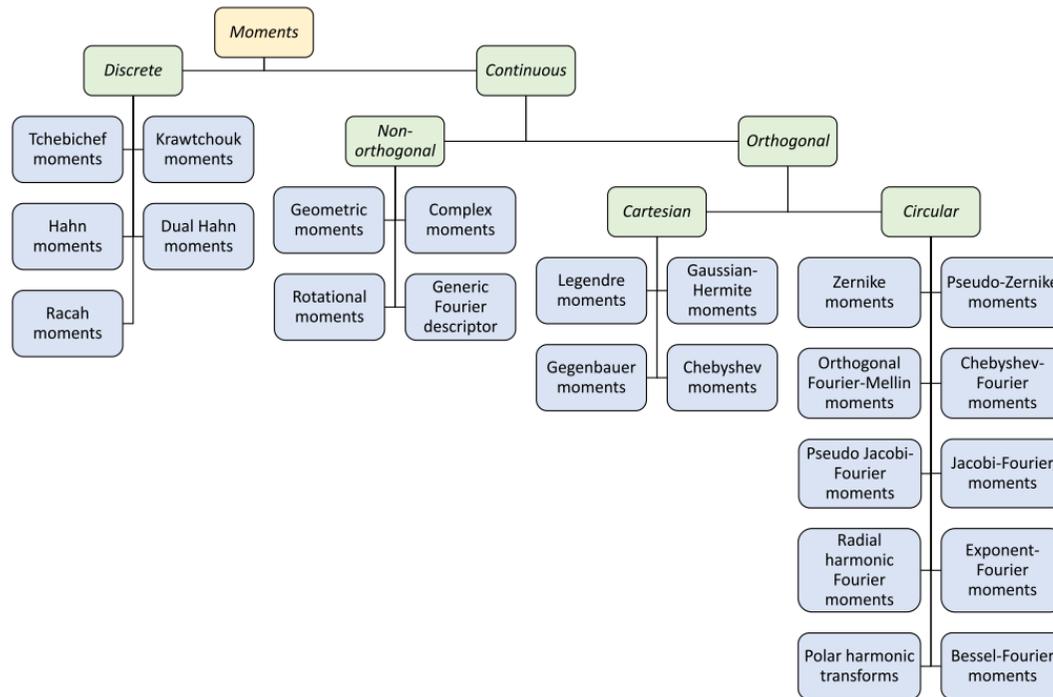
机器生成内容的
取证分析

汇报纲要

- 研究背景与意义
- **分析经典不变量**
- 完善全局不变量
- 设计局部不变量
- 探索层次不变量
- 研究总结与展望

分析经典不变量

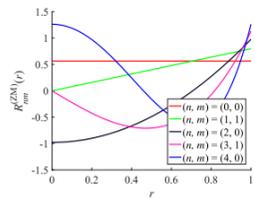
- 深入分析了经典不变量方法，为后续三个核心课题提供了理论和实践的基础。



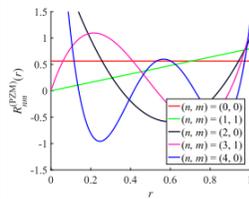
Method	Radial Basis Function
ZM	$R_{nm}^{(ZM)}(r) = \sqrt{\frac{n+1}{\pi}} \sum_{k=0}^{\frac{n- m }{2}} \frac{(-1)^k (n-k)! r^{n-2k}}{k! (\frac{n+ m }{2}-k)! (\frac{n- m }{2}-k)!}$
PZM	$R_{nm}^{(PZM)}(r) = \sqrt{\frac{n+1}{\pi}} \sum_{k=0}^{\frac{n- m }{2}} \frac{(-1)^k (2n+1-k)! r^{n-k}}{k! (n+ m +1-k)! (n- m -k)!}$
OFMM	$R_n^{(OFMM)}(r) = \sqrt{\frac{n+1}{\pi}} \sum_{k=0}^n \frac{(-1)^{n+k} (n+k+1)! r^k}{k! (n-k)! (k+1)!}$
CHFM	$R_n^{(CHFM)}(r) = \frac{2}{\pi} \left(\frac{1-r}{r}\right)^{\frac{1}{4}} \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} \frac{(-1)^k (n-k)! (4r-2)^{n-2k}}{k! (n-2k)!}$
PJFM	$R_n^{(PJFM)}(r) = \sqrt{\frac{(n+2)(r-r^2)}{\pi(n+3)(n+1)}} \sum_{k=0}^n \frac{(-1)^{n+k} (n+k+3)! r^k}{k! (n-k)! (k+2)!}$
JFM	$R_n^{(JFM)}(p, q, r) = \sqrt{\frac{r^{q-2} (1-r)^{p-q} (p+2n) \cdot \Gamma(q+n) \cdot n!}{2\pi \Gamma(p+n) \cdot \Gamma(p-q+n+1)}} \sum_{k=0}^n \frac{(-1)^k \Gamma(p+n+k) r^k}{k! (n-k)! \Gamma(q+k)}$
RHFM	$R_n^{(RHFM)}(r) = \begin{cases} \frac{1}{\sqrt{2\pi r}} & n = 0 \\ \sqrt{\frac{1}{\pi r}} \sin(\pi(n+1)r) & n > 0 \text{ \& } n \text{ odd} \\ \sqrt{\frac{1}{\pi r}} \cos(\pi n r) & n > 0 \text{ \& } n \text{ even} \end{cases}$
EFM	$R_n^{(EFM)}(r) = \frac{1}{\sqrt{2\pi r}} \exp(j2n\pi r)$
PCET	$R_n^{(PCET)}(r) = \frac{1}{\sqrt{\pi}} \exp(j2n\pi r^2)$
PCT	$R_n^{(PCT)}(r) = \begin{cases} \frac{1}{\sqrt{\pi}} & n = 0 \\ \sqrt{\frac{2}{\pi}} \cos(n\pi r^2) & n > 0 \end{cases}$
PST	$R_n^{(PST)}(r) = \sqrt{\frac{2}{\pi}} \sin(n\pi r^2)$
BFM	$R_n^{(BFM)}(r) = \frac{1}{\sqrt{\pi} J_{\nu+1}(\lambda_n)} J_{\nu}(\lambda_n r), J_{\nu}(x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k! \Gamma(\nu+k+1)} \left(\frac{x}{2}\right)^{\nu+2k}$

分析经典不变量

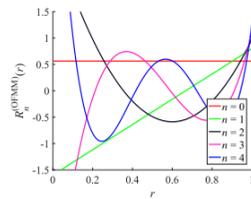
- 在理论层面，首先回顾不变量的发展历史并梳理其中的核心知识。



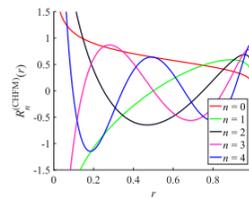
(a) ZM



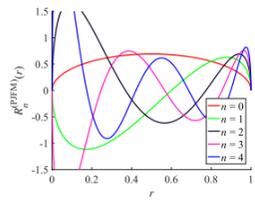
(b) PZM



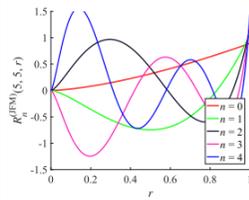
(c) OFMM



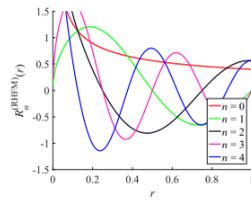
(d) CHFMM



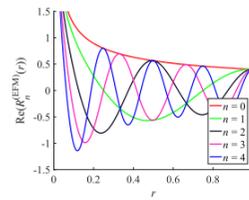
(e) PJFM



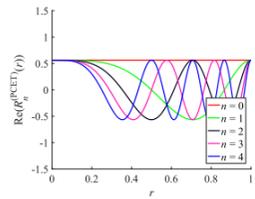
(f) JFM ($p, q = 5$)



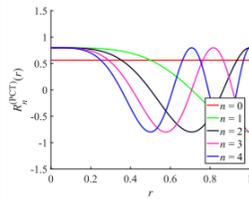
(g) RHFMM



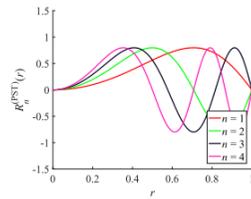
(h) EFM (real part)



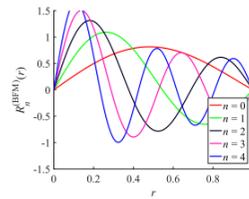
(i) PCET (real part)



(j) PCT



(k) PST

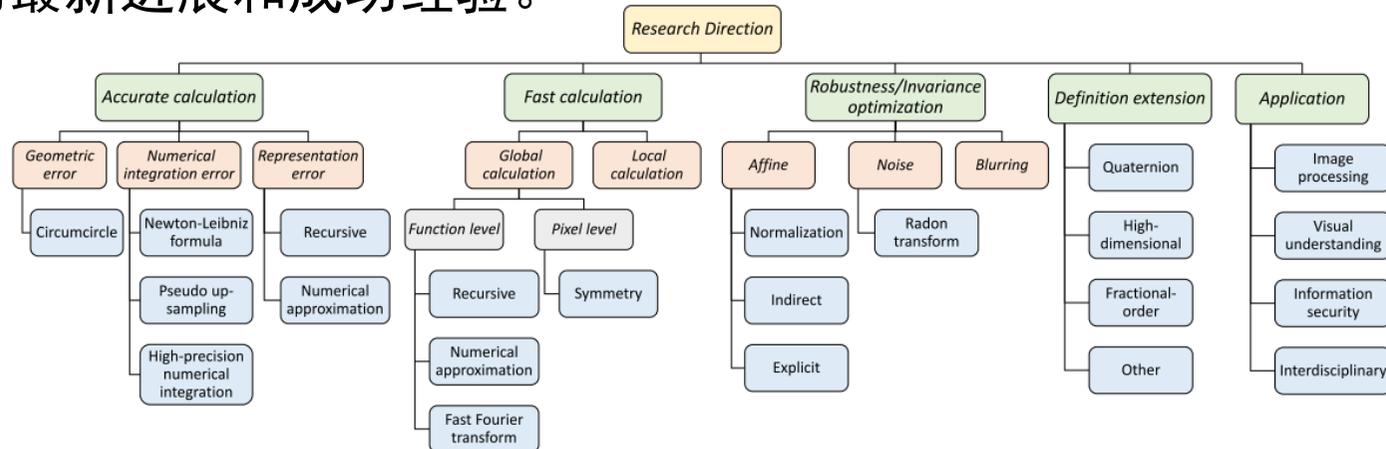


(l) BFM

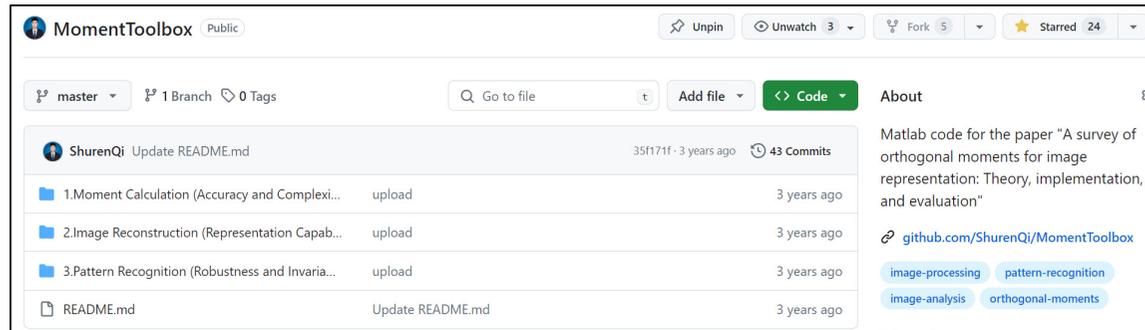
Method	Parameter	Complexity	Stability	Number of Zeros	Distribution of Zeros
ZM	$n \in \mathbb{N}, m \in \mathbb{Z}, m \leq n, n - m = \text{even}$	high	poor	$(n - m)/2$	biased
PZM	$n \in \mathbb{N}, m \in \mathbb{Z}, m \leq n$	high	poor	$n - m $	biased
OFMM	$n \in \mathbb{N}, m \in \mathbb{Z}$	high	poor	n	basically uniform
CHFMM	$n \in \mathbb{N}, m \in \mathbb{Z}$	high	poor	n	basically uniform
PJFM	$n \in \mathbb{N}, m \in \mathbb{Z}$	high	poor	n	basically uniform
JFM	$n \in \mathbb{N}, m \in \mathbb{Z}, p, q \in \mathbb{R}, p - q > -1, q > 0$	high	poor	n	basically uniform
RHFMM	$n \in \mathbb{N}, m \in \mathbb{Z}$	low	medium	n	uniform
EFM	$n \in \mathbb{Z}, m \in \mathbb{Z}$	low	medium	$2n$	uniform
PCET	$n \in \mathbb{Z}, m \in \mathbb{Z}$	low	good	$2n$	biased
PCT	$n \in \mathbb{N}, m \in \mathbb{Z}$	low	good	n	biased
PST	$n \in \mathbb{N}^+, m \in \mathbb{Z}$	low	good	$n - 1$	biased
BFM	$n \in \mathbb{N}, m \in \mathbb{Z}, v \in \mathbb{R}, \lambda_n = n\text{-th zero of } J_v(x)$	very high	medium	n	basically uniform

分析经典不变量

- 随后，重点分析不变量研究中快速/精确计算、稳健性/不变性优化、定义扩展和应用方面的最新进展和成功经验。

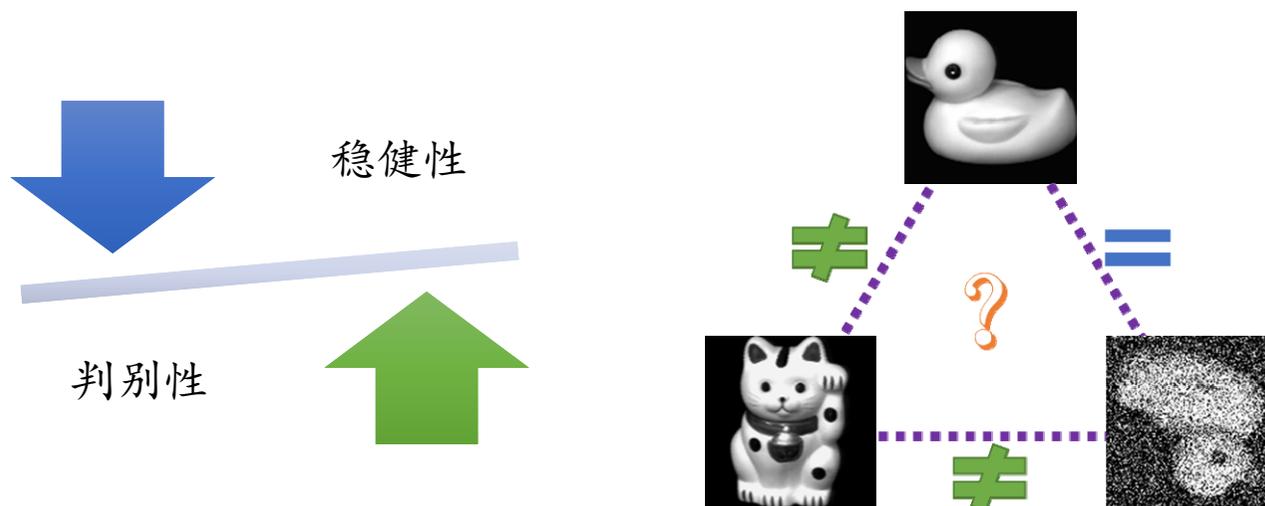


- 最后在实践中，给出了系统性的开源实现 MomentToolbox，并基于此提供了广泛的数值基准。



分析经典不变量

- 基于以上对经典不变量研究现状的综述，可以总结出以下两方面的**判别性**问题：
 - 理论上，经典不变量通常基于图像的全局表征假设，而对于视觉任务中更具信息量的图像**局部和层次表征**，缺乏相应的理论扩展；
 - 实践上，经典不变量通常用于小规模视觉任务，而对于具有不变性或对称性先验的**更大规模视觉任务**，缺乏相应的实践扩展。



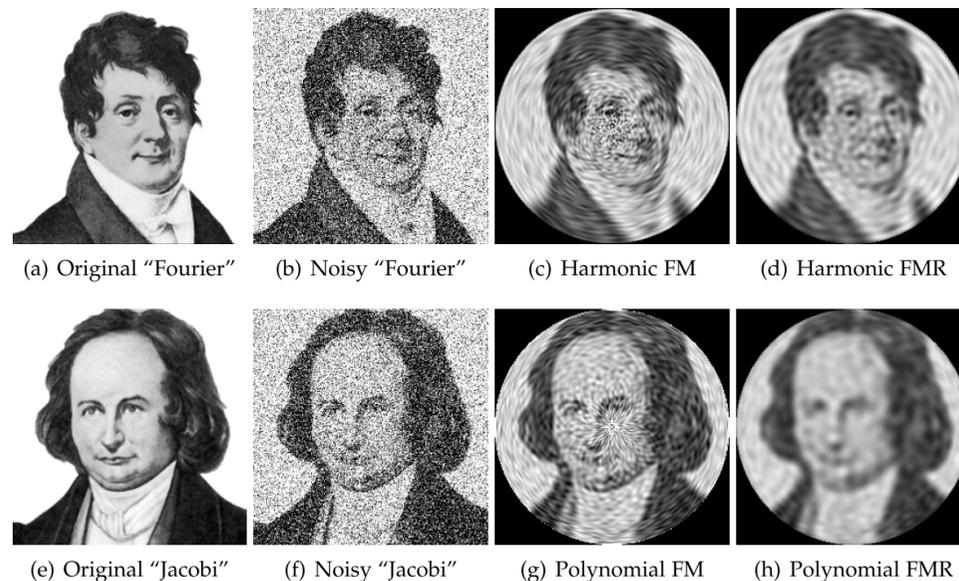
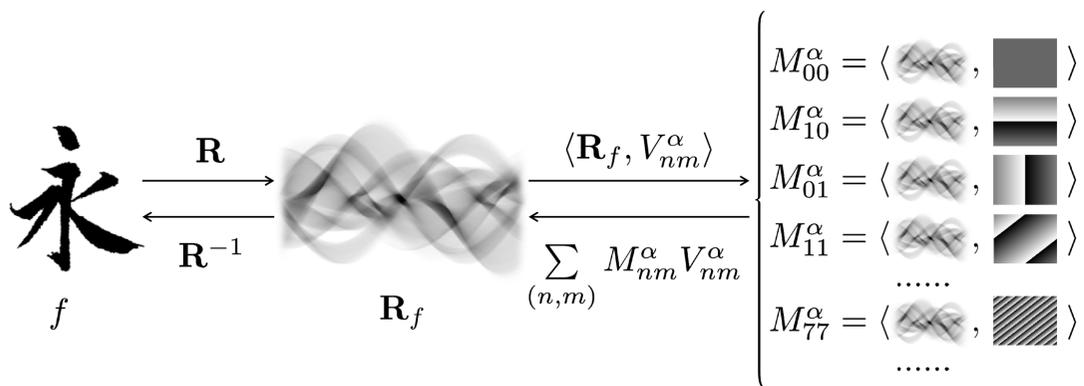
汇报纲要

- 研究背景与意义
- 分析经典不变量
- **完善全局不变量**
- 设计局部不变量
- 探索层次不变量
- 研究总结与展望

完善全局不变量

- 完善现有的全局不变量理论：
- 首先，将经典频域不变量推广为新的**时频分析**方法，从而表现出更高的判别性。该推广基于分数阶多项式基和傅里叶基的研究。此外，还将经典几何不变量推广为**噪声不变量**，从而更适用于非理想成像条件。该推广基于不变子空间的研究。

方法	拉东统计量 [245,246]	拉东正交变换 [114,247,248]	图像分数阶矩 [80,159]	拉东整数阶矩 [113]	本章方法
广义性			✓		✓
旋转不变性		✓	✓	✓	✓
噪声稳健性	✓	✓		✓	✓
时频判别性			✓		✓



完善全局不变量

- 形式化方面，分别从**隐式**和**显式**方式构造上述全局不变量，并考虑了两种路径的**计算**和**关系**问题。

隐式定义及其快速精确计算

定理 3.1: 谐波 FMR 的傅里叶定义——解析公式。由定义 3.2 和 3.3 给出的谐波 FMR 可被重写为以下傅里叶形式^[259]:

$$\langle \mathbf{R}_f, V_{nm}^\alpha \rangle = 2\pi \mathcal{F}(S(\gamma, \vartheta)), \quad (3.16)$$

其中, $\gamma = r^\alpha$, $\vartheta = \theta/2\pi$, $S(\gamma, \vartheta) = \sqrt{\gamma^{\alpha/2-1}/2\pi\alpha} \times \mathbf{R}_f\left(\underbrace{\gamma^{\frac{1}{\alpha}}}_r, \underbrace{2\pi\vartheta}_\theta\right)$.

定理 3.2: 多项式 FMR 的递归定义——解析公式。由定义 3.4 给出的多项式径向基函数可被重写为以下递归形式^[159]:

$$R_n^\alpha(p, q, r) = \sqrt{\frac{(p+2n)\alpha r^{\alpha q-1}(1-r^\alpha)^{p-q}}{r}} C_n(p, q) P_n^\alpha(p, q, r), \quad (3.19)$$

$$P_n^\alpha = (L_1 r^\alpha + L_2) P_{n-1}^\alpha + L_3 P_{n-2}^\alpha, n \geq 2, \quad (3.20)$$

$$C_n = \sqrt{\frac{n(q+n-1)}{(p+n-1)(p-q+n)}} C_{n-1}, n \geq 1, \quad (3.21)$$

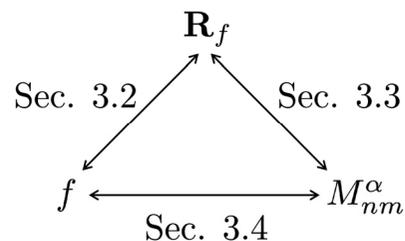
显式定义及与隐式、几何矩的关系

定理 3.3: 谐波 FMR 的显式定义。由定义 3.2 和 3.3 给出的谐波 FMR 可被重写为如下显式形式^[113]:

$$\langle \mathbf{R}_f, V_{nm}^\alpha \rangle = \sum_{k=0}^{\infty} W_1^+(n, k) \sum_{t=0}^{\infty} \binom{\alpha k + \frac{\alpha}{2}}{t} \Theta_{\alpha k + \frac{\alpha}{2} - t, t} G_{\alpha k + \frac{\alpha}{2} - t, t}, \quad (3.22)$$

定理 3.4: 多项式 FMR 的显式定义。由定义 3.2 和 3.4 给出的多项式 FMR 可被重写为如下显式形式^[113]:

$$\langle \mathbf{R}_f, V_{nm}^\alpha \rangle = \sum_{k=0}^n W_2(\alpha, p, q, n, k) \sum_{s=0}^{\infty} W_3(p, q, s) \sum_{t=0}^{\infty} \binom{\alpha(s+k+\frac{\alpha}{2})}{t} \Theta_{\alpha(s+k+\frac{\alpha}{2})-t, t} G_{\alpha(s+k+\frac{\alpha}{2})-t, t}, \quad (3.28)$$



可以被解释为几何矩的
无穷线性组合

$$G_{\xi_1, \xi_2} = \int_x \int_y f(x, y) x^{\xi_1} y^{\xi_2} dx dy$$

完善全局不变量

- 与数据增广训练或增强预处理的深度表征相比，其在小规模的经典视觉任务以及版权保护任务中表现出优势。

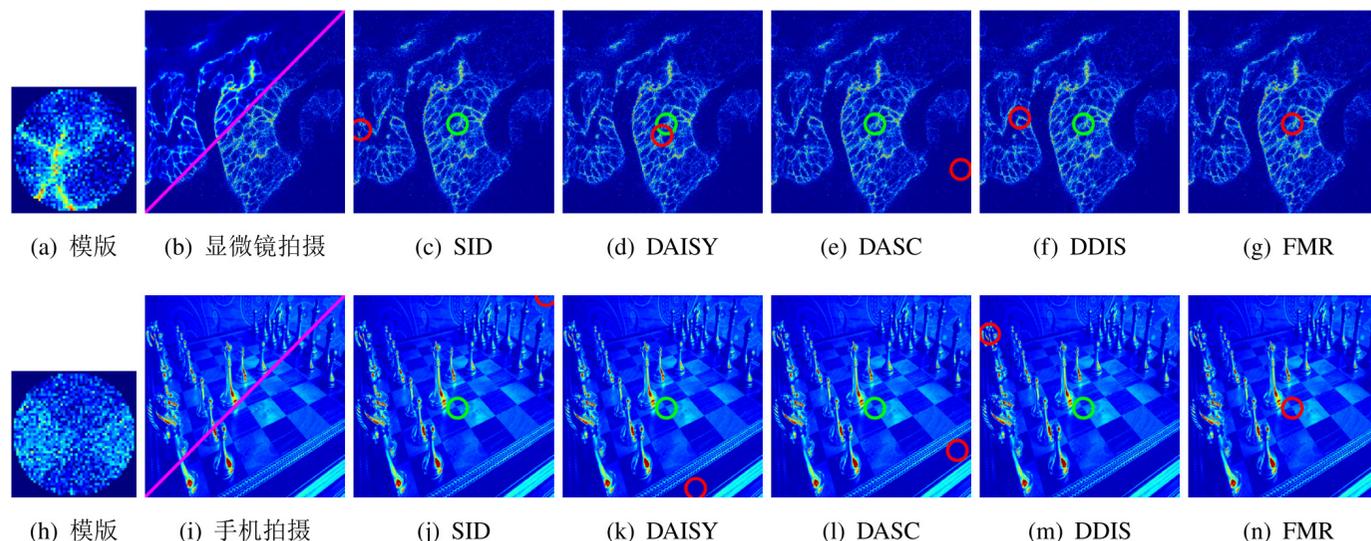


表 3.4 真实荧光显微镜图像上不同表征的模板匹配分数 (%)。

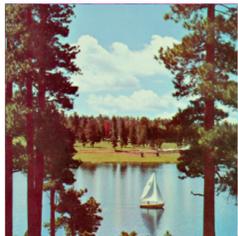
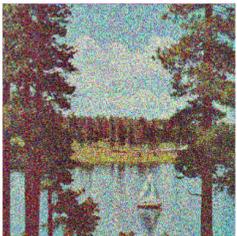
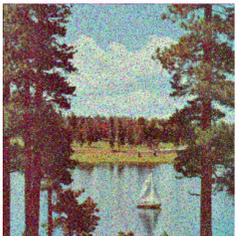
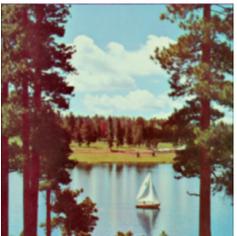
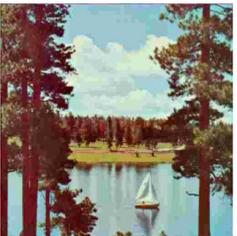
表征方法	理想模版	旋转模版	含噪模版	旋转并含噪模版
SID	50	38	31	20
DAISY	73	11	50	11
DASC	80	7	22	2
DDIS	68	67	65	65
FMR	99	99	96	93

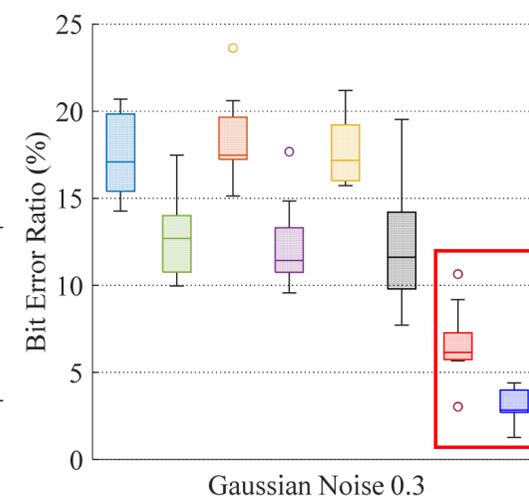
表 3.5 真实手机图像上不同表征的模板匹配分数 (%)。

表征方法	理想模版	旋转模版	含噪模版	旋转并含噪模版
SID	81	64	27	20
DAISY	99	13	7	0
DASC	99	10	5	1
DDIS	71	71	71	70
FMR	100	99	89	83

完善全局不变量

- 与数据增广训练或增强预处理的深度表征相比，其在**小规模的经典视觉任务以及版权保护任务**中表现出优势。

					
Original / PSNR	Rotation / 8.3	Gauss. Noise / 9.7	S. & P. Noise / 11.9	Blurring / 26.6	JPEG / 24.4
					
FM: BER % ↓	1.3	11.7	9.1	2.4	2.6
					
FMR: BER % ↓	0.9	2.9	3.1	0.6	0.5



汇报纲要

- 研究背景与意义
- 分析经典不变量
- 完善全局不变量
- 设计局部不变量**
- 探索层次不变量
- 研究总结与展望

设计局部不变量

- 将全局不变量理论推广到具有尺度空间的局部假设：

通过变量与坐标转换
给出局部不变量的广义形式

方法	启发式 局部不变量 ^[267]	笛卡尔矩个例扩展为 局部不变量 ^[91]	圆谐矩个例扩展为 局部不变量 ^[93]	本章方法
广义性				✓
旋转不变性	✓		✓	✓
常数阶复杂性		✓	✓	✓

全局定义

$$\langle f, V_{nm} \rangle = \iint_D R_n^*(r) A_m^*(\theta) f(r, \theta) r dr d\theta$$

变量与坐标转换

$$(x', y') = \frac{(x, y) - (u, v)}{w}$$

$$\begin{cases} r' = \sqrt{(x')^2 + (y')^2} = \frac{1}{w} \sqrt{(x-u)^2 + (y-v)^2} \\ \theta' = \arctan\left(\frac{y'}{x'}\right) = \arctan\left(\frac{y-v}{x-u}\right) \end{cases}$$

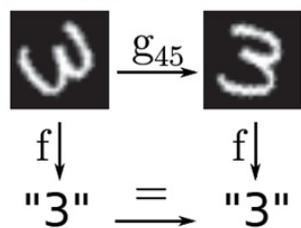
局部定义

$$\langle f, V_{nm}^{uvw} \rangle = \iint_D \underbrace{R_n^*\left(\frac{\sqrt{(x-u)^2 + (y-v)^2}}{w}\right)}_{(V_{nm}^{uvw}(x,y))^*} \underbrace{A_m^*\left(\arctan\left(\frac{y-v}{x-u}\right)\right)}_{\theta'} f(x, y) dx dy$$

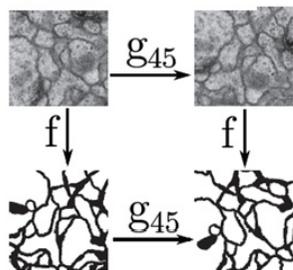
设计局部不变量

- 给出了上述局部定义在几何不变性、等变性和协变性方面的核心结论。

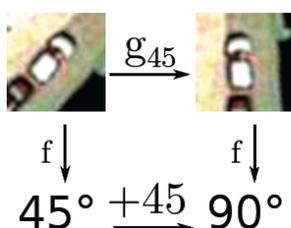
不变性



等变性



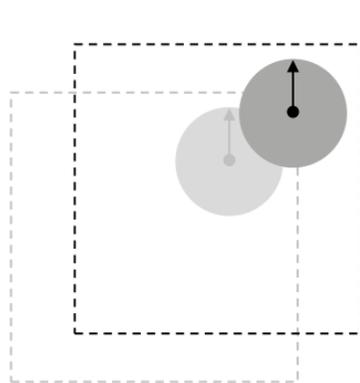
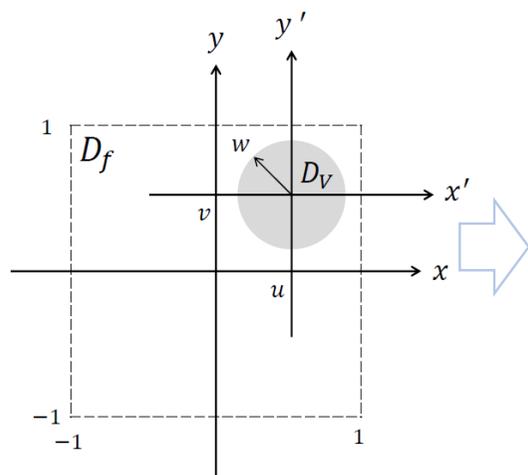
协变性



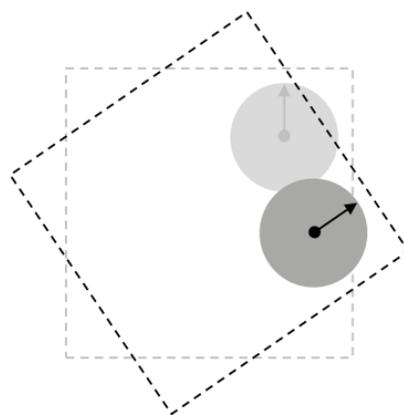
- 不变性—— $\mathcal{R}(\mathcal{D}(f)) \equiv \mathcal{R}(f)$;
- 等变性—— $\mathcal{R}(\mathcal{D}(f)) \equiv \mathcal{D}(\mathcal{R}(f))$;
- 协变性—— $\mathcal{R}(\mathcal{D}(f)) \equiv \mathcal{D}'(\mathcal{R}(f))$ 。

协变性是不变性和等变性的推广

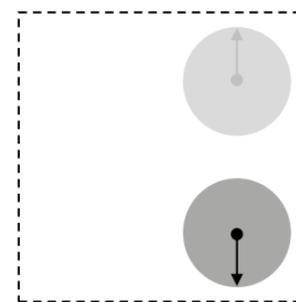
协变性/等变性成立，则构造不变表征可行



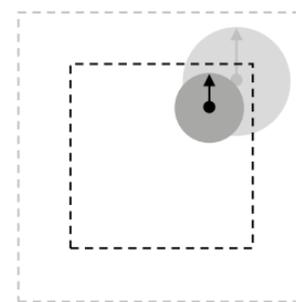
平移等变性



旋转不变性



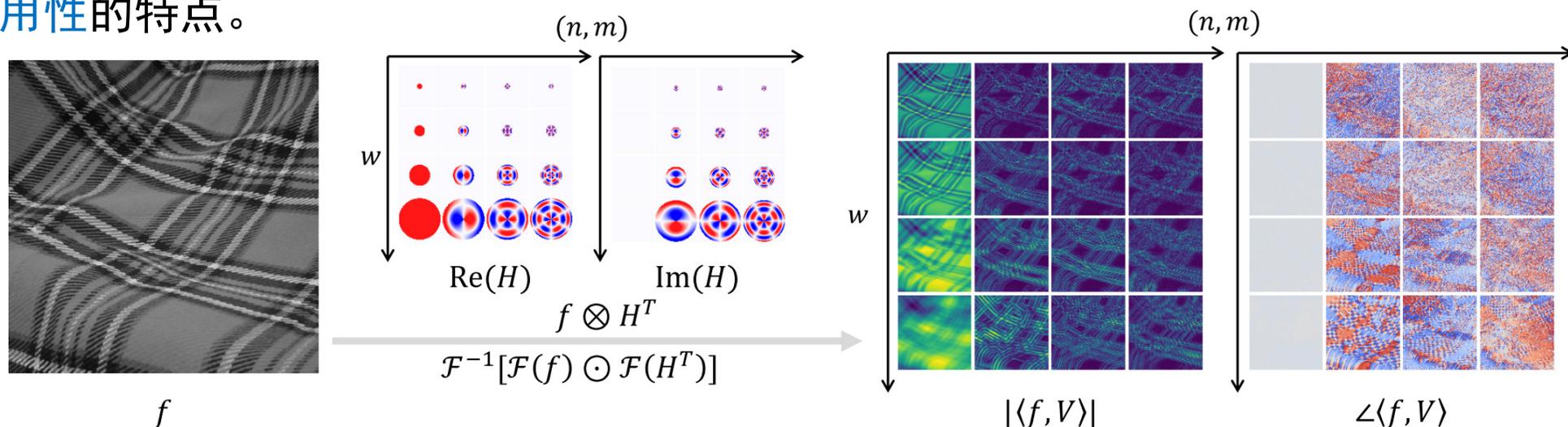
翻转不变性



缩放协变性

设计局部不变量

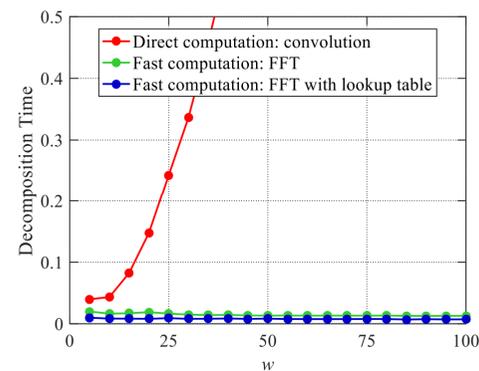
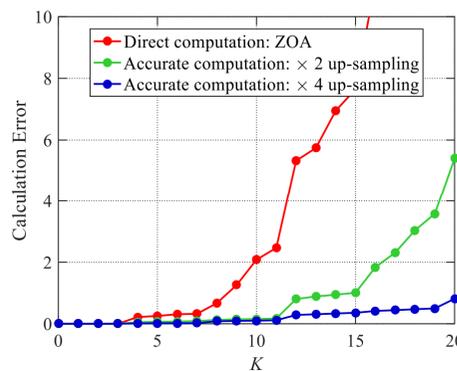
- 此外，推导出了高效的离散计算策略，具有低误差、常数阶复杂性以及基函数通用性的特点。



通过FFT实现快速高精度计算

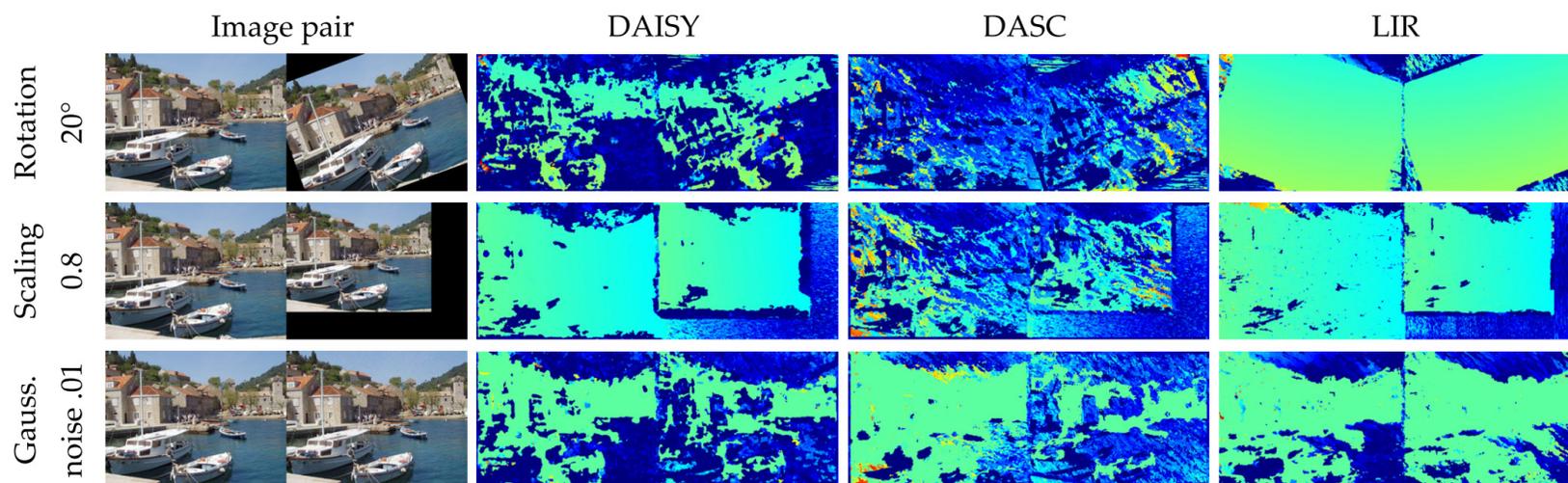
$$\mathcal{O}(w_{\max}^2 \#_{uv} \#_w) \quad \text{VS} \quad \mathcal{O}(\#_w \#_{uv} \log \#_{uv})$$

$$w_{\max}^2 \quad \text{VS} \quad \log(\#_{uv})$$



设计局部不变量

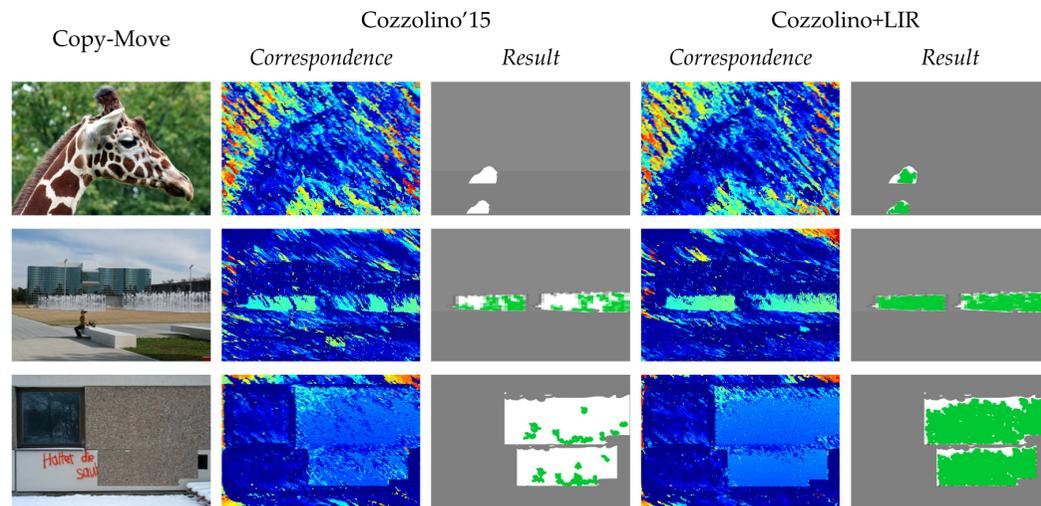
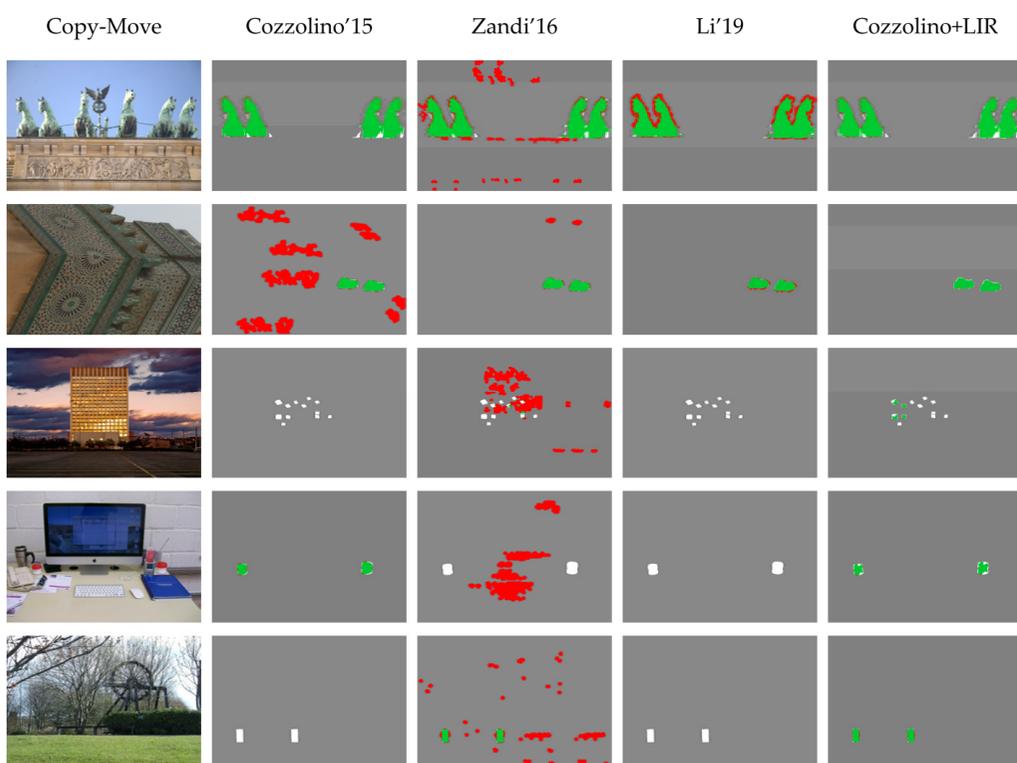
- 与先进的稠密局部表征相比，其在中等规模的经典视觉任务以及人为篡改取证任务中表现性能优势。



稠密匹配	旋转	翻转	缩放	信号劣化	特征维度	特征时间	匹配时间
DAISY	~ 10%	~ 0%	~ 70%	~ 40%	200	6秒	79秒
DASC	~ 0%	~ 0%	~ 0%	~ 50%	128	106秒	77秒
LIR	~ 80%	~ 90%	~ 40%	~ 40%	10	6秒	20秒

设计局部不变量

- 与先进的稠密局部表征相比，其在中等规模的经典视觉任务以及人为篡改取证任务中表现性能优势。



取证方法	Cozzolino'15 ^[217]	Cozzolino+LIR	引入本章方法的增益率
Precision	84.47	90.86	7.56
Recall	46.06	63.32	37.47
F1	56.15	71.18	26.77
平均匹配数	144519.75	270730.23	87.33

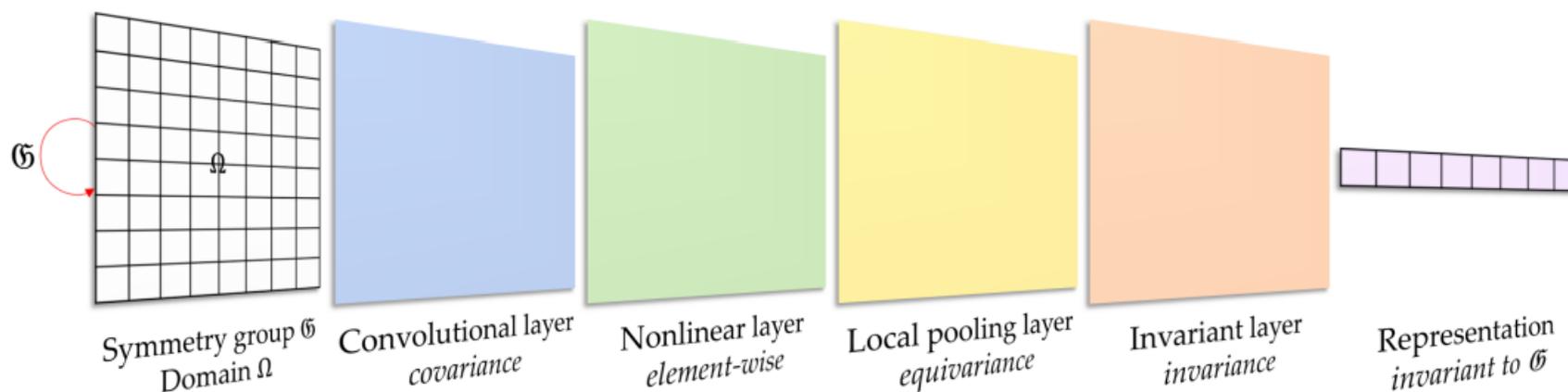
汇报纲要

- 研究背景与意义
- 分析经典不变量
- 完善全局不变量
- 设计局部不变量
- 探索层次不变量**
- 研究总结与展望

探索层次不变量

- 将全局和局部不变量的基本理论统一推广到层次假设：
- 首先，以对称性和不变性视角反思了 CNN 的几种典型模块，并形式化了这些模块在层次不变性中的对称性原则。

方法	经典不变量 ^[16]	经典卷积网络 ^[21]	散射网络 ^[321]	等变网络 ^[322]	本章方法
判别性		✓	✓	✓	✓
稳健性	✓		✓	✓	✓
可解释性	✓		✓	✓	✓
高效性	✓				✓



探索层次不变量

- 随后，重新定义了新的表征模块以满足层次不变性蓝图，群论分析证明其在任意中间层都表现出连续且直接的几何对称性。

性质 5.1: 平移、旋转和翻转联合等变性。对表征单元 $\mathbf{U} \triangleq \mathbf{P} \circ \mathbf{S} \circ \mathbf{C}$ ，具有任意的卷积层参数 λ ，其任意次级联都满足平移、旋转和翻转联合等变性（忽略边际效应和采样近似），即下列恒等式：

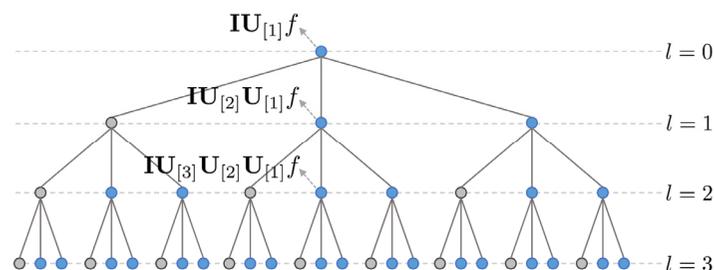
$$\mathbf{U}_{[L]} \circ \cdots \circ \mathbf{U}_{[2]} \circ \mathbf{U}_{[1]}(\mathbf{g}_1 M) \equiv \mathbf{g}_1 \mathbf{U}_{[L]} \circ \cdots \circ \mathbf{U}_{[2]} \circ \mathbf{U}_{[1]}(M). \quad (5.16)$$

对于任意级联长度 $L \geq 1$ 、任意 $\mathbf{g}_1 \in \mathfrak{G}_1$ 和 $M \in X$ 均成立，其中 \mathfrak{G}_1 为平移、旋转和翻转对称群。

性质 5.2: 缩放协变性。对表征单元 \mathbf{U} ，我们指定其卷积层的尺度参数为 w 并记为 $\mathbf{U}^w \triangleq \mathbf{P} \circ \mathbf{S} \circ \mathbf{C}^w$ ，其任意次级联都满足缩放协变性（忽略边际效应和采样近似），即下列恒等式：

$$\begin{aligned} & \mathbf{U}_{[L]}^w \circ \cdots \circ \mathbf{U}_{[2]}^w \circ \mathbf{U}_{[1]}^w(\mathbf{g}_2 M) \\ & \equiv \mathbf{g}'_2 \mathbf{U}_{[L]}^w \circ \cdots \circ \mathbf{U}_{[2]}^w \circ \mathbf{U}_{[1]}^w(M) \\ & \equiv \mathbf{g}_2 \mathbf{U}_{[L]}^{ws} \circ \cdots \circ \mathbf{U}_{[2]}^{ws} \circ \mathbf{U}_{[1]}^{ws}(M), \end{aligned} \quad (5.18)$$

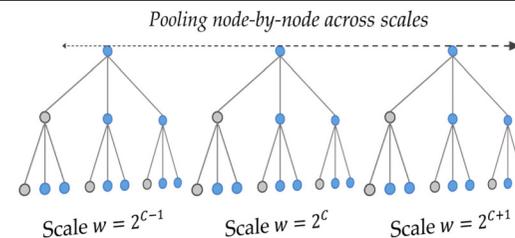
对于任意级联长度 $L \geq 1$ 、任意 $\mathbf{g}_2 \in \mathfrak{G}_2$ 和 $M \in X$ 成立，其中 \mathbf{g}'_2 为可预测操作，与 \mathbf{g}_2 相关，可重写为显式形式 $\mathbf{g}'_2 \mathbf{U}^w \triangleq \mathbf{g}_2 \mathbf{U}^{ws}$ ， s 为对应于 \mathbf{g}_2 的图像缩放因子， \mathfrak{G}_2 为缩放对称群。



性质 5.3: 层次不变性。对表征单元 \mathbf{U} 的任意次级联，设计感兴趣对称群 $\mathfrak{G}_0 \subseteq \mathfrak{G}_1 \times \mathfrak{G}_2$ 的全局不变映射 \mathbf{I} 在实践中是可行的，由于性质 5.1 和 5.2 保证的输入图像和深度表征（任意层）之间可预测的几何对称性关系。更具体的说，基于定义 5.4，我们假设存在 \mathbf{I} 使得 $\mathbf{I}(\mathbf{g}'_0 M) = \mathbf{I}(M)$ 对于任意 $\mathbf{g}'_0 \in \mathfrak{G}_0$ 和 $M \in X$ 成立，其中 \mathbf{g}'_0 为相对于 \mathbf{g} 和 \mathbf{U} 的可预测操作，即不变性在单层上成立，则有不变性：

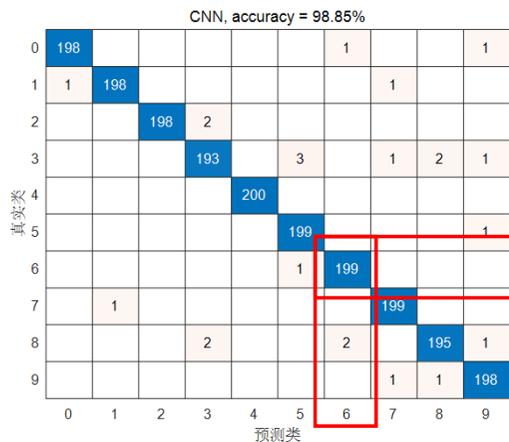
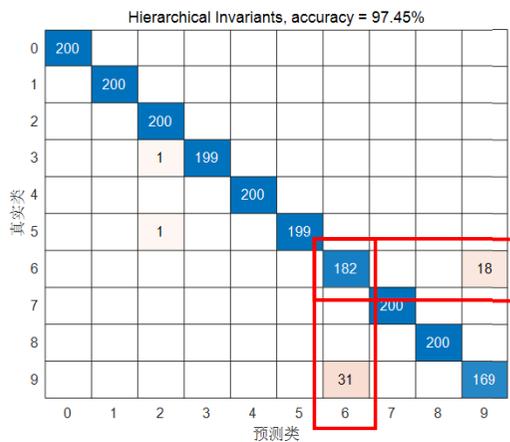
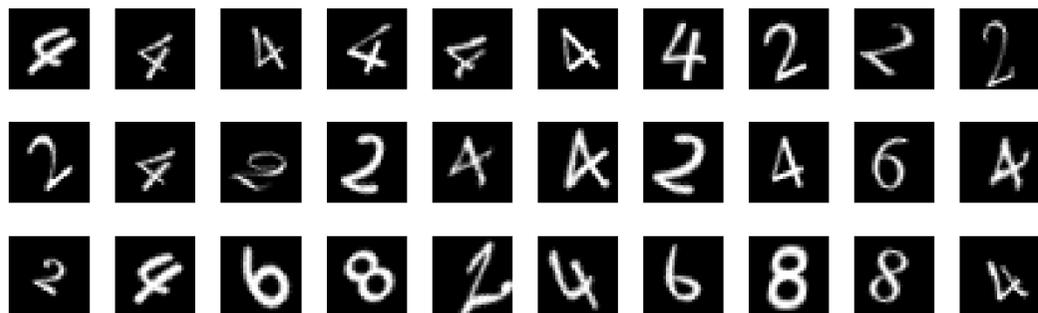
$$\mathbf{I}(\mathbf{g}'_0 M)_{[L]} \equiv \mathbf{I}M_{[L]}, \quad (5.20)$$

对于任意级联长度 $L \geq 1$ 成立。



探索层次不变量

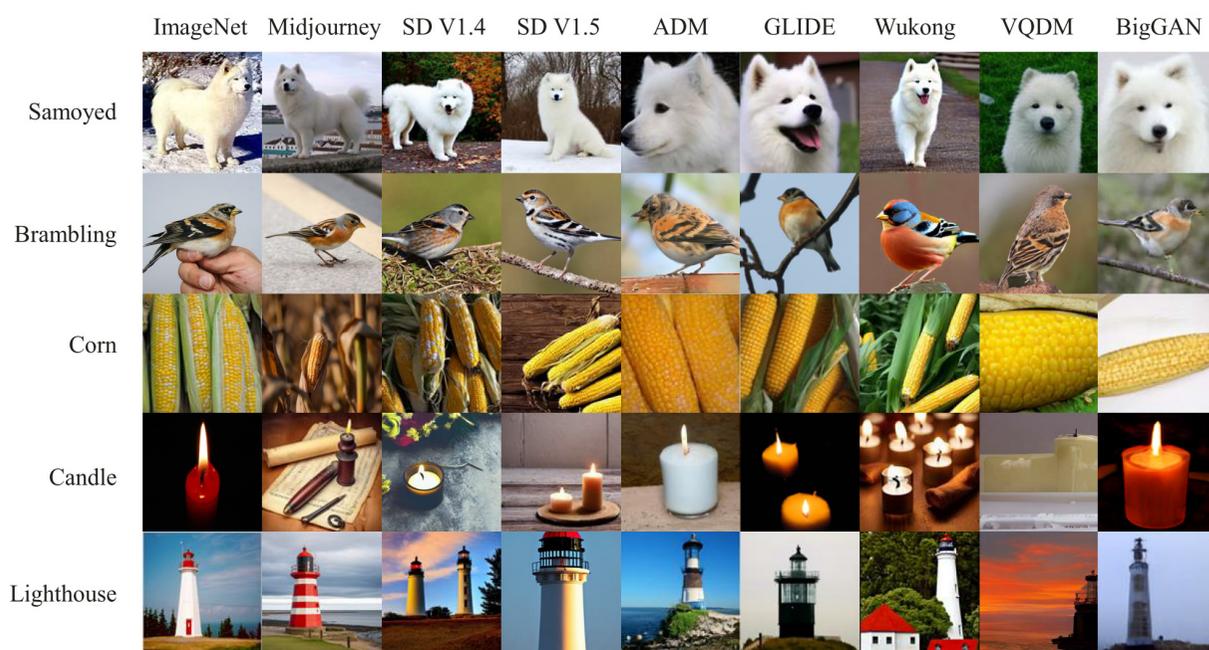
- 与现有的经典不变量、学习表征网络、不变表征网络相比，其在更大规模的经典视觉任务以及机器生成取证任务中能够更好地权衡不变性和判别性。



表征方法	时间 GPU†	直接测试			随机平移和旋转变体测试		
		Precision	Recall	F1	Precision	Recall	F1
经典表征							
Cosine	15	45.68	45.35	45.43	32.50	31.20	30.83
Wavelet	16	67.11	66.75	66.75	38.31	35.25	35.44
Krawtchouk	15	71.73	69.85	69.69	27.05	26.30	25.83
学习表征							
SimpleNet	535†	98.60	98.60	98.60	35.42	33.50	33.72
SimpleNet+	551†	52.70	48.90	48.82	54.26	50.75	50.78
AlexNet	393†	100	100	100	66.18	64.45	64.27
AlexNet+	392†	93.07	92.00	91.80	94.23	93.10	92.93
VGGNet	3610†	100	100	100	70.74	70.25	69.93
VGGNet+	7731†	95.98	95.70	95.68	95.53	95.20	95.13
不变表征							
ScatterNet	115	98.96	98.95	98.95	57.20	56.95	56.23
HIR	57	97.48	97.45	97.45	95.05	94.95	94.98

探索层次不变量

- 与现有的经典不变量、学习表征网络、不变表征网络相比，其在更大规模的经典视觉任务以及机器生成取证任务中能够更好地权衡不变性和判别性。



AIGC图像取证任务
 在小样本、多类样本混合检测场景
 以“即插即用”的方式实现94.73%的F1分数

表征方法	训练/测试 = 5/5			训练/测试 = 1/9		
	Precision	Recall	F1	Precision	Recall	F1
经典表征						
Cosine NN	0.00	0.00	0.00	0.00	0.00	0.00
Cosine SVM	94.95	94.57	94.76	94.36	91.06	92.68
Wavelet NN	48.70	94.17	64.20	48.69	94.13	64.18
Wavelet SVM	94.03	94.57	94.30	83.55	93.48	88.24
Krawtchouk NN	0.00	0.00	0.00	0.00	0.00	0.00
Krawtchouk SVM	75.24	74.77	75.00	71.56	68.57	70.03
学习表征						
SimpleNet	61.79	40.70	49.08	56.40	60.48	58.37
AlexNet	80.76	77.63	79.16	71.83	72.50	72.17
VGGNet	84.75	86.67	85.70	72.45	72.37	72.41
GoogLeNet	74.15	80.40	77.15	67.84	68.83	68.33
ResNet	85.10	83.03	84.06	76.88	73.67	75.24
DenseNet	86.83	85.23	86.02	76.84	75.37	76.10
InceptionNet	82.69	86.63	84.62	68.62	68.56	68.59
MobileNet	81.54	82.47	82.00	68.52	68.57	68.55
不变表征						
ScatterNet NN	83.68	83.73	83.71	79.37	79.70	79.53
ScatterNet SVM	90.31	85.17	87.67	85.28	79.70	82.40
HIR NN	96.79	96.47	96.63	95.66	93.04	94.33
HIR SVM	96.92	96.37	96.64	95.21	94.26	94.73

汇报纲要

- 研究背景与意义
- 分析经典不变量
- 完善全局不变量
- 设计局部不变量
- 探索层次不变量
- 研究总结与展望**

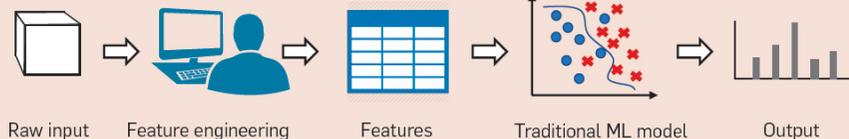
总结与展望

研究目标

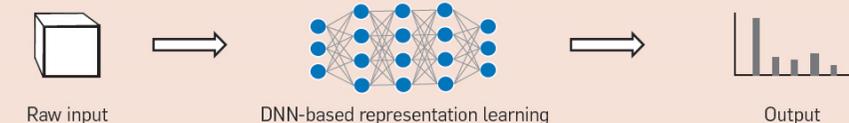
受数学和物理学的“爱尔兰根纲领”启发
试图在计算机视觉系统中嵌入物理世界的**不变性和对称性先验**

研究难点

知识驱动路径



数据驱动路径



如何同时
兼顾二者的优势
摒弃二者的劣势

研究总结

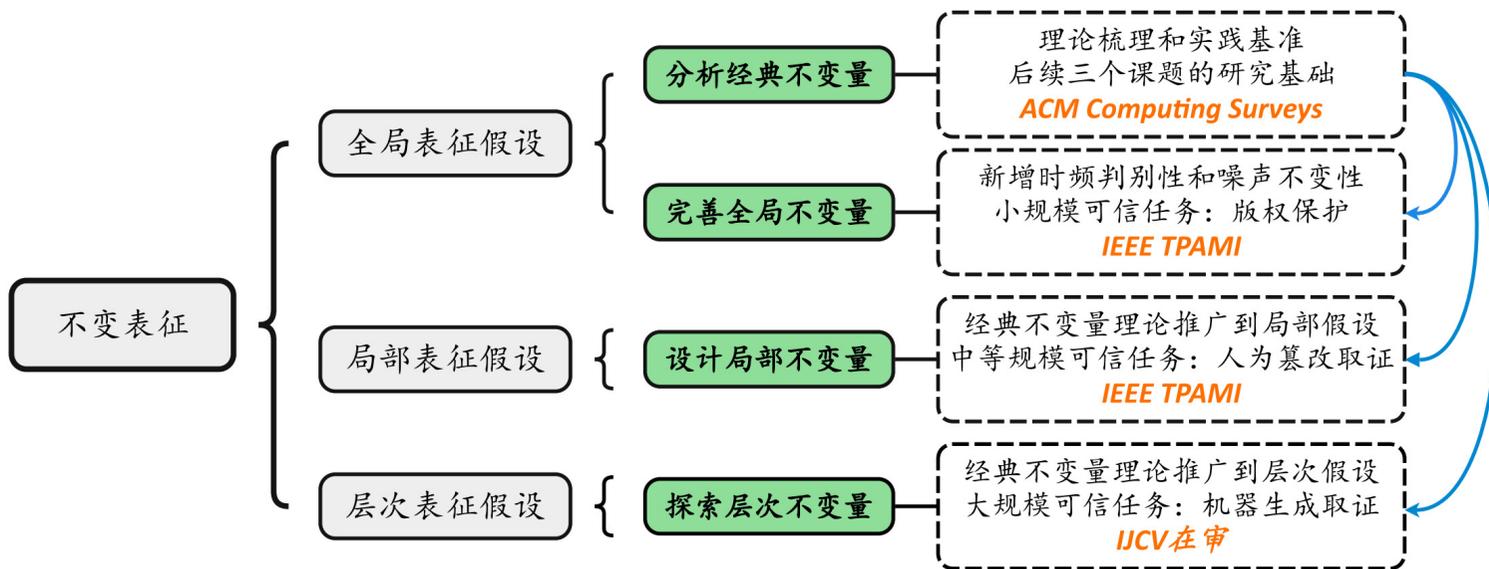
在**理论、实践和应用**层面的扩展了经典不变量，聚焦于增强**判别性**
分别发展出**全局**不变量、**局部**不变量、**层次**不变量

主要学术论文

1. Shuren Qi, Yushu Zhang, Chao Wang, Jiantao Zhou, and Xiaochun Cao. A survey of orthogonal moments for image representation: Theory, implementation, and evaluation. ACM Computing Surveys ([CSUR](#)), 2023, 55(1): 1-35. [ESI高被引论文](#)
2. Shuren Qi, Yushu Zhang, Chao Wang, Tao Xiang, Xiaochun Cao, and Yong Xiang. Representing noisy image without denoising. IEEE Transactions on Pattern Analysis and Machine Intelligence ([TPAMI](#)), 2024, Early Access. [CCF-A](#)
3. Shuren Qi, Yushu Zhang, Chao Wang, Jiantao Zhou, and Xiaochun Cao. A principled design of image representation: Towards forensic tasks. IEEE Transactions on Pattern Analysis and Machine Intelligence ([TPAMI](#)), 2023, 45(5): 5337-5354. [CCF-A](#)
4. Shuren Qi, Yushu Zhang, Chao Wang, Zhihua Xia, Xiaochun Cao, Jian Weng, and Feng-Lei Fan. Hierarchical invariance for robust and interpretable vision tasks at larger scales. International Journal of Computer Vision ([IJCV在审](#)). [CCF-A](#)



感谢观看，恳请批评指正！



by Shuren Qi

shurenqi@nuaa.edu.cn | shurenqi.github.io