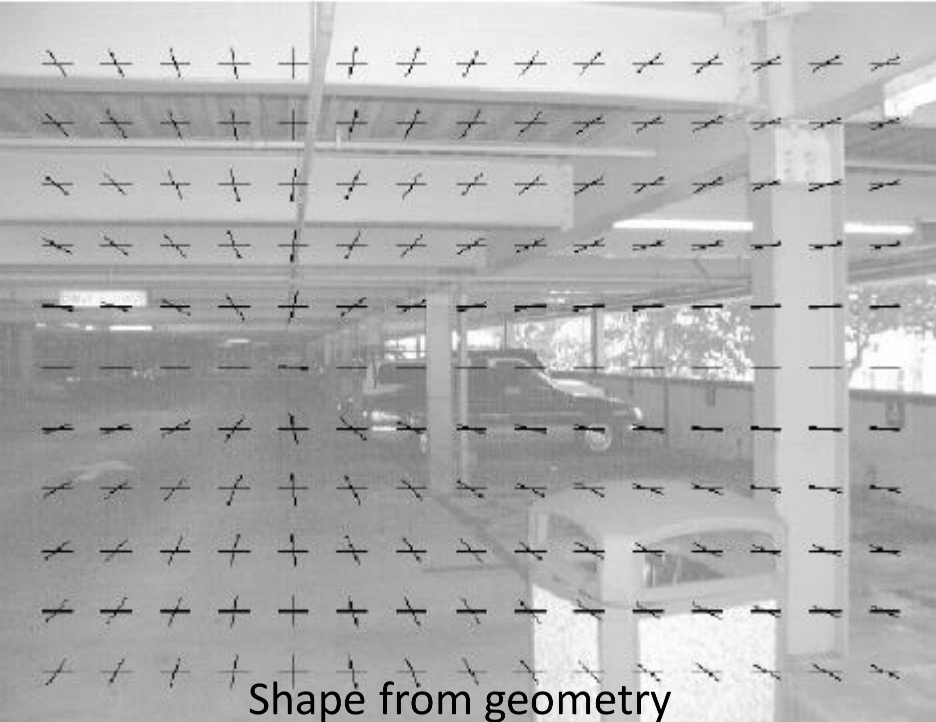
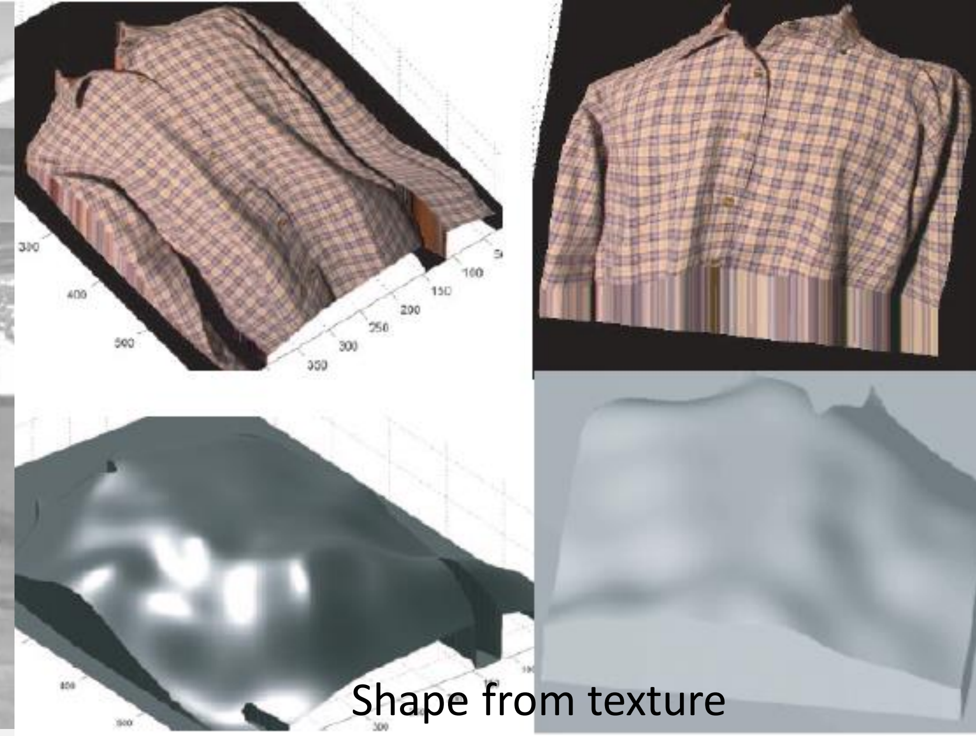
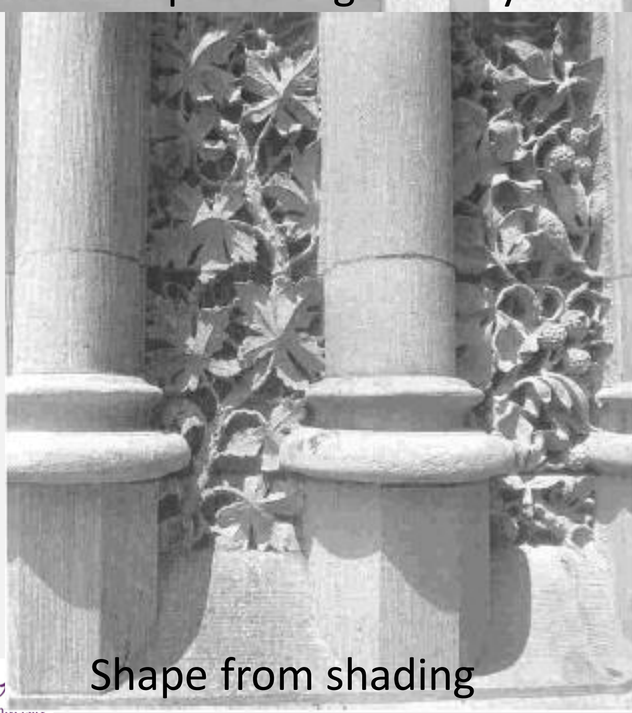
Shape from Texture



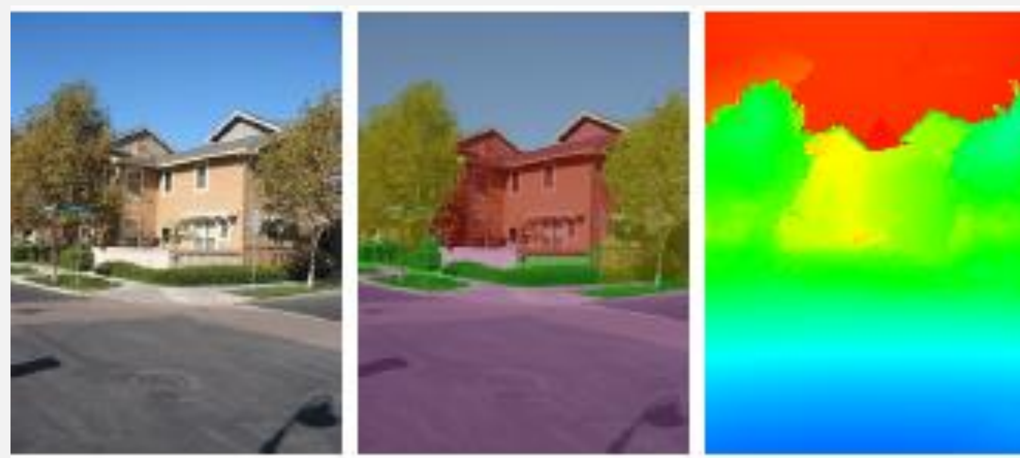
Shape from geometry



Shape from texture



Shape from shading

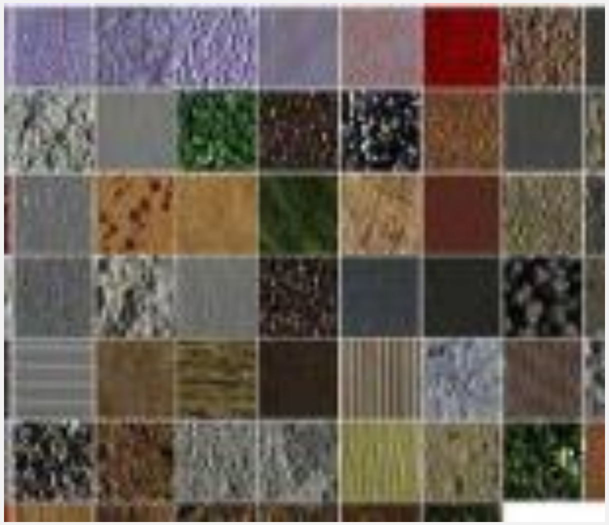


Learning based

Shape from Texture

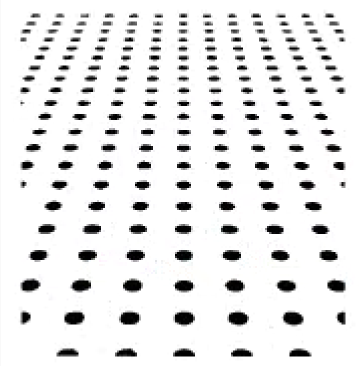
- 从**纹理**图案恢复深度信息
  - 表面光滑，深度渐变 smooth
  - 纹理元素均匀分布 homogeneous
  - 纹理元素之间没有关联 isotropic

LBP

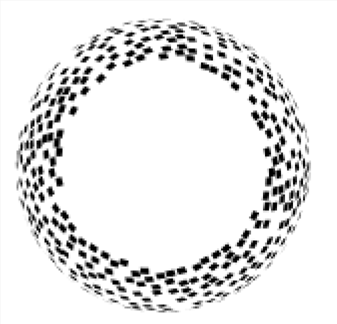


Shape from Texture

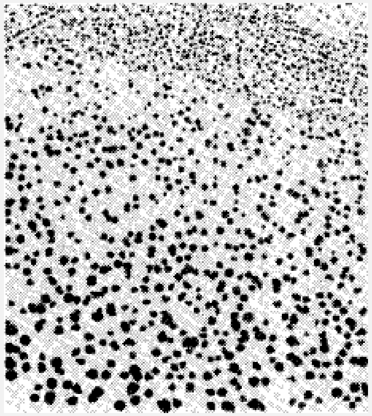
- 最简单的纹理Periodic—频谱法



- 一般纹理Cyclostationary—texton法

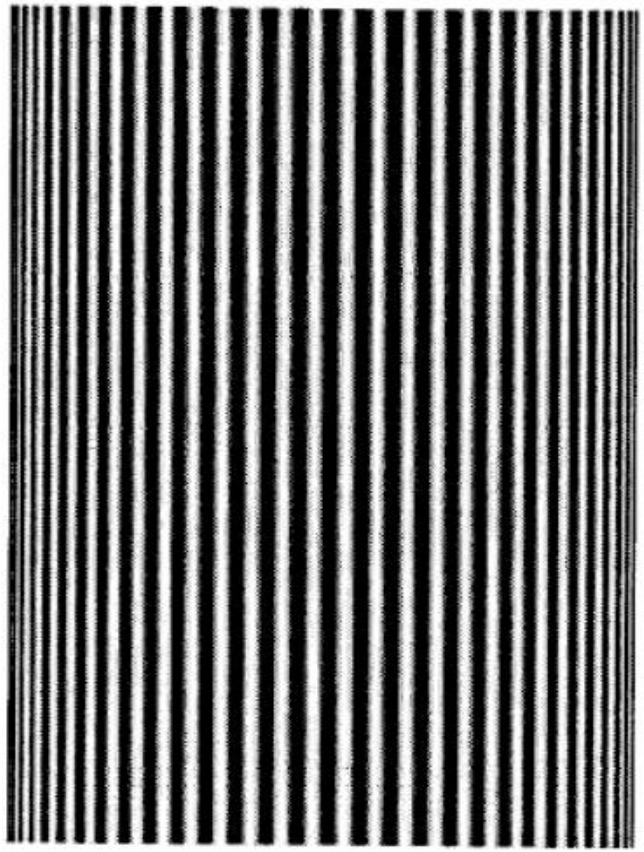


- 复杂纹理

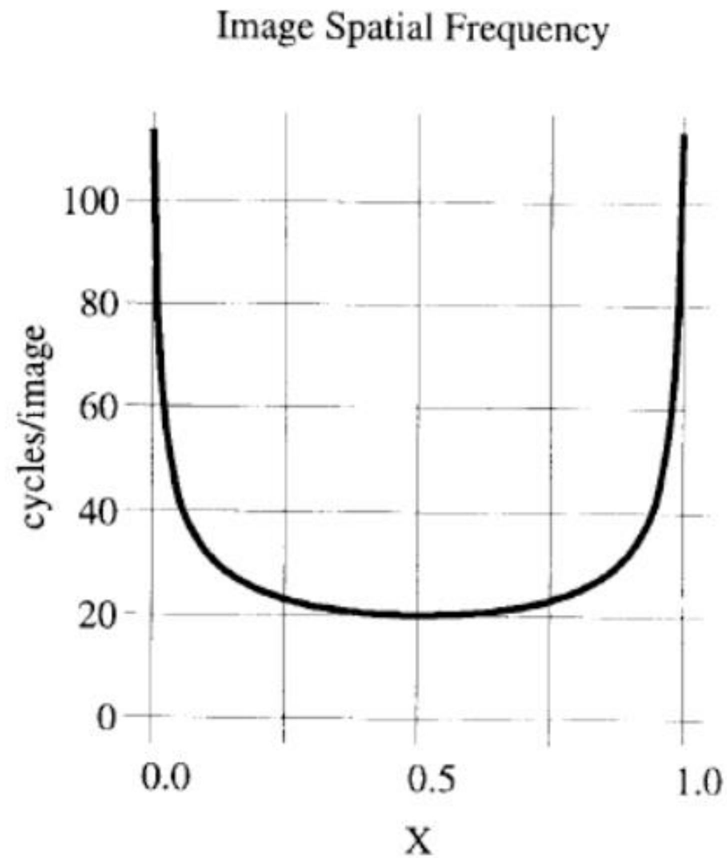


Texture description

Use filter outputs to measure local spatial frequency.



(a)



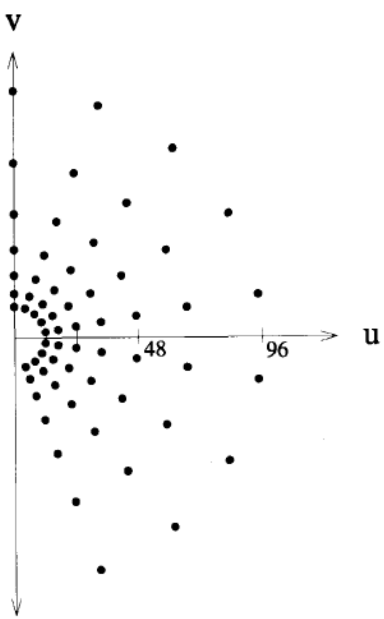
(b)

Fig. 2. (a) Cylinder with sinusoidal grating texture. (b) Horizontal component of image spatial frequency on center cross-section of (a).

Texture projection

Assume orthographic projection.

Gabor滤波响应



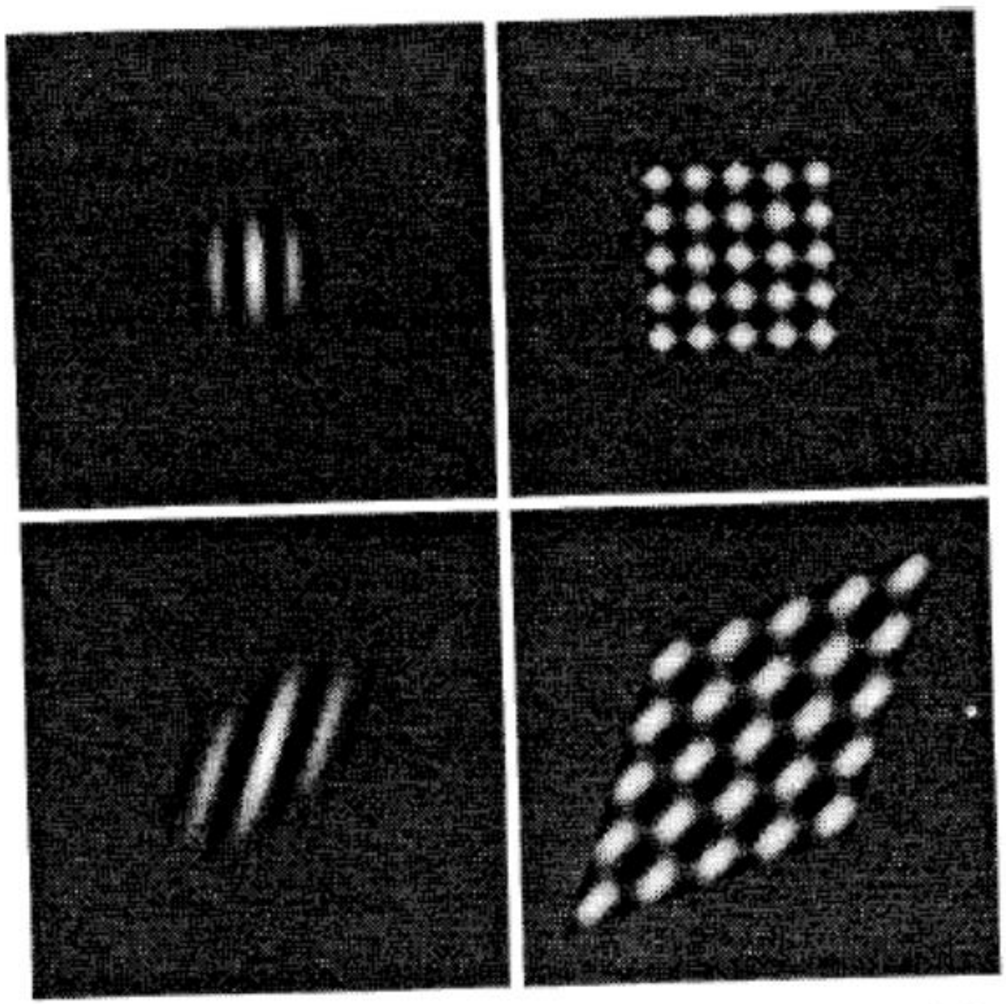
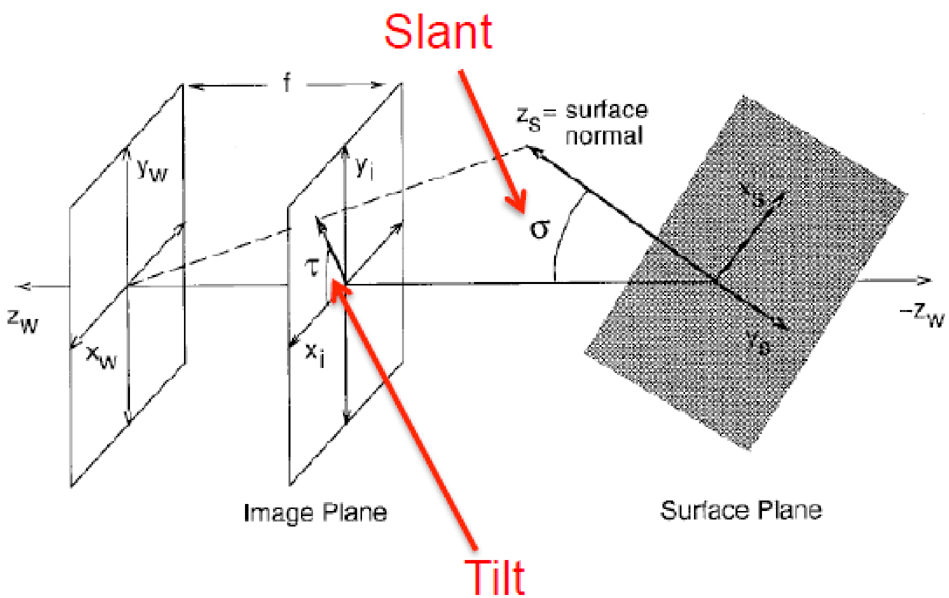
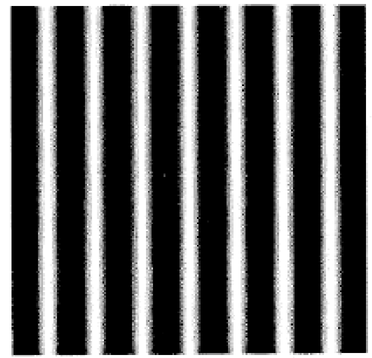


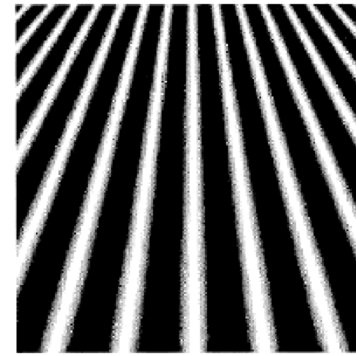
Fig. 5. Top row: real part of Gabor filter with radial frequency of 12 cycles/image, and a texture patch. Bottom row: back-projections of Gabor filter and texture patch onto a plane with orientation $(\sigma, \tau) = (60^\circ, 45^\circ)$.

Slant and tilt

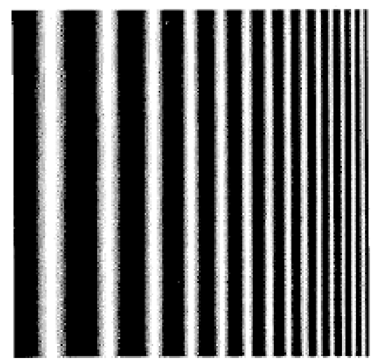




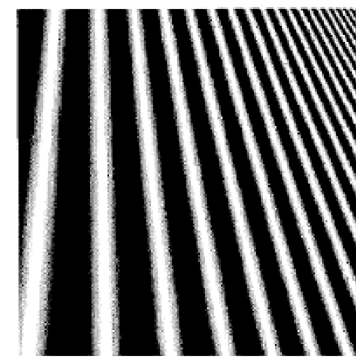
$\sigma=0°$



$\sigma=45°$，$\tau=90°$



$\sigma=45°$，$\tau=0°$



$\sigma=45°$，$\tau=45°$

由Gabor滤波结果确定倾角

实际算法考虑到噪声等干扰，采用二阶矩计算

$$a(\boldsymbol{x}) = \sum_i u_i^2(\boldsymbol{x})A_i^2(\boldsymbol{x})$$

$$b(\boldsymbol{x}) = 2\sum_i u_i(\boldsymbol{x})v_i(\boldsymbol{x})A_i^2(\boldsymbol{x})$$

$$c(\boldsymbol{x}) = \sum_i v_i^2(\boldsymbol{x})A_i^2(\boldsymbol{x}),$$

$$M = \frac{1}{2}\left(a+c+\sqrt{b^2+(a-c)^2}\right)$$

$$m = \frac{1}{2}\left(a+c-\sqrt{b^2+(a-c)^2}\right)$$

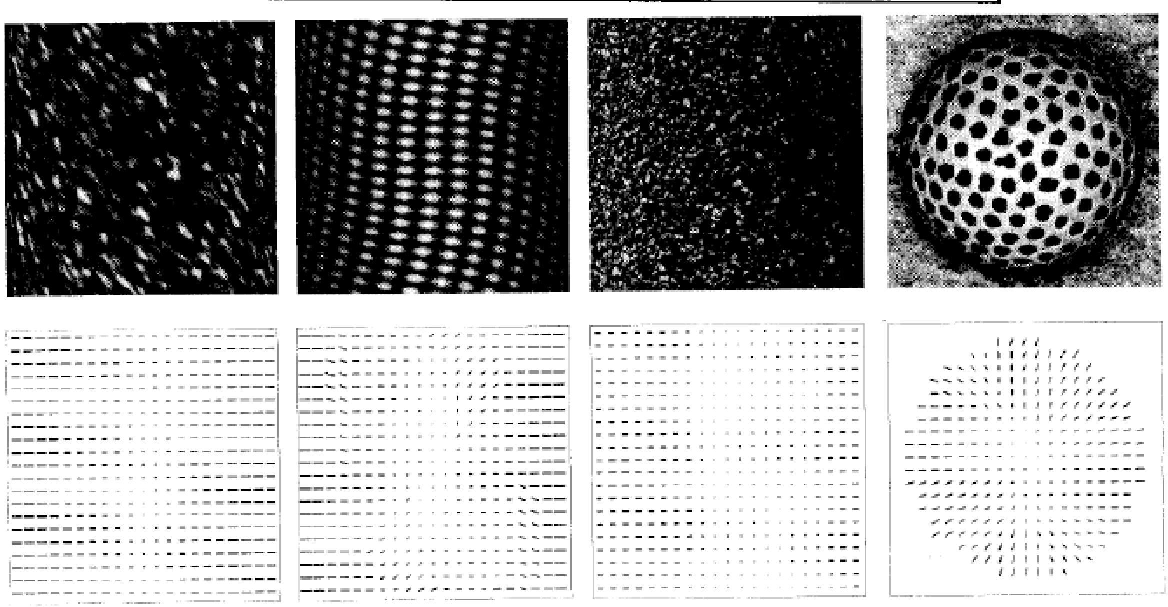
$$\theta = \frac{1}{2}\arctan\frac{b}{a-c}.$$

$$\cos\sigma = \sqrt{\frac{4a_s c_s - b_s^2}{4ac-b^2}} = \sqrt{\frac{M_s m_s}{Mm}}.$$

$$\tau = \left\{\theta \pm \frac{1}{2}\arccos\lambda,\ \theta \pm \frac{1}{2}\arccos\lambda + \pi\right\} \quad \lambda = \frac{(\cos^2\sigma+1)(M+m)-2(M_s+m_s)}{\sin^2\sigma(M-m)}$$

Box 1. Summary of algorithm

1. Convolve the image with Gabor functions and their partial derivatives, and smooth the filter output amplitudes (to reduce noise) by convolving them with a Gaussian.
2. Select the Gabor filter h_k with the largest amplitude output at each point.
3. Compute the (signed) instantaneous frequency $\mathbf{u}_i(\mathbf{x}_i)$ at each point using equation (6).
4. Sample (σ, τ)-space, backprojecting $\mathbf{u}_i(\mathbf{x}_i)$ to compute $\mathbf{u}_s(\mathbf{x}_s)$ using equation (20). Compute the variance $V_{\sigma,\tau}$ of $\mathbf{u}_s(\mathbf{x}_s)$. Coarse-to-fine sampling in multiple stages may be used.
5. Output the values of (σ, τ) for which $V_{\sigma,\tau}$ is a minimum.



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

Texture description

Non-occluded textons, and approximated as flat.

The two pieces of the solution

每个texton有三个旋转自由度

If we knew the transformations

- We can find the textons
- We can find the local intensity contrast

By minimization of:

$$\Sigma_i \; || \; \lambda_i \mathcal{I}_\mu - \mathcal{I}_i \; ||^2$$

Local contrast

Texton

Local texture with inverse transform

If we knew the texton and contrast

- Recover the transformation by transforming the texton to match each local patch.

EM

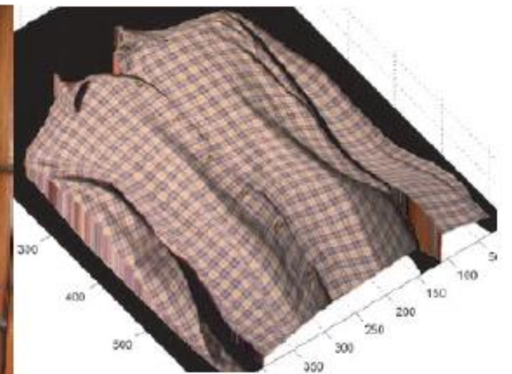
纹理元素 变形

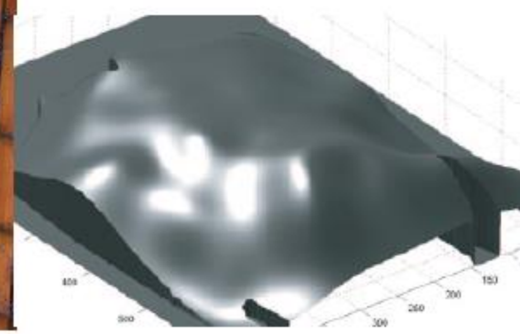
$$\frac{1}{2\sigma_{im}^2}\sum_i \left(\| \lambda_i \mathcal{I}_\mu - \mathcal{T}_i^{-1}\mathcal{I} \|^2 \delta_i\right) + \sum_i (1-\delta_i)K + \frac{1}{2\sigma_{light}^2}(\lambda_i - 1)^2 + L$$

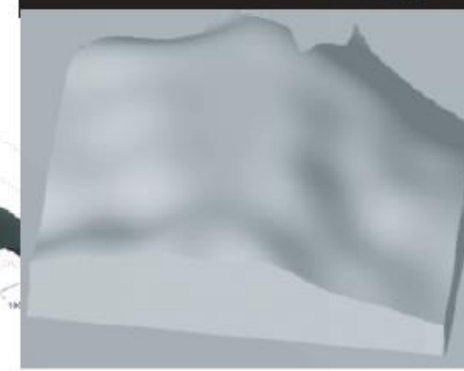
对比度 隐变量

Find interest points

EM iterations

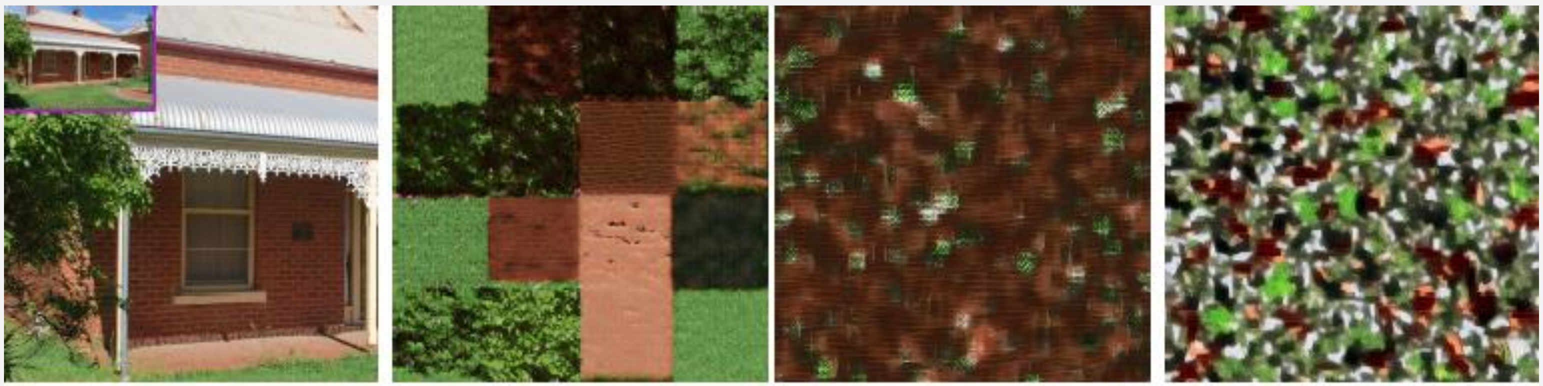




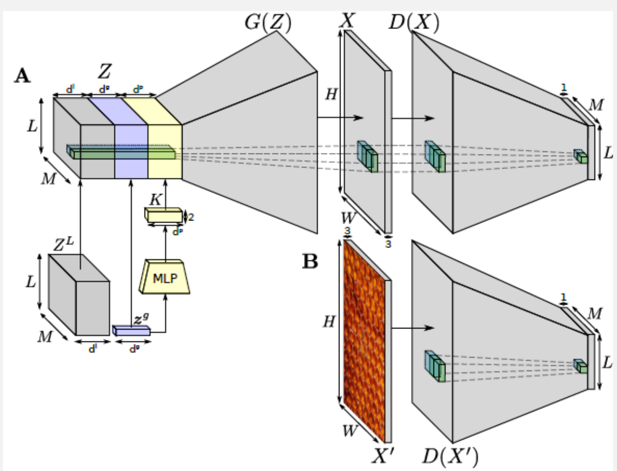


EM→GAN

- Learning Texture Manifolds with the Periodic Spatial GAN, CVPR2017
- GAN生成纹理



Input PSGAN SGAN Gatys et al.



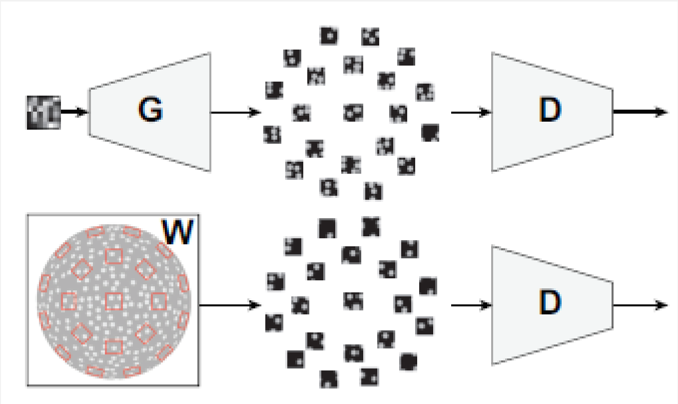
2.3. Spatially Periodic Dimensions

The third part of Z, Z^p, contains spatial periodic functions, or plane waves in each channel i:

$$Z^{p}_{\lambda\mu i} = \zeta_{\lambda\mu i}(K) = \sin\left(\boldsymbol{k}_i^T \begin{pmatrix}\lambda\\ \mu\end{pmatrix} + \phi_i\right), \qquad (2)$$

Model

- Toward a Universal Model for Shape from Texture, CVPR2020
- Generator
- Discriminator
- Warper



- 需要求解
  - 展平后的texture分布
  - 扭曲场

$$
\begin{aligned}
G^{(t)}, D^{(t)} &= \arg\min_{G}\max_{D} \mathbb{E}_{I\sim q(I;W^{(t-1)})}[\log D(I)] \\
&\quad + \mathbb{E}_{Z\sim p(Z)}[\log(1-D(G(Z)))], \\
W^{(t)} &= \arg\max_{W} \mathbb{E}_{I\sim q(I;W)}[\log D^{(t)}(I)] - C(W),
\end{aligned}
$$

Generator

- M+3，减少加窗效应
- 三种随机输入
  - local maps [-1,1] 调节纹理图案，噪声
  - periodic maps, 变频正弦波

$$Z_p^{(i,j)}(\lambda,\mu) = \sin\left(2\pi k_1^{(i,j)}\lambda + 2\pi k_2^{(i,j)}\mu + \phi^{(i,j)}\right),$$
$$k_1^{(i,j)} = a^{(i,j)}\cos(\theta^{(i,j)}),$$
$$k_2^{(i,j)} = a^{(i,j)}\sin(\theta^{(i,j)}),$$

  - global map [-1,1]，全局噪声，比如阴影

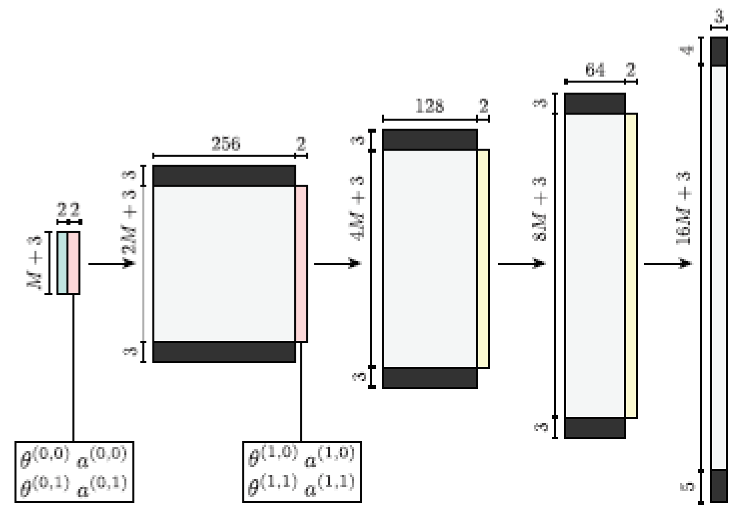


Figure 3: Side view of generator, showing only one spatial dimension (spatial operations are square). Arrows denote 5×5 stride-2 deconvolution followed by a ReLU and a batch normalization layer. Each arrow's output is cropped (darkly shaded) to avoid windowing effects. Adjacent blocks denote concatenation, and colored blocks are input samples from predefined stochastic processes: green are spatially-constant random samples ("global maps"); yellow are spatially i.i.d. random samples ("local maps"); and pink are randomly-shifted sinusoids with learnable frequencies ("periodic maps"). See text for details.

Discriminator

- Symmetrically to the generator
  - four stride-2 convolutions
  - leaky-ReLU activation with slope 0.2
  - use “valid” convolution
- 输出概率

Warper

- 3D normal and tangent vectors

$$\hat{n} = \frac{-p\hat{x} - q\hat{y} + \hat{z}}{\sqrt{p^2 + q^2 + 1}}, \quad \hat{t} = \frac{cn_z\hat{x} + sn_z\hat{y} - (cn_x + sn_y)\hat{z}}{\sqrt{c^2 n_z^2 + s^2 n_z^2 + (cn_x + sn_y)^2}}.$$

- 3×3 rotation $\hat{x} \mapsto \hat{t}$, $\hat{z} \mapsto \hat{n}$, and $\hat{y} \mapsto \hat{n} \times \hat{t} \triangleq \hat{b}$.
- 正交投影 $W(x,y) = \begin{bmatrix} t_x(x,y) & b_x(x,y) \\ t_y(x,y) & b_y(x,y) \end{bmatrix}$

- 连续性约束+光滑性约束
  - Integrability

$$C^{(I)} = \frac{\alpha^{(I)}}{hw} \sum_{i,j} [p_{i,j+1} - p_{i+1,j+1} + p_{i,j} - p_{i+1,j} + q_{i,j+1} + q_{i+1,j+1} - q_{i,j} - q_{i+1,j}]^2,$$

  - Smoothness

$$C^{(S)} = \alpha_n^{(S)} \frac{1}{hw} \|\nabla \hat{n}\|_2^2 + \alpha_t^{(S)} \frac{1}{hw} \|\nabla \hat{t}\|_2^2,$$

- 高斯金字塔多尺度优化

Experiments

- generator output of size 192×192 (M = 12)
- unwarper patch size {96, 128, 160, 192}
- Synthetic images created with Blender

$$\mathrm{MAE}(\hat{n}, \hat{n}^{(gt)}) = \frac{1}{hw}\sum_{i=1}^{h}\sum_{j=1}^{w} |\cos^{-1}(\hat{n}_{ij} \cdot \hat{n}_{ij}^{(gt)})|$$

- 有效恢复形状
  - 只有纹理图难以辨认形状的
  - 没有边缘的
  - 对非周期纹理也有效
- 缺陷
  - 对于接近平坦的恢复成平面
  - 局部极值产生褶皱

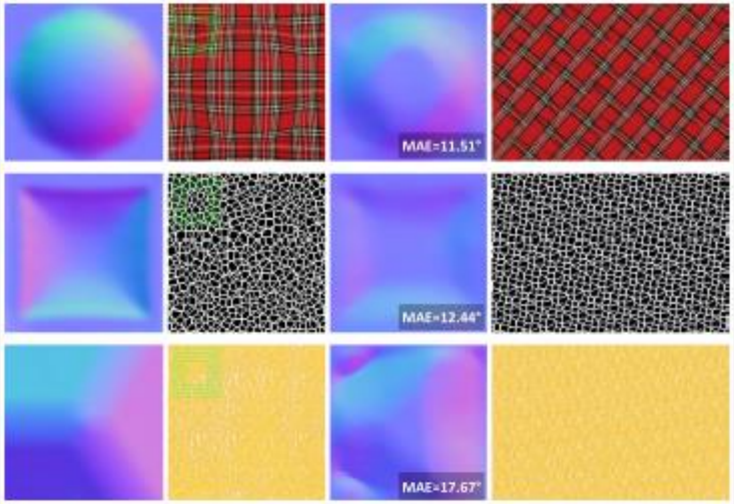


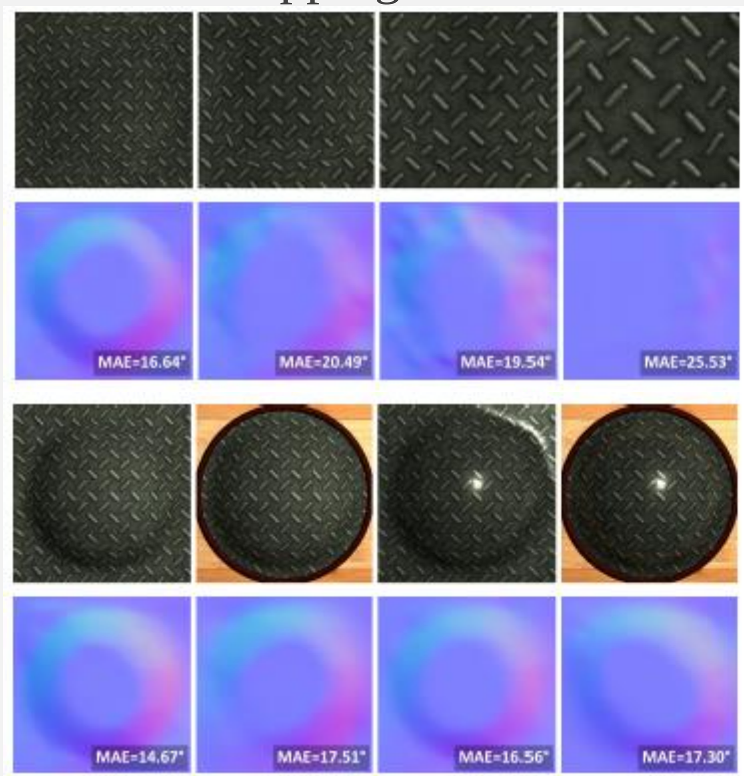
Figure 5: Shape and texture results for synthetic images. From left to right: surface normals of the true shape; input image; output surface normals (with MAE); and sample from output texture generator. Green inset squares show patch sizes used by the unwarper.

Experiments

- isolated textons [16, 17] and stationarity [4]
- Stationarity [4] works well for periodic textures but breaks for general cyclostationary ones.
- Isolated textons [17] works well for non-overlapping texture elements but breaks when they overlap.

[4]	30.5	41.2	45.4	35.6	32.7	37.8	36.1	38.7	29.5	29.0	24.4
[17]	-	-	12.9	35.9	21.5	23.7	-	19.8	22.6	21.6	**6.9**
[16]	27.7	29.8	38.2	40.4	41.7	23.5	33.1	17.6	32.0	47.6	9.2
ours	**15.2**	**16.8**	**12.6**	**15.0**	**18.6**	**17.4**	**19.5**	**17.1**	**14.0**	**20.1**	14.9

Table 1: Shape accuracy (MAE, in degrees) of our and other algorithms for various textures, including those of Figs. 5 & 6 and four additional ones. Each entry is average error over four images with the same texture and four different shapes (see supplement for complete visualizations). Missing entries for [17] are failures to identify a frontal textons.



Conclusion

- 三方对抗
- 可适用各种纹理
- Our results suggest that it is worth considering how multi-player games might be used to address other types of intrinsic image tasks, and how they might be combined with perceptual grouping for higher-level vision tasks in real-world environments.