基于视皮层方位选择特性的模版提取及其在质量评价中的应用

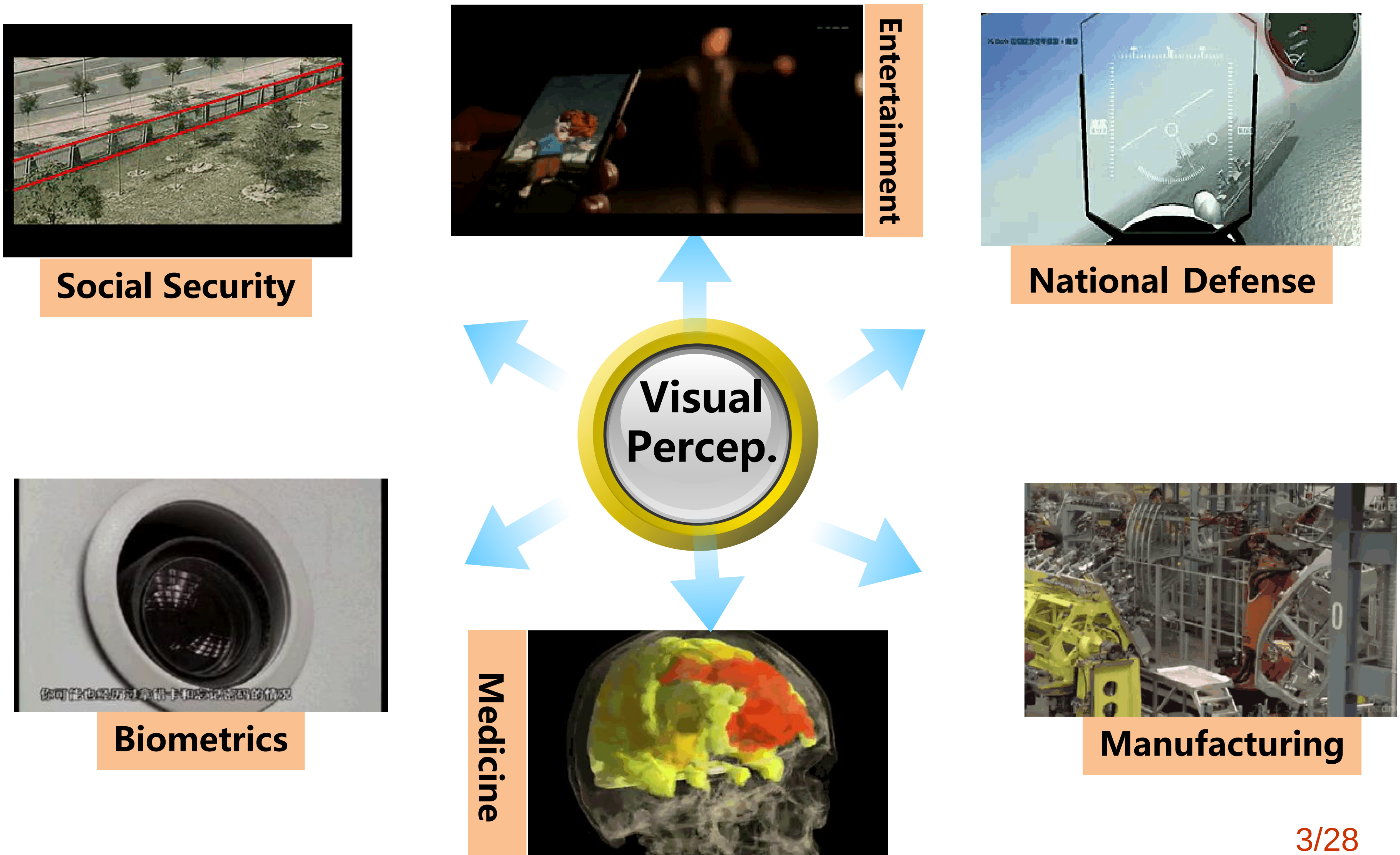
吴金建
西安电子科技大学
http://web.xidian.edu.cn/wjj/index.html

Outline

- Background: Local Structural Descriptors
- The Orientation Selectivity Mechanism
- The Visual Pattern Modeling
- Image Content Extraction
- Application 1: Texture Classification
- Application 2: Quality Assessment
- Conclusion

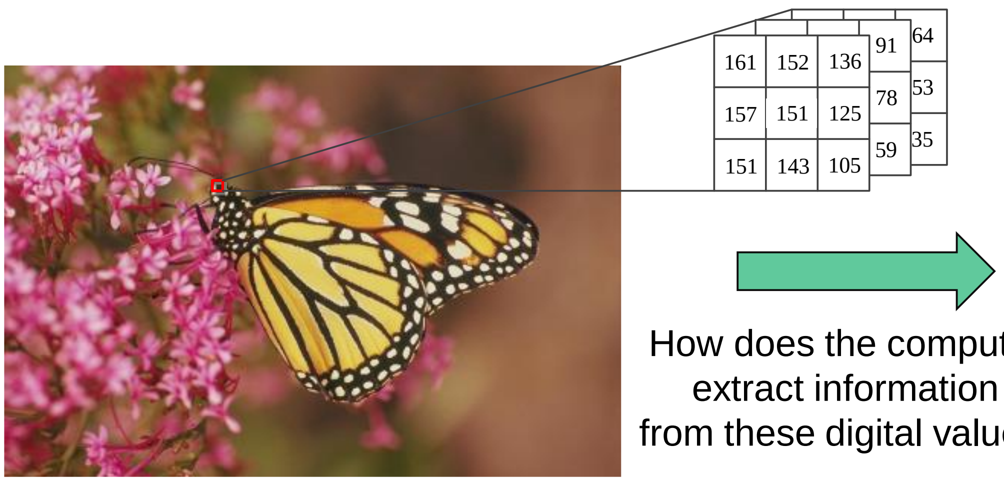
Background: Local Structural Descriptors

Video/Image is everywhere (more than 80% info. that we received)



Background: Local Structural Descriptors

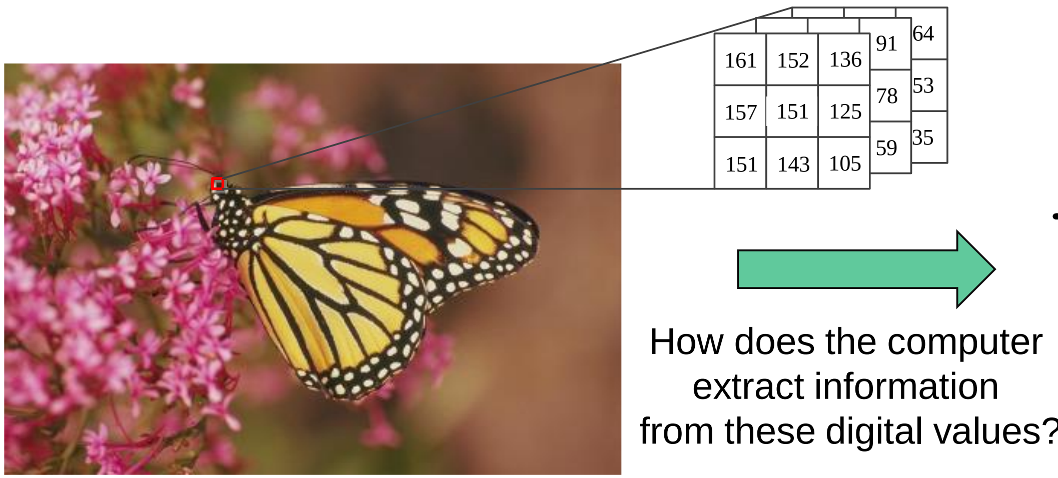
How does the computer understand the digital images?



How does the computer
extract information
from these digital values?

Background: Local Structural Descriptors

How does the computer understand the digital images?



- Global descriptor:
  - Color
  - Entropy
  - …
- Regional descriptor:
  - Edge
  - Contour
  - …
- Local descriptor:
  - Corner (SIFT)
  - Texton (LBP)
  - …

Background: Local Structural Descriptors

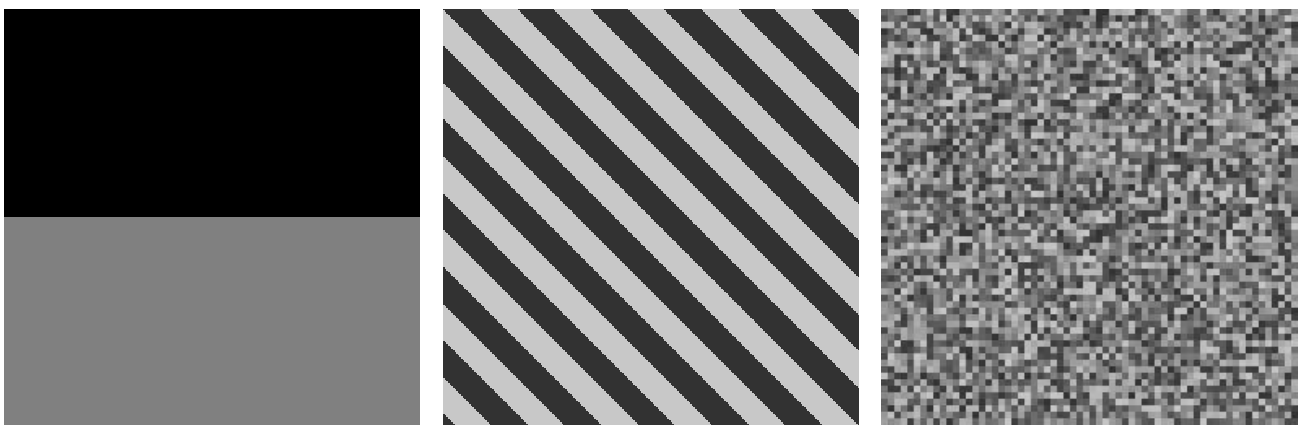
Local structural descriptor:

- The human visual system (HVS) is extremely adaptive to extract structures for image perception and understanding.
- The HVS is sensitive to luminance changes:
  - First order statistic values, e.g., mean, variance, etc.
  - The gray-level difference (luminance contrast)…

Background: Local Structural Descriptors

Local structural descriptor:

- The human visual system (HVS) is extremely adaptive to extract structures for image perception and understanding.
- The HVS is sensitive to luminance changes:
  - First order statistic values, e.g., mean, variance, etc.
  - The gray-level difference (luminance contrast)…

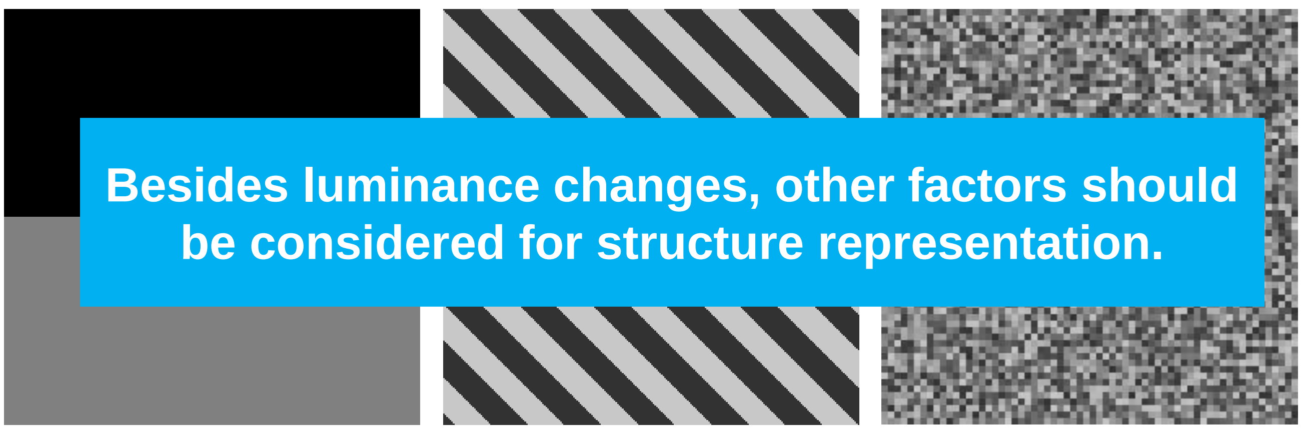


Similar luminance changes, while quite different structure.

Background: Local Structural Descriptors

Local structural descriptor:

- The human visual system (HVS) is extremely adaptive to extract structures for image perception and understanding.
- The HVS is sensitive to luminance changes:
  - First order statistic values, e.g., mean, variance, etc.
  - The gray-level difference (luminance contrast)…

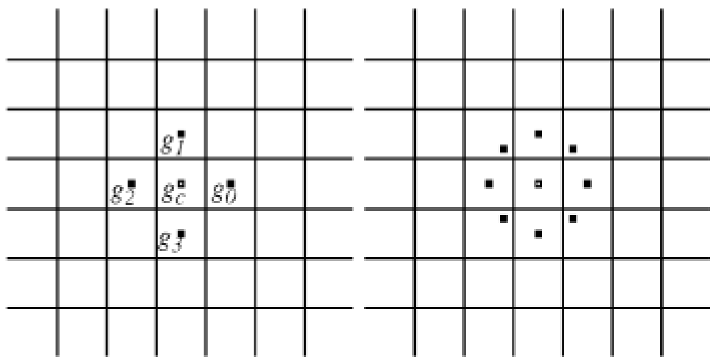


Similar luminance changes, while quite different structure.

Background: Local Structural Descriptors

Spatial distribution within neighbor pixels:

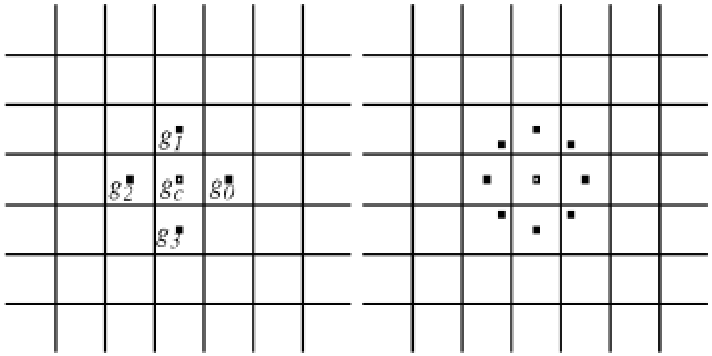
- The spatial joint distribution is analyzed with the co-occurring pixel values from neighborhood;
- By using the signed gray-level differences, the famous local binary pattern (LBP) is introduced.



Background: Local Structural Descriptors

Spatial distribution within neighbor pixels:

- The spatial joint distribution is analyzed with the co-occurring pixel values from neighborhood;
- By using the signed gray-level differences, the famous local binary pattern (LBP) is introduced.

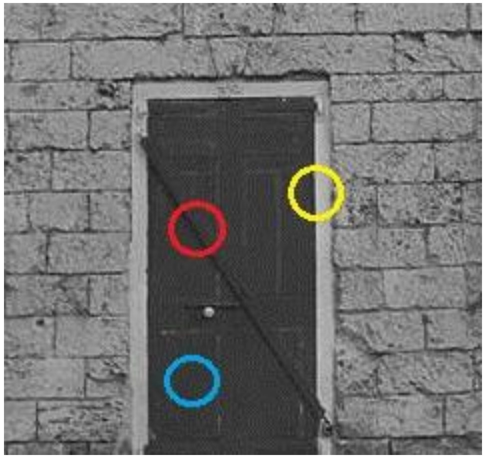


- **The LBP succeeds in describing the spatial correlation of the structure.**
- **However, the signed gray-level difference procedure is too sensitive to disturbance.**

Background: Local Structural Descriptors

Visual orientation:

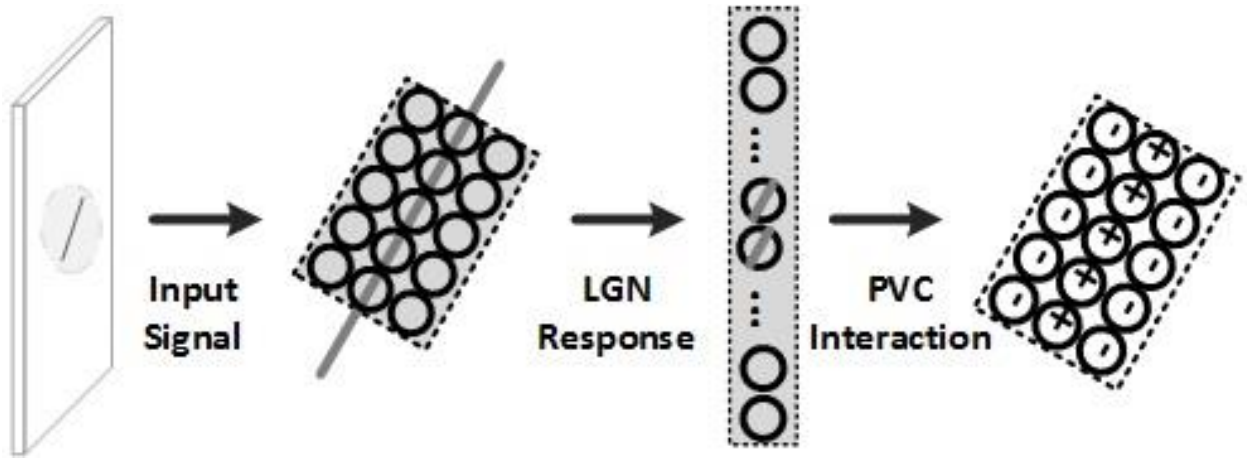
- The HVS is sensitivity to visual orientation.
- Orientation information has been widely used in structural descriptor:
  - SIFT, HOG, etc.
- The HVS exhibits substantial orientation selectivity for scene perception.



The Orientation Selectivity Mechanism

The OS mechanism in the primary visual cortex :

- The HVS responds preferentially to the edge regions of an input scene;
- Because neurons in the primary visual cortex (especially in layer 4) exhibit substantial orientation selectivity

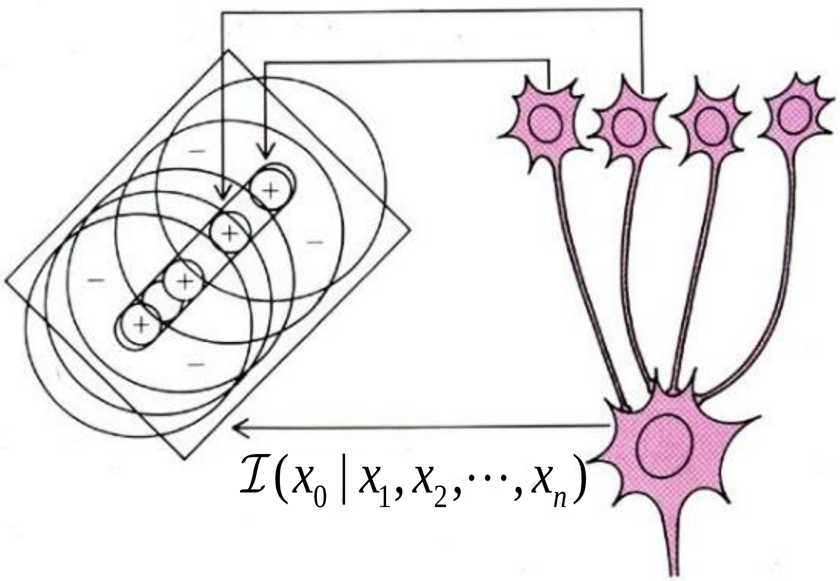


The Orientation Selectivity Mechanism

The OS mechanism in the primary visual cortex :

- Orientation selectivity arises from the arrangement of neuron interaction;

$$\mathcal{A}(\mathcal{I}(x\,|\,\mathcal{X})) = \mathcal{A}(\mathcal{I}(x\,|\,x_1, x_2, \cdots, x_n))$$



The Orientation Selectivity Mechanism

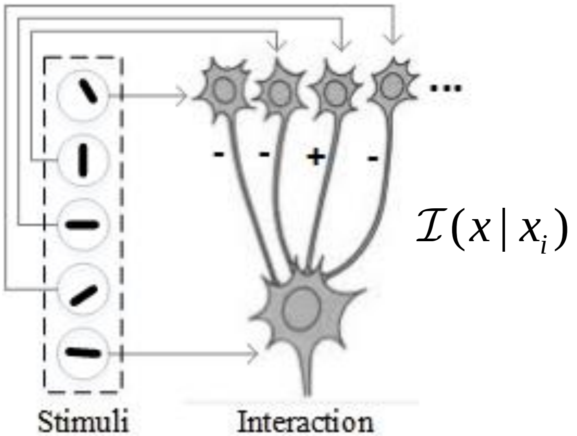
The OS mechanism in the primary visual cortex :

- Orientation selectivity arises from the arrangement of neuron interaction;

$$\mathcal{A}(\mathcal{I}(x|\mathcal{X}))=\mathcal{A}(\mathcal{I}(x|x_1,x_2,\cdots,x_n))$$

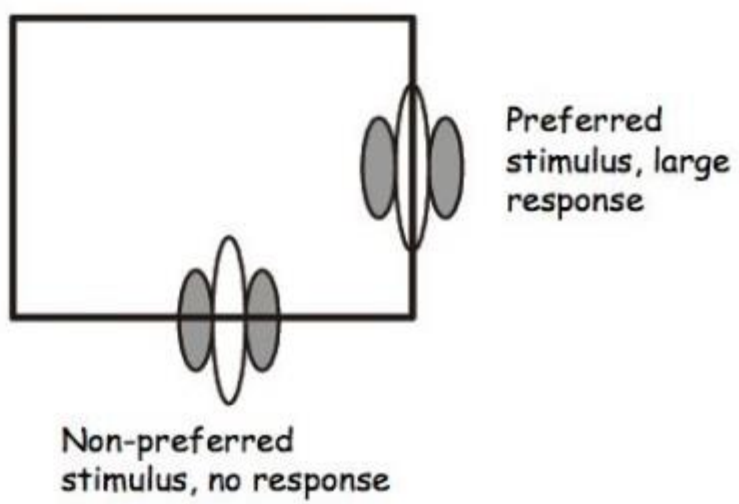
- For simplicity, only consider the synapses between the central cell and its excited cells;

$$\mathcal{A}(\mathcal{I}(x|\mathcal{X}))\approx\mathcal{A}(\mathcal{I}(x|x_1),\mathcal{I}(x|x_2),\cdots,\mathcal{I}(x|x_n))$$



The OS modeling : Interaction

- The interaction between two cortical neurons:
  - **Excitation** if with similar preferred stimuli/orientations;
  - **Inhibition** if with dissimilar preferred stimuli/orientations.

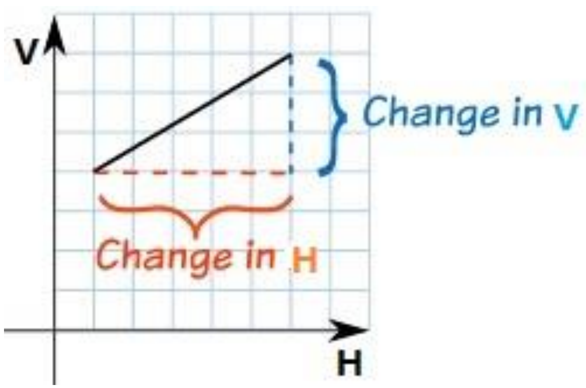


The OS based Visual Pattern Modeling

The OS modeling : Interaction

- The interaction between two cortical neurons:
  - **Excitation** if with similar preferred stimuli/orientations;
  - **Inhibition** if with dissimilar preferred stimuli/orientations.
- Preferred orientation of each pixel in an image:

$$\theta(x) = \arctan \frac{\mathcal{L}_v(x)}{\mathcal{L}_h(x)}$$



The OS modeling : Interaction

- The interaction between two cortical neurons:
  - **Excitation** if with similar preferred stimuli/orientations;
  - **Inhibition** if with dissimilar preferred stimuli/orientations.
- Preferred orientation of each pixel in an image:

$$\theta(x) = \arctan \frac{\mathcal{L}_v(x)}{\mathcal{L}_h(x)}$$

- Interaction between two pixels:

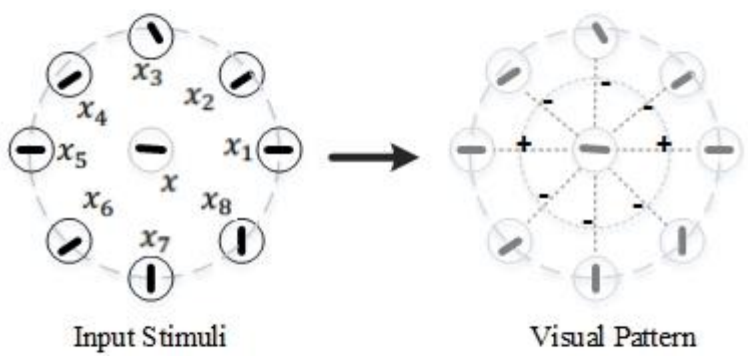
$$\mathcal{I}(x \mid x_i) = \begin{cases} + & \text{if } \left| \theta(x) - \theta(x_i) \right| < \mathcal{T} \\ - & \text{else} \end{cases}$$

The OS based Visual Pattern Modeling

The OS modeling : OS based visual pattern

- The arrangement of interactions among the central pixel and its surrounding pixels:

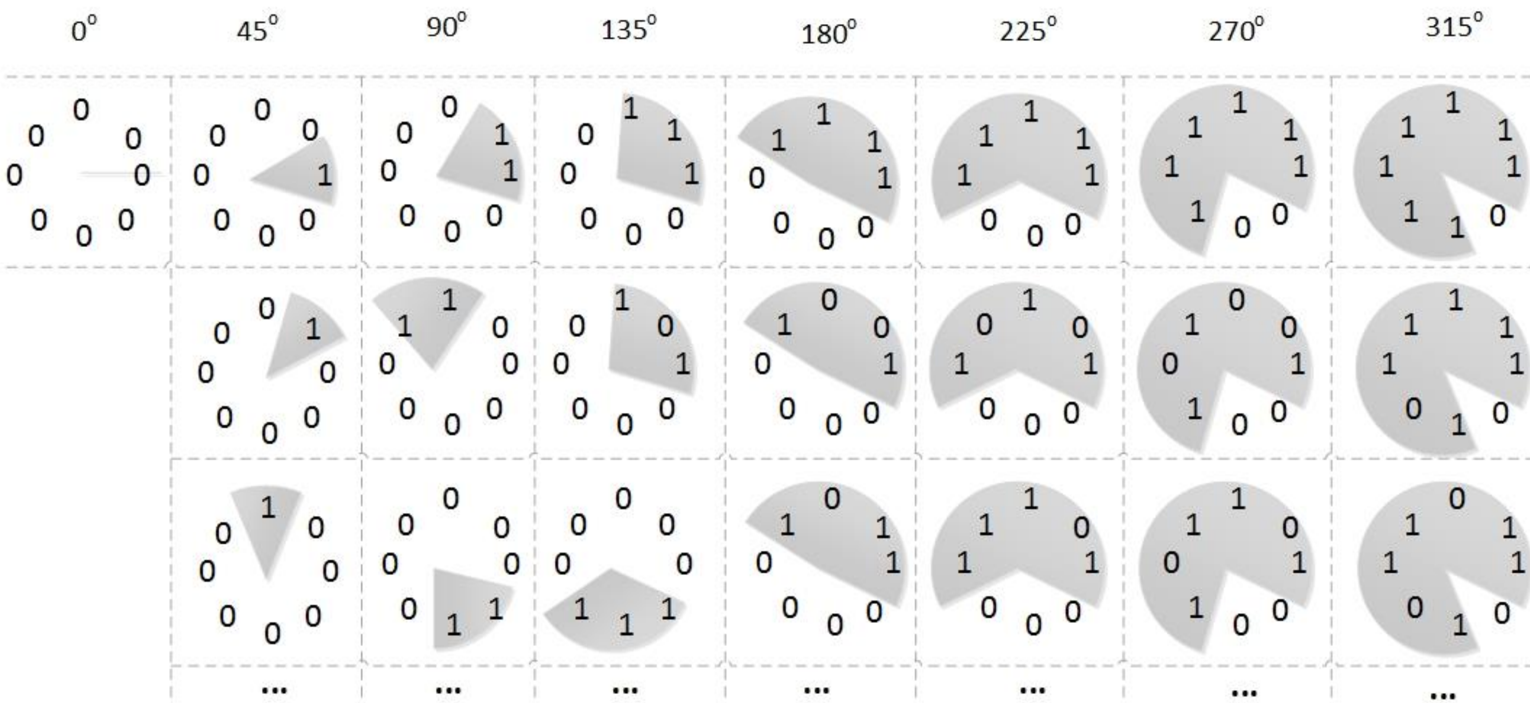
$$\mathcal{P}(x \mid \mathcal{X}) = \mathcal{A}(\mathcal{I}(x \mid x_1), \mathcal{I}(x \mid x_2), \cdots, \mathcal{I}(x \mid x_n))$$



The OS modeling : fundamental pattern extraction

- Too many original patterns: the number of pattern increases exponentially with the neighbor number.

- Pattern reduction: fundamental pattern
  - Patterns with same excitatory subfield represent similar response



The OS based Visual Pattern Modeling

The OS modeling : fundamental pattern extraction

- Too many original patterns: the number of pattern increases exponentially with the neighbor number.

- Pattern reduction: fundamental pattern
  - Patterns with same excitatory subfield represent similar response

- Fundamental patterns:
  - A N-neighborhood local receptive field corresponds to N sector domains of the excitatory subfield;
  - E.g., for an 8-neighborhood, there are 8 excitatory subfield, {0^o, 45^o, 90^o, 135^o, 180^o, 225^o, 270^o, 315^o}

OS based structure extraction:

- A successful image structure descriptor should effectively represent:
  - Intensity change;
  - Spatial distribution.
- The OS based pattern: represents the spatial distribution.
- For intensity change representation: gradient magnitude,

$$\mathcal{M}(x) = \sqrt{(G_h(x))^2 + (G_v(x))^2}$$

Image Content Extraction

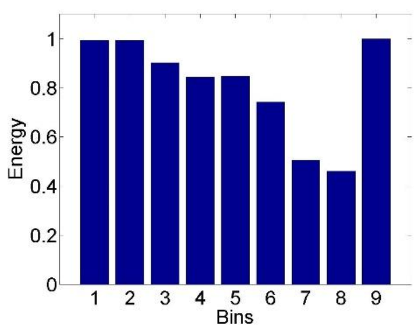
OS based structural histogram:

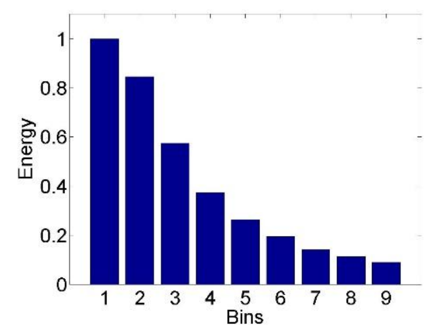
- Mapping an image into a structural histogram: with both OS based pattern and gradient magnitude, the structural information is extracted:

$$H_w(k) = \sum_{x=1}^{N} \mathcal{M}(x)\ \delta(\mathcal{P}_f(x), \mathcal{P}_f^k)$$

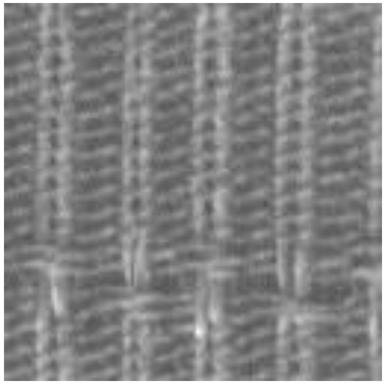
where:

$$\delta(\mathcal{P}_f(x), \mathcal{P}_f^k) = \begin{cases} 1 & \text{if } \mathcal{P}_f(x) = \mathcal{P}_f^k \\ 0 & \text{else} \end{cases}$$

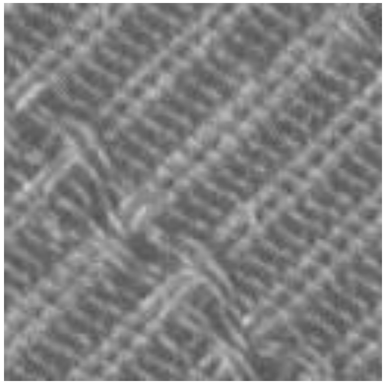




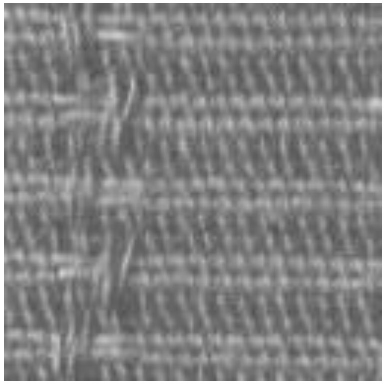
Rotation invariance:



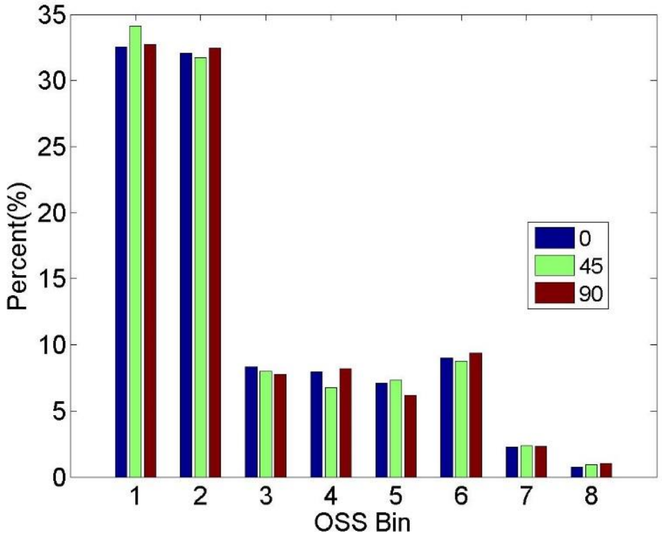
0^{o}



45^{o}

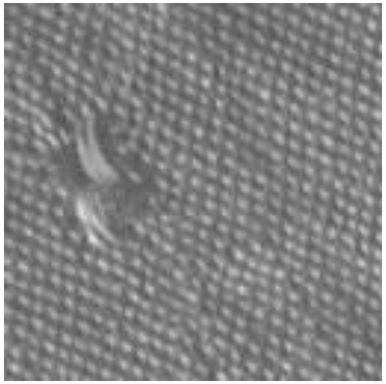


90^{o}

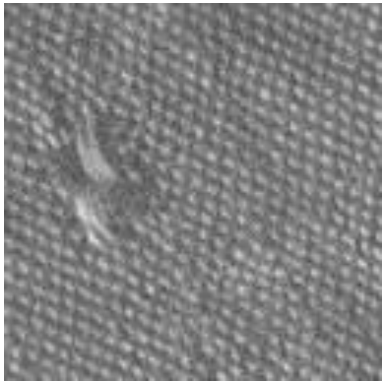


Application 1: Texture Classification

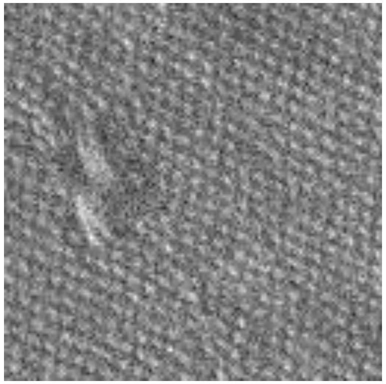
Robustness:



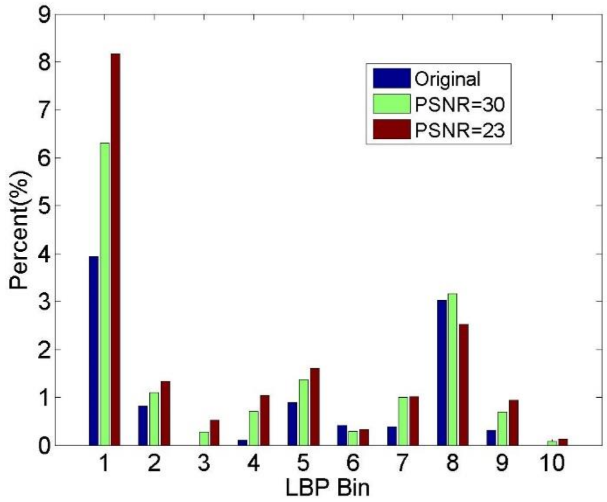
Original

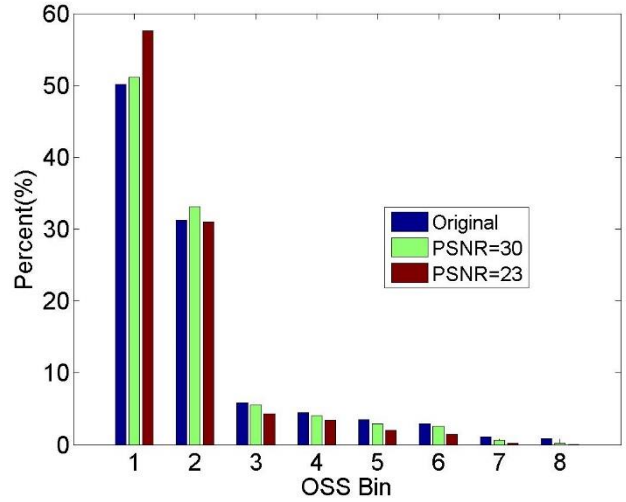


PSNR=30dB



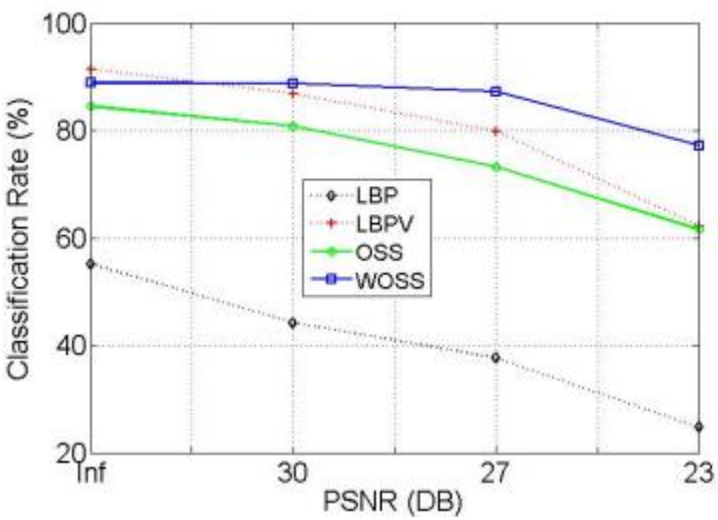
PSNR=23dB



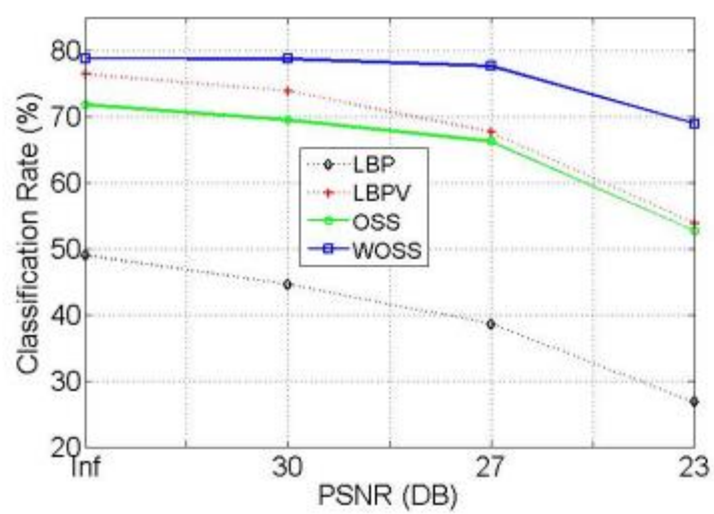


Texture Classification:

- TC10 database:
  - Training: images under 'inca' illumination and '0◦' angle (480 images);
  - Testing: the other images (480×8 images).



- TC12 database:
  - Training: same as that in TC10;
  - Testing: images under 't184' and 'horizon' illuminations and nine angles (480×2×9 images).



Jinjian Wu et al., "Visual Orientation Selectivity based Structure Description", IEEE TIP2015.

Structural degradation caused by different distortions:

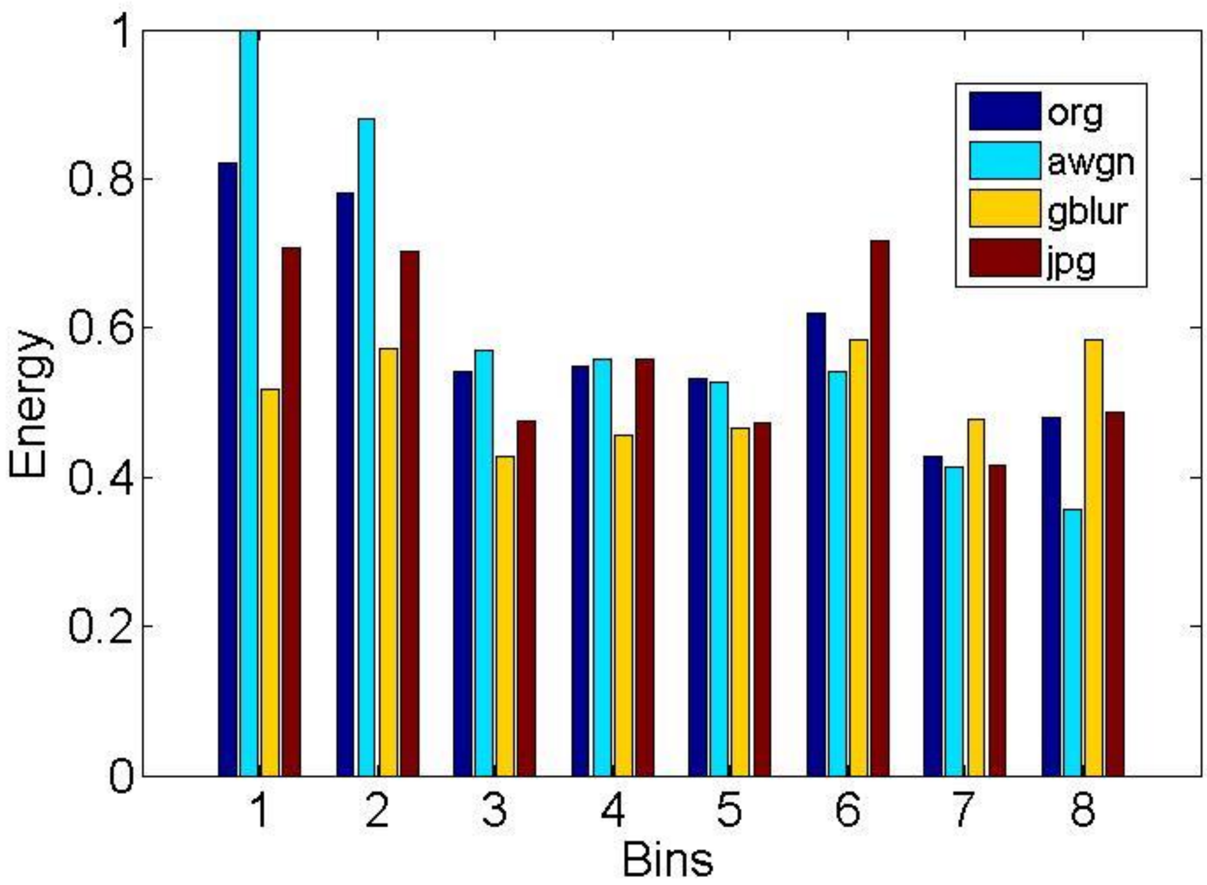
Application 2: Quality Assessment

Structural degradation caused by different distortions:


Though under a same level of noise, their qualities are different.

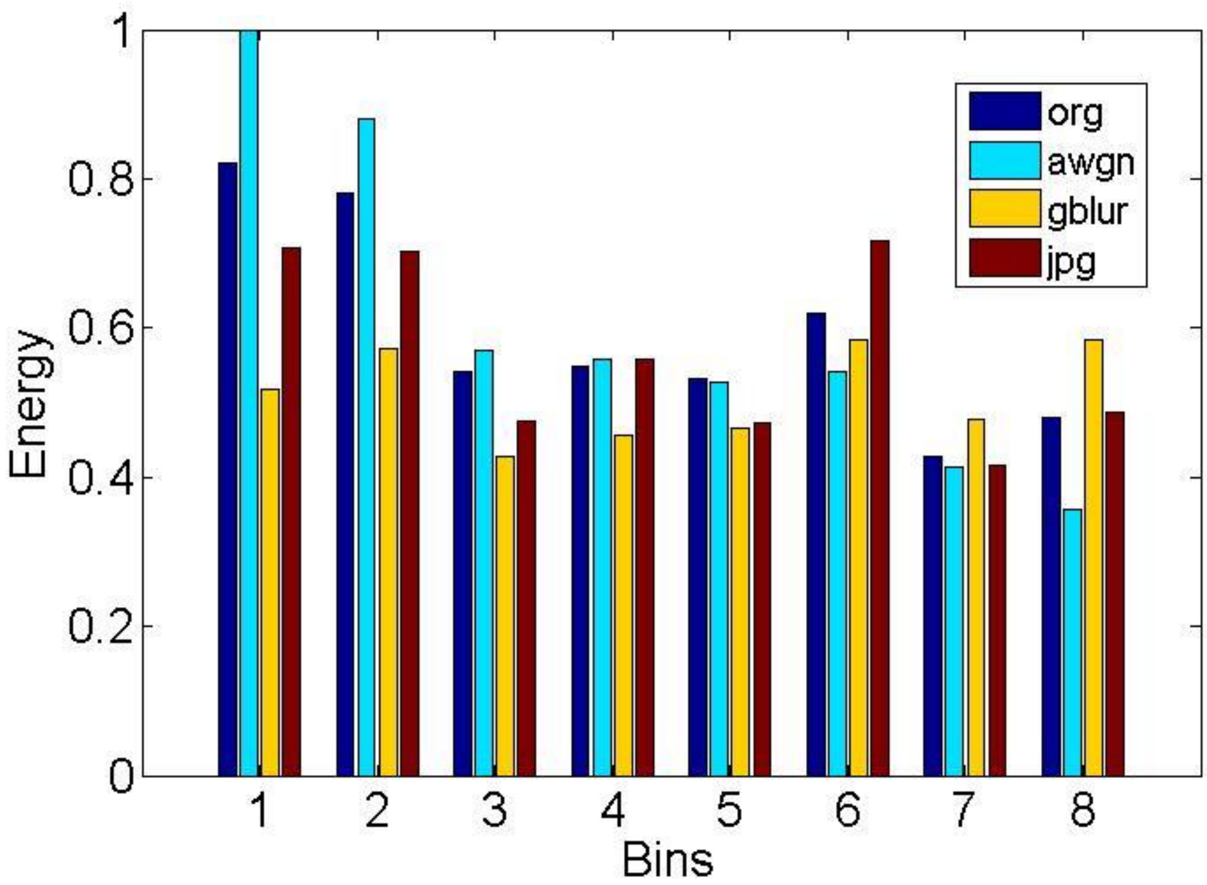

Application 2: Quality Assessment

Structural degradation caused by different distortions:



Application 2: Quality Assessment

Structural degradation caused by different distortions:



Different distortion types cause different structural degradation

Application 2: Quality Assessment

西安电子科技大学 XIDIAN UNIVERSITY

Quality assessment according to structural degradation:

Performance on LIVE database

Dist.	Crit. \ Algo.		RR			FR			
		OSS	RRED [22]	WNISM [21]	RRVIF [38]	PSNR	SSIM [2]	MSSIM [39]	FSIM [40]
No. of scalars		16	16	18	2	N	N	N	N
j2k	CC	*0.932*	0.952	0.924	0.932	0.896	0.941	0.957	0.978
	SRCC	*0.921*	0.946	0.920	0.950	0.889	0.936	0.953	0.971
	RMSE	*8.91*	7.74	6.18	5.88	11.0	8.51	4.69	5.22
jpg	CC	*0.958*	0.959	0.876	0.895	0.860	0.951	0.943	0.984
	SRCC	*0.944*	0.953	0.851	0.885	0.841	0.944	0.942	0.983
	RMSE	*8.91*	9.08	7.71	7.15	14.5	9.89	5.33	5.72
awgn	CC	*0.979*	0.946	0.890	0.957	0.982	0.969	0.974	0.965
	SRCC	*0.972*	0.946	0.870	0.946	0.985	0.963	0.973	0.965
	RMSE	*5.58*	9.08	7.29	4.66	4.334	6.901	3.65	7.34
gblur	CC	*0.972*	0.956	0.888	0.955	0.784	0.874	0.955	0.969
	SRCC	*0.962*	0.952	0.915	0.961	0.782	0.894	0.954	0.971
	RMSE	*4.19*	5.42	7.22	4.66	11.5	8.96	4.69	4.57
ff	CC	*0.924*	0.895	0.925	0.944	0.890	0.945	0.947	0.946
	SRCC	*0.907*	0.918	0.923	0.941	0.890	0.941	0.947	0.950
	RMSE	*10.4*	12.7	6.25	5.42	13.0	9.36	5.30	9.21
overall	CC	*0.863*	0.827	0.710	0.725	0.872	0.904	0.943	0.960
	SRCC	*0.865*	0.830	0.703	0.732	0.876	0.910	0.945	0.963
	RMSE	*12.4*	15.3	18.4	17.6	13.4	11.7	9.09	7.68

Application 2: Quality Assessment

Quality assessment according to structural degradation:

Overall performance on CSIQ and TID databases

DB	Algo. / Crit.	RR				FR			
		WOSS	RRED [22]	WNISM [21]	RRVIF [38]	PSNR	SSIM [2]	MSSIM [39]	FSIM [40]
CSIQ	CC	*0.901*	0.695	0.696	0.698	0.800	0.815	0.900	0.928
	SRCC	*0.895*	0.773	0.705	0.733	0.806	0.838	0.914	0.940
	RMSE	*0.113*	0.189	0.189	0.182	0.158	0.152	0.115	0.098
TID	CC	*0.782*	0.712	0.572	0.535	0.573	0.641	0.842	0.886
	SRCC	*0.759*	0.702	0.495	0.500	0.579	0.627	0.853	0.890
	RMSE	*0.829*	0.943	1.101	1.134	1.100	1.03	0.723	0.623

Jinjian Wu et al., "Orientation Selectivity based Visual Pattern for Reduced-Reference Image Quality Assessment", Information Science 2015.

Conclusion

- In this paper, a novel structure descriptor is introduced.
- Inspired by the orientation selectivity mechanism in the primary visual cortex, the interactions among neurons are analyzed.
- By mimicking the excitation/inhibition in a local receptive field, a novel visual pattern is proposed to describe the spatial correlation among pixels.
- Taking both the luminance change and visual pattern into account, a novel structural descriptor is designed.
- Applications on texture classification and quality assessment demonstrate the effectiveness of the proposed structural descriptor.

Related papers

西安电子科技大学
XIDIAN UNIVERSITY

- Jinjian Wu, et al., "Orientation Selectivity based Visual Pattern for Reduced-Reference Image Quality Assessment", **Information Science**, vol.351, pp.18-29, July 2016. (source code is available at **http://web.xidian.edu.cn/wjj/index.html**)
- Jinjian Wu, , et al., "Visual Orientation Selectivity based Structure Description", **IEEE TIP**, Vol. 24, No. 11, PP. 4602-4613, Nov. 2015.
- Jinjian Wu, et al., “Reduced-Reference Image Quality Assessment with Orientation Selectivity based Visual Pattern”, IEEE ChinaSIP 2015, Chengdu, Sichuan.
- Jinjian Wu, et al., “Visual Pattern Degradation based Image Quality Assessment”, OIT’2015, Beijing, China.
- Jinjian Wu, et al., "Orientation Selectivity based Structure for Texture Classification", SPIE /COS Photonics Asia Conference, 2014, Beijing, China.

Thanks