

# WICI网络智能和脑信息学暑期研讨班

## 面向Web的语义表示和内容理解

### Web-Oriented Semantic Representation and Content Understanding

史忠植

[shizz@ics.ict.ac.cn](mailto:shizz@ics.ict.ac.cn)

中国科学院计算技术研究所



中国科学院计算所  
INSTITUTE OF COMPUTING TECHNOLOGY

# 内容提要

---

- 引言
- 语义表示
- 事件来龙去脉生成
- 图像内容理解
- 结束语

# 致谢

## 感谢合作者:

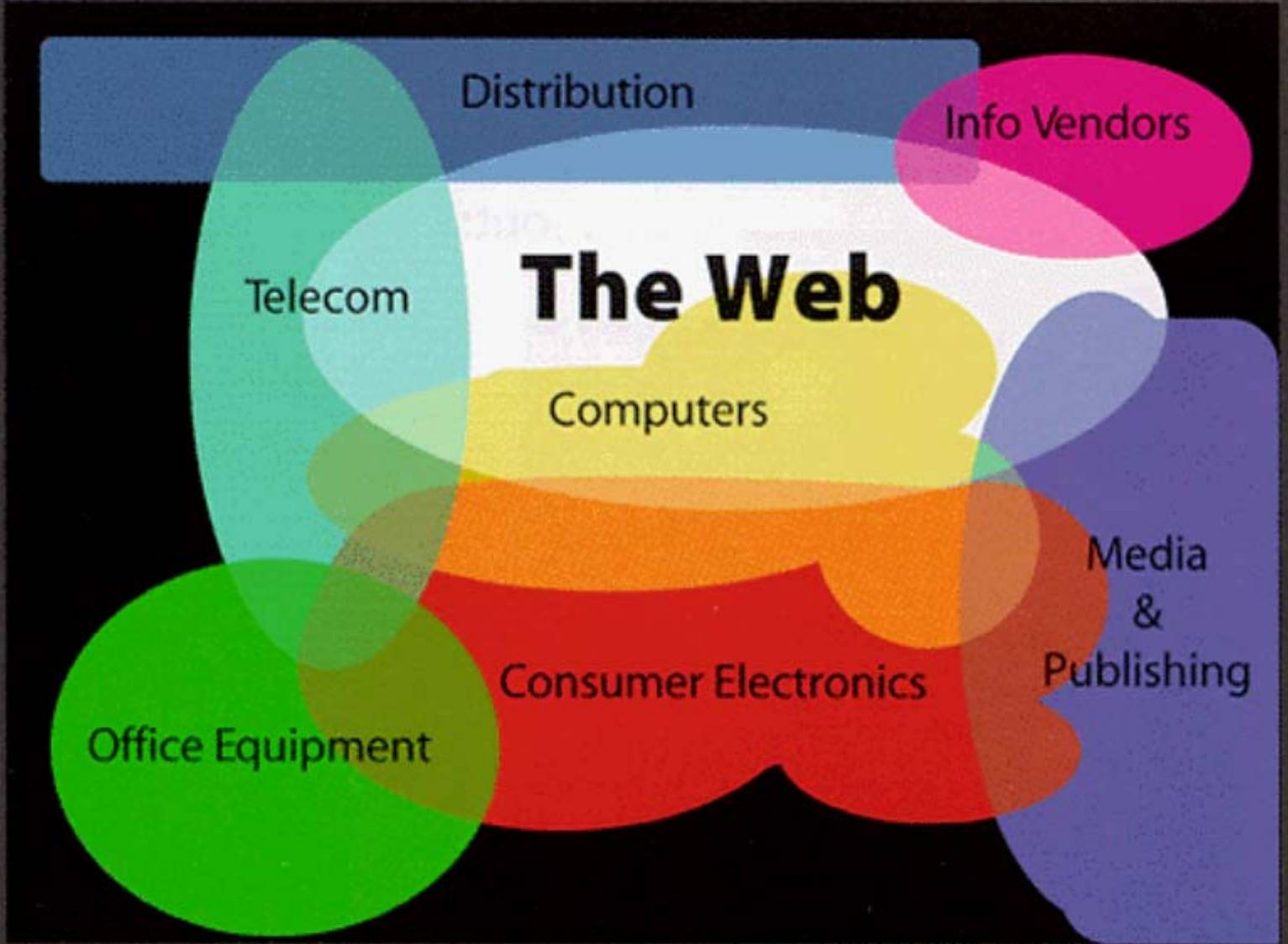
- 贾自艳 (2004). **Web**信息智能获取若干关键问题研究. 博士学位论文, 中国科学院计算技术研究所
- 李清勇 (2006). 视觉感知的稀疏编码理论及其应用研究. 博士学位论文, 中国科学院计算技术研究所.
- 张素兰 (2007). 视知觉组织模型及应用研究. 博士学位论文, 北京: 中科院计算技术研究所.

# 致谢

## 感谢下列资助:

- 国家自然科学基金重点项目  
课题编号: **60435010**, 课题名称: 基于感知学习和语言认知的智能计算模型
- **973**课题  
编号: **2003CB317004**, 课题名称: 语义网格资源描述模型、形式化理论和支撑技术.
- **973**课题  
编号: **2007CB311004**, 课题名称: 非结构化(图像)信息的内容理解与语义表征.

Services  
Products



'90s

Container/Medium    Transport|Translate|Transform    Content/Message

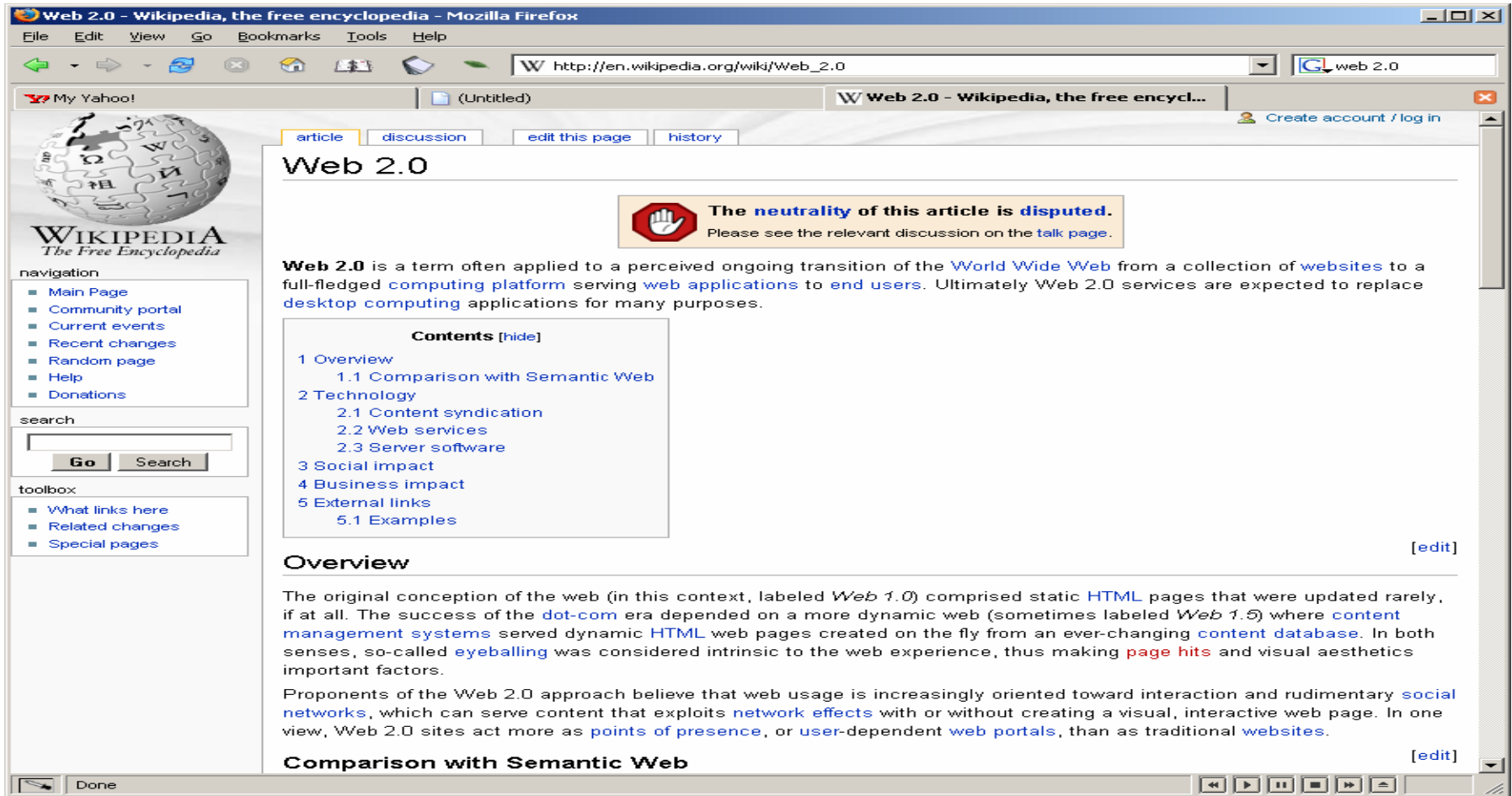
# Web 2.0

- 奥莱理是技术图书出版商、奥莱理媒体公司的主席兼**CEO**，**2003**年，他灵感突发，提出**Web2.0**的概念。
- **2004**年，由奥莱理媒体公司和**CMP**技术公司举办的第一届**Web2.0**大会在洛杉矶召开。
- **Web2.0**的应用在博客、播客、**RSS**等上面。通过用户聚合数据来建立价值。
- 新一轮**Web2.0**，一些非常轻巧的桌面应用如**badges**、**widgets**和**gadgets**等开始流行，用户可以复制、粘贴自己喜欢的内容到个人博客、空间或者网站上。

# Web 2.0: Evolution Towards a Read/Write Platform

Web 1.0 (1993-2003)		Web 2.0 (2003- beyond)
Pretty much HTML pages viewed through a browser		Web pages, plus a lot of other "content" shared over the web, with more interactivity; more like an application than a "page"
"Read"	Mode	"Write" & "Contribute"
"Page"	Primary Unit of content	"Post / record"
"static"	State	"dynamic"
Web browser	Viewed through...	Browsers, RSS Readers, anything
"Client Server"	Architecture	"Web Services"
Web Coders	Content Created by...	Everyone
"geeks"	Domain of...	"mass amatuerization"

# Wikipedia



Web 2.0 - Wikipedia, the free encyclopedia - Mozilla Firefox


File Edit View Go Bookmarks Tools Help

http://en.wikipedia.org/wiki/Web\_2.0

My Yahoo! (Untitled) Web 2.0 - Wikipedia, the free encycl... Create account / log in

article discussion edit this page history

## Web 2.0

 **The neutrality of this article is disputed.**  
Please see the relevant discussion on the [talk page](#).

**Web 2.0** is a term often applied to a perceived ongoing transition of the [World Wide Web](#) from a collection of [websites](#) to a full-fledged [computing platform](#) serving [web applications](#) to [end users](#). Ultimately Web 2.0 services are expected to replace [desktop computing](#) applications for many purposes.

**Contents** [\[hide\]](#)

- Overview
  - 1.1 Comparison with Semantic Web
- Technology
  - 2.1 Content syndication
  - 2.2 Web services
  - 2.3 Server software
- Social impact
- Business impact
- External links
  - 5.1 Examples

[\[edit\]](#)

### Overview

The original conception of the web (in this context, labeled *Web 1.0*) comprised static [HTML](#) pages that were updated rarely, if at all. The success of the [dot-com](#) era depended on a more dynamic web (sometimes labeled *Web 1.5*) where [content management systems](#) served dynamic [HTML](#) web pages created on the fly from an ever-changing [content database](#). In both senses, so-called [eyeballing](#) was considered intrinsic to the web experience, thus making [page hits](#) and visual aesthetics important factors.

Proponents of the Web 2.0 approach believe that web usage is increasingly oriented toward interaction and rudimentary [social networks](#), which can serve content that exploits [network effects](#) with or without creating a visual, interactive web page. In one view, Web 2.0 sites act more as [points of presence](#), or [user-dependent web portals](#), than as traditional [websites](#).

[\[edit\]](#)

### Comparison with Semantic Web

# Blogging



## US Internet Users Who Read Blogs, 2004 & 2005 (as a % of respondents)



Source: Pew Internet & American Life Project, May 2005

064578 ©2005 eMarketer, Inc.

www.eMarketer.com

## Frequency with which US Internet Users Read Political Blogs, March 2005 (as a % of respondents)



Note: n=2,630

Source: Harris Interactive, April 2005

064580 ©2005 eMarketer, Inc.

www.eMarketer.com

# Social Networks



## LinkedIn

Find out more

Find People

## ryze

Business Networking

Ryze Home Invite

About Ryze

New to F

Ryze  
up No

Sign

## CONSUMATING

beta

Join or Login  
People: Tags  
Updates | Questions | News

# DATE BETTER!

Meet hipsters, bloggers, freaks and geeks in your area by showing off your quirky personality with zany answers to our constantly rotating questions. Find cute nerds to hook up with in our people browser!

SIGN UP NOW >>

Already Got An Account?

Welcome back! To access your account and all the hotties that wait inside, log in below.

e-mail

password

Log In >>

Groups of people...  
living near you...  
sharing an interest...  
and meeting r

Join over a million people who:

Learn something

Do something

French Speakers in Seattle, WA

Share something

Pug Owners in Boulder, CO

Scrapbookers in Singapore

Change something

Voters in San Bernardino, CA

What are you inter

e.g., Knitting, Elvis, French,  
[Browse all interests](#) or [Brows](#)

Want to learn more? [Tak](#)

TIME

"A convenien  
non-threaten  
connect to ot  
who share sit  
interests and

Chicago  
Tribune

"No matter w  
interest, ther  
for you."

"The people were incredible... we'll be  
friends for life. Meetups are a great idea!"  
—Stay-at-home Mom in Raleigh-Durham, NC

Question Of The Week:

## George Lucas: Savior Of Film or Destroyer Of Your Childhood?

What do you think? [Read and rate the responses to this week's question and write your own!](#)

Get started! Click on a characteristic you like!

cactus lucky13 spirituality katebush bettendorf blackcats play zen blake yebisu  
almost-sarcasm t-tests stencilart rachmaninov cutie theirongiant cheapredwine  
wearing cocktails drinkingsocially alsacelorraine tribe silversurfer cocacola dunno

Done

Read www.ryze.com

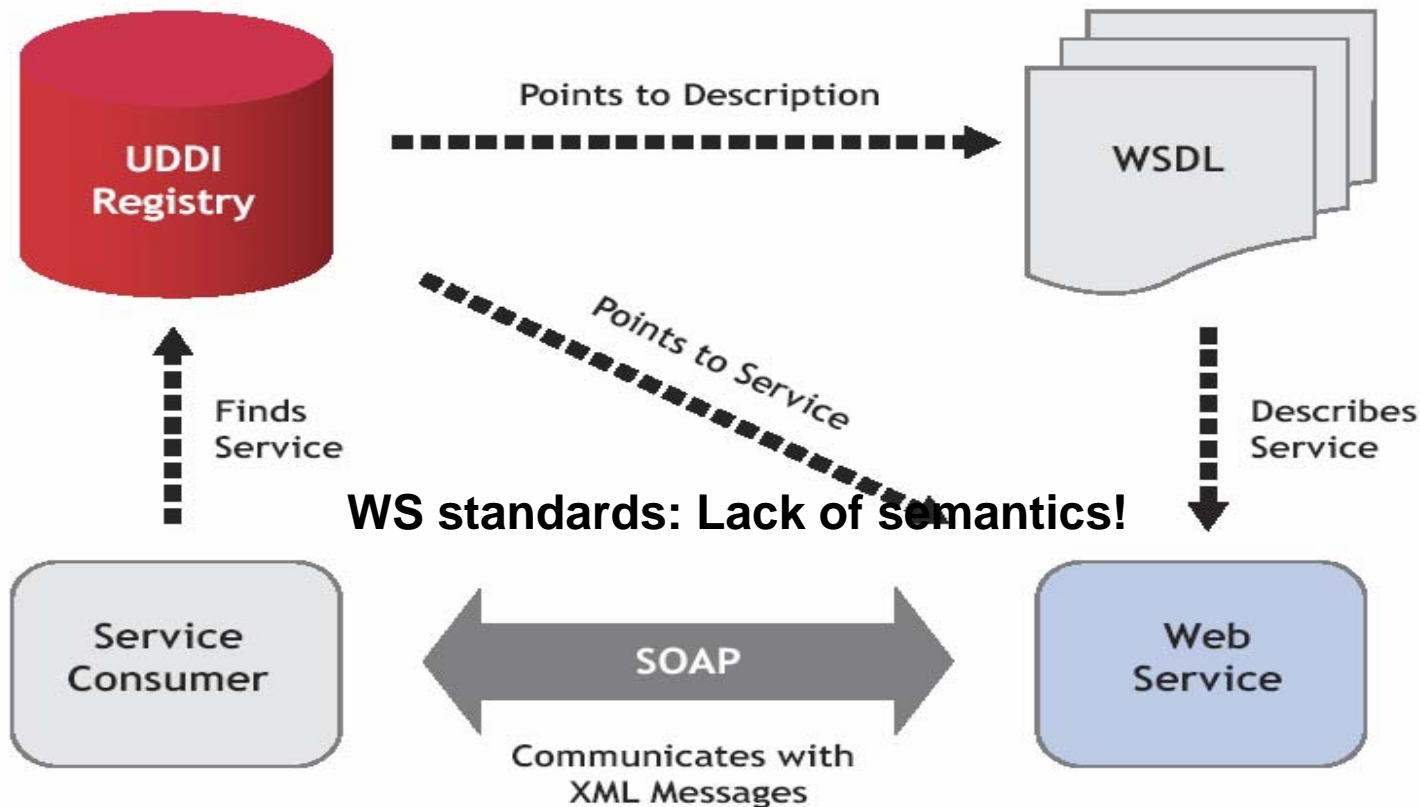
# Web 服务



中科院计算所  
INSTITUTE OF COMPUTING  
TECHNOLOGY



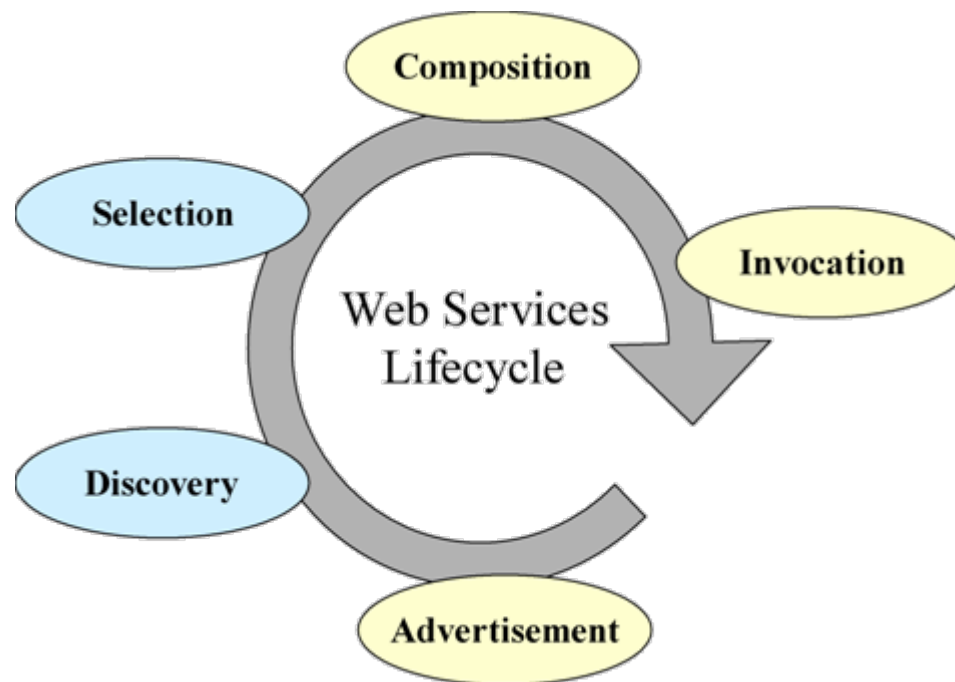
# Web 服务



## Web Service Architecture

# Web 服务

- *programmatic interfaces for applications (i.e., business logic), available over the WWW infrastructure and developed with XML technologies.*



# 内容提要

---

- 引言
- 语义表示
- 事件来龙去脉生成
- 图像内容理解
- 结束语

# 语义 Web的目的

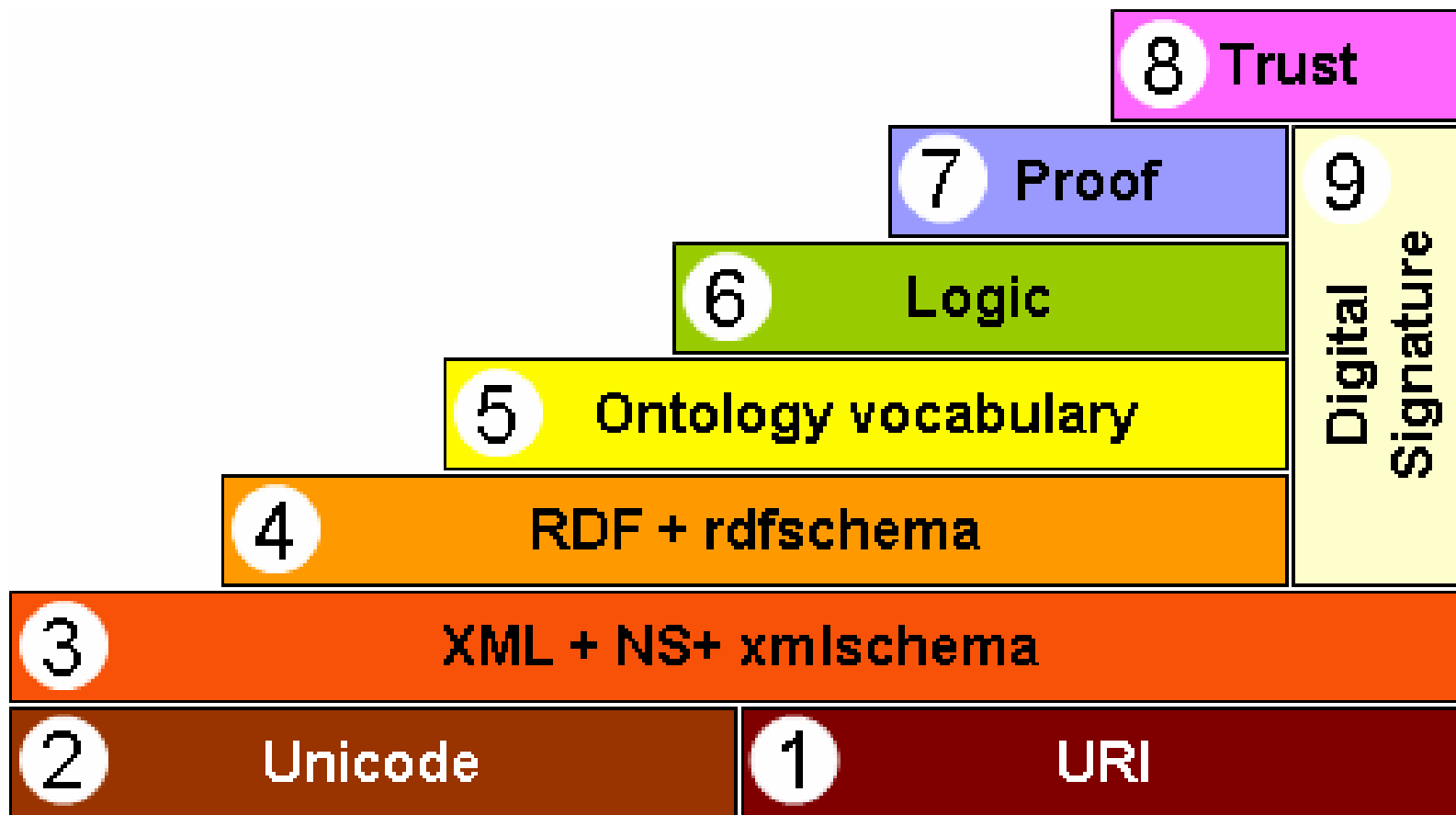
	Static	Dynamic	Syntax	Semantic
Encoding	HTML	+ RDBMS	+ XML	+ RDF/OWL
Creation	Manually	Generated by server-side applications	Generated by applications based on schema	Generated by applications based on models
Users	Humans	Humans	Humans and applications	Humans and applications
Paradigm	Browse	Create/Query/Update	Integrate	Interoperate
Applications	Browsers	Browsers	Process Integration, EAI, BPMS, Workflows	Intelligent agents, Semantic engines

1995

2000

2005

# 语义 Web 的架构



# HTML标识

.....

`<h2>Zhongzhi Shi</h2>`

`<b>Affiliation</b>:`

Laboratory of Intelligent Information Processing`<br>`

Institute of Computing Technology`<br>`

Chinese Academy of Sciences`<p>`

`<b>Email</b>:` shizz @ ics.ict.ac.cn`<br>`

`<b>Phone</b>:` 86-10-82610254(office)

.....

`</html>`

8/27/2007

# XML标注

---

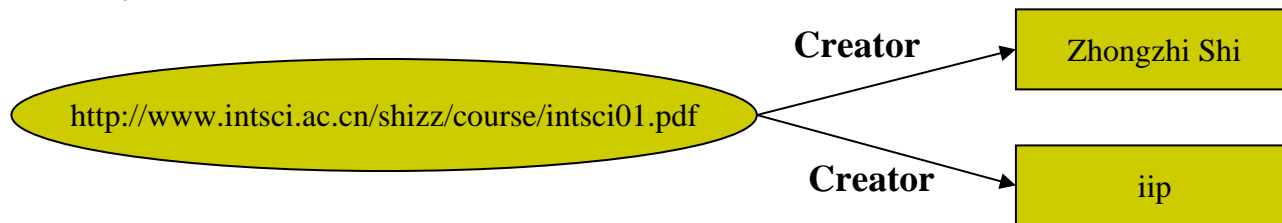
```
<researcher><name>Zhongzhi Shi</name>  
<affiliation>  
<Laboratory>Laboratory of Intelligent Information  
  Processing</Laboratory>  
<Institute>Institute of Computing Technology</Institute>  
<Academy>Chinese Academy of Sciences </Academy>  
</affiliation>  
<email>shizz @ ics.ict.ac.cn</email>  
<phone id="office"> (86)-10-82610254</phone>  
.....</researcher>  
</html>
```

# XML Schema

```
<xsd:element name="reseracher">
  <xsd:complexType>
    <xsd:element name="name" type="xsd:String"/>
    <xsd:element name="affiliation" type="affil"
      minOccurs="1" maxOccurs="unbounded"/>
    <xsd:element name="phone" type="xsd:String"/>
    <xsd:element name="email" type="xsd:String"/>
  </xsd:complexType>
</xsd:element>
<xsd:complexType name="affil">
  <xsd:element name=" department"
type="xsd:String"/>
  <xsd:element name=" faculty" type="xsd:String"/>
  <xsd:element name="university" type="xsd:String"/>
</xsd:complexType>
```

- *Metadata is machine understandable information about web resources or anything that has an URI, it is represented as a set of independent assertions:*

Triple:  $T(\text{subject, attribute, values})$



```
<rdf:Description about="http://www.intsci.ac.cn/shizz/course/intsci01.pdf">
  <dc:Creator rdf:resource="http://www.intsci.ac.cn/shizz"/>
  <dc:Creator rdf:resource="mailto:iip@ics.ict.ac.cn"/>
</rdf:Description>
```

# RDF: Dublin Core

---

- The Dublin Core provides properties for describing network objects, suitable for use by network search engines.
- The Dublin Core is a set of predefined properties for describing documents.
- The first Dublin Core properties were defined at the Metadata Workshop in Dublin, Ohio in 1995 and is currently maintained by the [Dublin Core Metadata Initiative](#).

# Dublin Core Metadata Initiative

---

- **The Dublin Core Metadata Initiative** is an open forum engaged in the development of interoperable online metadata standards that support a broad range of purposes and business models.
- <http://dublincore.org/>

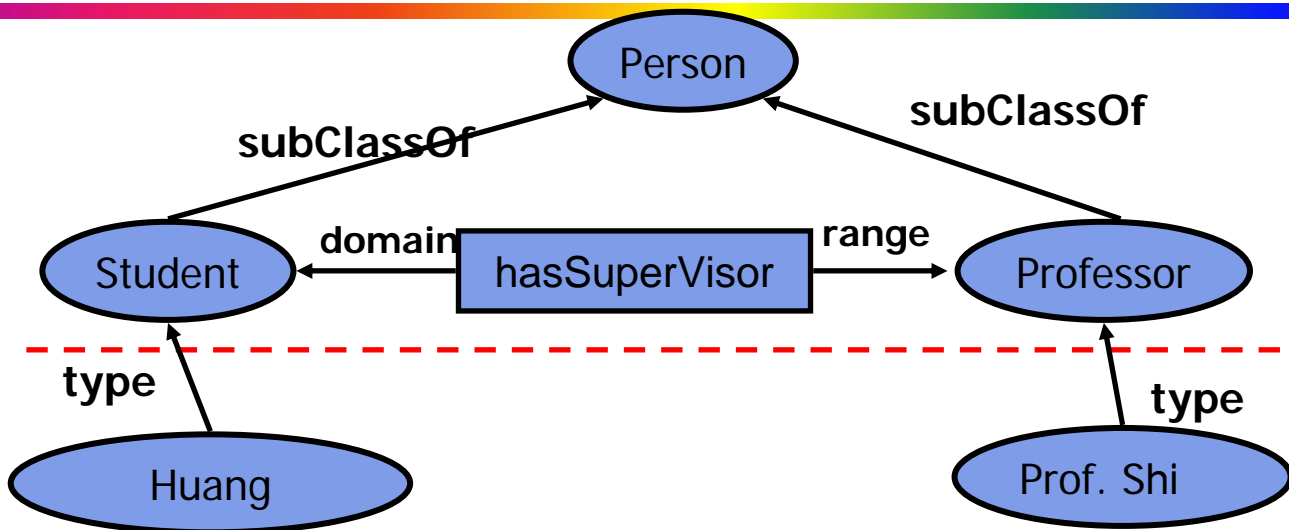
# Annotating Metadata

```
<rdf:Description rdf:about=...dc-rdf/">
  <dc:title>
    Guidance on expressing the Dublin Core within the Resource
    Description Framework (RDF)
  </dc:title>
  <dc:creator> Eric Miller </dc:creator>
  <dc:creator> Paul Miller </dc:creator>
  <dc:creator> Dan Brickley </dc:creator>
  <dc:subject> Dublin Core; RDF; XML </dc:subject>
  <dc:publisher> Dublin Core Metadata Initiative
</dc:publisher>
  <dc:contributor> Dublin Core Data Model Working Group
</dc:contributor>
  <dc:date> 1999-07-01 </dc:date>
  <dc:format> text/html </dc:format>
  <dc:language> en </dc:language>
</rdf:Description>
```

# 资源描述框架模式 RDF Schema (RDFS)

- RDFS defines **vocabulary** for RDF
- Organizes this vocabulary in a **typed hierarchy**
  - Class, subclassOf, type
  - Property, subPropertyOf
  - domain, range

# RDFS



# Web Ontology Language (OWL)

---

- OWL is built on top of RDF
- OWL is for processing information on the web
- OWL was designed to be interpreted by computers
- OWL was not designed for being read by people
- OWL is written in XML
- OWL is a web standard

# OWL

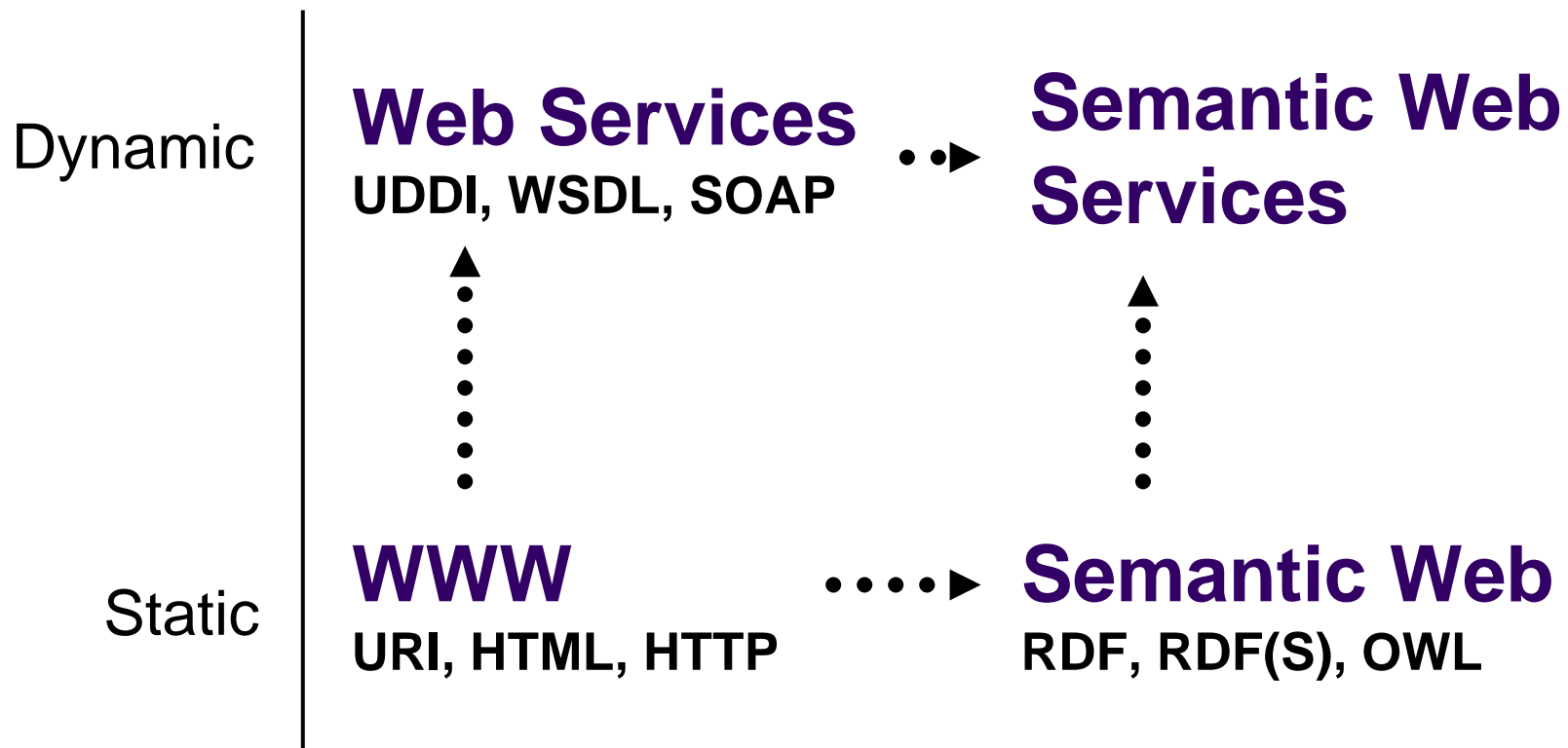
- Ability to be distributed across many systems
- Scalability to Web needs
- Compatibility with Web standards for accessibility and internationalization
- Openness and extensibility

# OWL 推理机

- [KAON2](#) is a reasoner for OWL extended with the DL-safe subset of SWRL; it also provides an OWL API.
- FaCT -- a DL reasoner. see [WonderWeb project](#), [Bechhofer 15 Sep](#).
- Racer -- a DL reasoner. see [Horrocks 12Sep](#)
- Cerebra from [Network Inference](#) - **owl syntax checker**, nearly complete OWL DL [Horrocks 12Sep](#)
- [cwm](#) -- useful but incomplete OWL Full
- Euler -- useful but incomplete OWL Full, see [De Roo 11 Jul](#): 51 / 234 tests
- [surnia](#) -- OWL full reasoner based on otter. [Hawke 26Aug](#)
- [Jena/HP](#) ( [Reynolds/HP 7 May](#)) will support OWL reasoning.
- Vampire [Horrocks 17 Jul](#) - uses a first-order theorem prover to do OWL DL
- [Pellet](#) is a reasoner built in Java that was designed specifically for OWL reasoning. [Hendler/Sirin/Parsia 15Sep](#)).
- [SWI-Prolog Semantic Web Library](#) contains *owl.pl* - an OWL reasoning package.
- [F-OWL](#) is an f-logic based Owl tool from UMBC.
- [E-wallet](#) is an e-commerce and mobile computing tool based on a rule-based OWL reasoner.



# 语义 Web 服务



# 语义 Web 服务

**Semantic Web Services**

**=**

**Semantic Web Technology**

**+**

**Web Service Technology**

# 语义 Web 服务

- Define exhaustive description frameworks for describing Web Services and related aspects (**Web Service Description Ontologies**)
- Support ontologies as underlying data model to allow machine supported data interpretation (**Semantic Web aspect**)
- Define semantically driven technologies for automation of the Web Service usage process (**Web Service aspect**)

# 语义 Web 服务

- What should S+WS ontologies provide?  
(Mainly) Automation of the Usage Process:
  - Publication
  - Discovery
  - Selection
  - Composition
  - Execution
  - Monitoring

# 内容提要

---

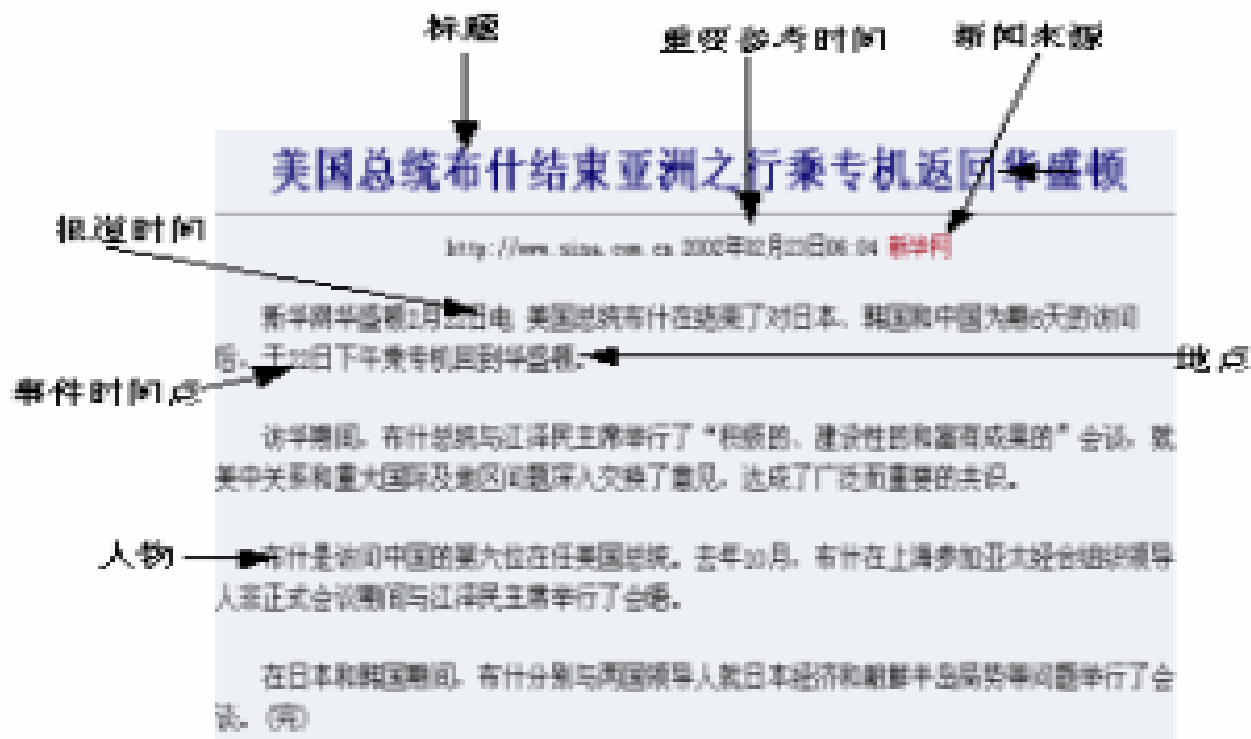
- 引言
- 语义表示
- 事件来龙去脉生成
- 图像内容理解
- 结束语

# 问题提出

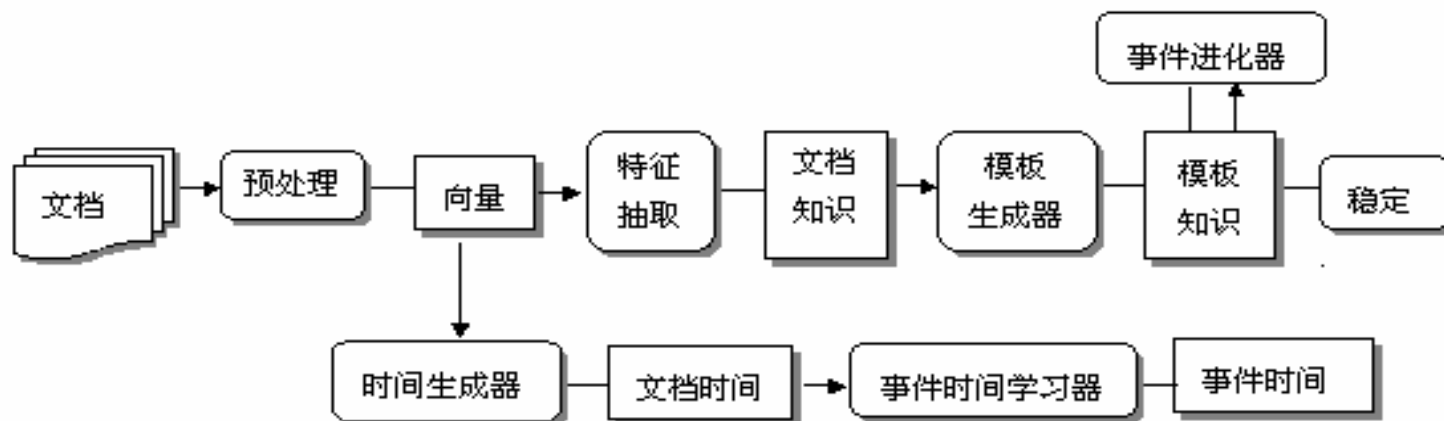
- 新闻占网上信息极大的比重（**84.38**）
- 几乎所有的人都关注新闻信息（**57.97%**）。
- 网络新闻内容丰富、形态多样。
- 新闻书写格式很平凡
- 适用于挖掘出新奇性信息（事件），同时能够跟踪该事件。

一般来说，网络新闻具有实效性、综合性、可存储、可持久、选择性和交互性等特征。互联网的优势不仅仅是能够获取“取之不尽，用之不竭”的信息和资料，更重要的是对信息资源的深度发掘。简单的“大而全”或者面面俱到都是行不通的，迟早要被观众所淘汰。

# 网络新闻示意图



# 事件主要元素的学习过程



- 文本权重计算:

$$weight_{t,d}' = \sum_{\Lambda} tf_{t,\Lambda} \times w_{\Lambda} \quad weight_{t,d} = \frac{weight_{t,d}'}{\text{MAX}_{t, in d} (weight_{t,d}')}$$

- 文本特征选择: StopList; 阈值 (0.3)

# 时间知识学习

- 文档中时间信息存在形式  
绝对时间；相对时间；模糊时间
- 时间参考信息  
最近修改时间；文档创建时间；正文中的时间

输入：文档集 DocS。

输出：TimeS，存放文档集中每篇文档所报道事件及其对应时间。

1. 顺序处理 DocS 中的每篇文档。
2.  $Get(Time_{last}) \rightarrow origin_{last}$  //得到文档最近修改时间，设为时间原点。
3.  $Get(Time_{create})$  //得到文档的创建时间。
4.  $If(Time_{create} \neq \phi) Time_{create} \rightarrow origin_{last}$  //如果存在，设为时间原点。
5.  $Match(Tag_{time}, Doc)$  //据时间标记符号对正文中时间表达进行定位。
6.  $Parse(STime_{content}, Doc)$  //解析得到所有文本中时间表示。
7.  $Format(STime_{content}, origin_{last})$  //根据时间原点，将所有事件表达为标准形式。
8.  $TimeS = TimeS \cup Best(origin_{last}, STime_{content})$  //分析所有时间挑选出最能表达该文档中事件发生的时间。

# 事件模板特征知识学习

- 权重计算:

$$\text{weight}_{t, NE}'' = \sum_{D_j} \text{weight}_{t, D_j} \quad \text{weight}_{t, NE}' = \frac{\text{weight}_{t, NE}''}{\text{MAX}_{t, in NE} \{ \text{weight}_{t, NE} \}}$$

- 词对事件支持度分布

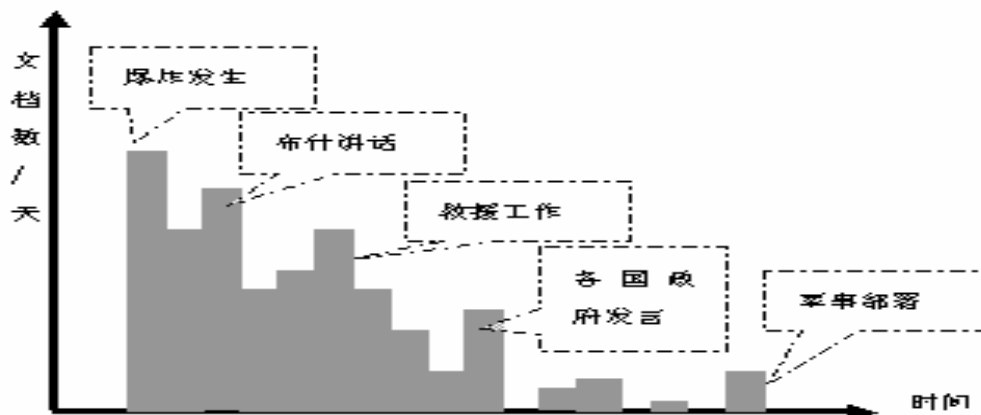
	[0.0 ~0.1)	[0.1 ~0.2)	[0.2 ~0.3)	[0.3 ~0.4)	[0.4 ~0.5)	[0.5 ~0.6)	[0.5 ~0.7)	[0.7 ~0.8)	[0.8 ~0.9)	[0.9 ~1.0]
NE1	311	15	8	3	3	2	4	1	1	3
NE2	421	9	5	5	2	1	0	1	1	2
NE3	347	13	9	5	5	5	3	2	2	1
NE4	285	10	9	7	7	5	2	2	3	2
NE5	315	12	7	4	9	5	4	2	1	2

- 事件特征选择原则

前n条原则 (Top n)

阈值限制 (Threshold =  $\Omega$ )

# 事件模板进化策略

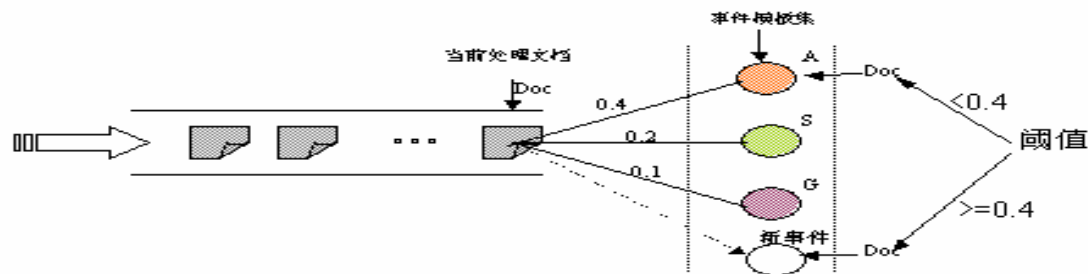


“9-11”事件报道动态变化图

$$INEF_{t_i} = \frac{\log\left(\frac{NUM_{ws} + 0.5}{freq_{t_i}}\right)}{\log(NUM_{ws} + 1)}$$

$$weight_{t_i, NE} = weight_{t_i, NE}' \times INEF_{t_i}'$$

# 事件探测和追踪



输入：新闻文档流。

输出：生成的事件模板集和新闻文档流的事件标示。

1. 文档依次被顺序处理。
2. 第一个文档表示成为事件模板。
3. 每个后续文档和它前面所有的事件模板进行匹配。
4. 根据相似度度量，文档被分配给最相似的事件，然后事件模板需要重新计算。
5. 如果所有的相似度都小于某个阈值，则它可看作新事件种子，并将其表示为模板形式。

# 基于事件距离的相似度计算

$$SIM_{d,NE} = \frac{\sum_{i, d \in \text{know}_i, n \in \text{know}_i} (\text{weight}_{i,d} \times \text{weight}_{i,NE})}{\sqrt{\sum_{i=1}^{|\text{know}_d|} \text{weight}_{i,d}^2} \times \sqrt{\sum_{i=1}^{|\text{know}_{NE}|} \text{weight}_{i,NE}^2}}$$

$$DIS_{\text{time}}(d, NE) = \min \left\{ |time_d - time_{NE}|, |time_{NE} - time_d| \right\}$$

$$SCORE_{d,NE} = \alpha SIM_{d,NE} - \beta DIS_{\text{time}}(d, NE) - 0.4$$

- 阈值判断 (threshold =  $\theta$ )
- 第一名原则 (Top 1)

# 基于“类间距离”提高算法效率

- 基于余弦公式计算事件模板之间的“类间距离”
- 文档与事件之间的相似度差的绝对值来计算“类间距离”

$$Dis(A, B) = |SCORE_{doc, A} - SCORE_{doc, B}|$$

其中，A 和 B 为事件，Doc 代表文档

	A	B	C	D	E	F
Doc	0.8757	0.8543	0.5475	0.1245	0.1434	0.9132

	A	B	C	D	E	F
}	0	0.0124	0.3291	0.7522	0.7333	0.0355
			0.0124	0.7398	0.7209	0.0489
				0.4231	0.4042	0.3555
					0.0189	0.7887
						0.7598

# 实验语料及其算法评价机制

事件 ID	事件名	起始日期	终止时间	文档数
IE1	厦门远华案	2001-03-18	2001-08-28	59
IE2	中英撞机事件	2001-06-27	2001-08-22	56
IE3	美国总统布什访华	2002-02-16	2002-02-26	221
IE4	英国恐怖事件	2002-09-11	2002-09-20	1042
IE5	国家主席江泽民访问英国	2002-10-11	2002-11-01	100
...	...	...	...	...

	相关	不相关
检索到的	a	b
未检索到的	c	d

$$\text{召回率 (Recall)} = R = \frac{a}{a+c}$$

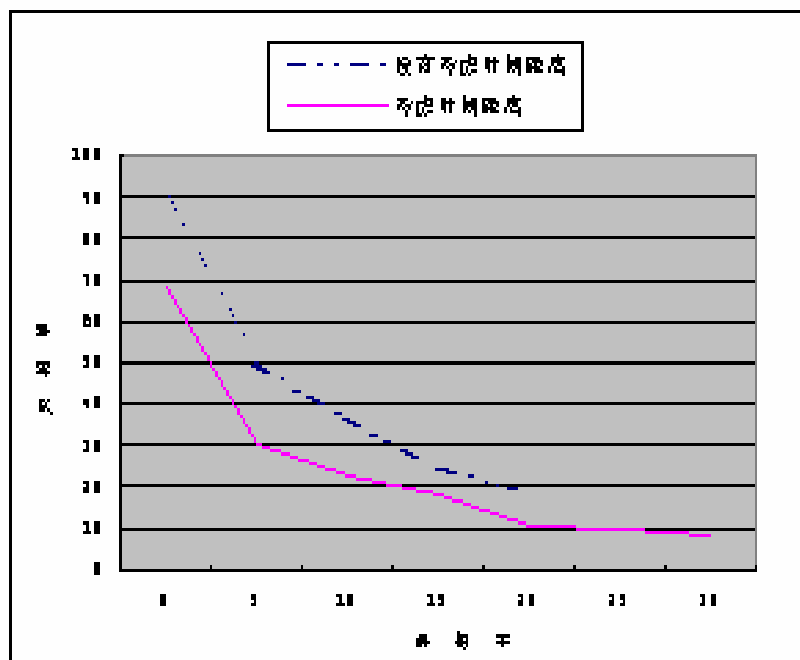
$$\text{正确率 (Precision)} = P = \frac{a}{a+b}$$

$$\text{F1-Measure} = \frac{2PR}{P+R}$$

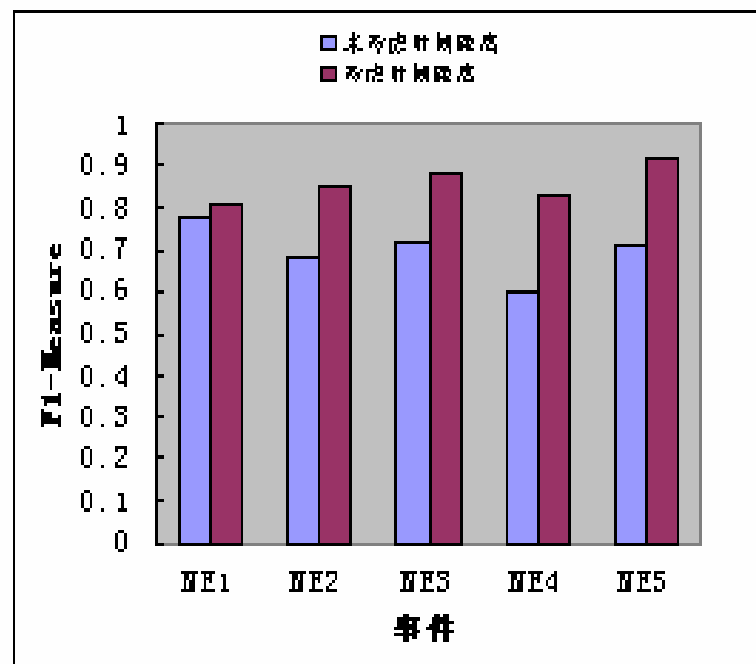
$$\text{失报率 (misses)} = \frac{c}{a+c}$$

$$\text{错报率 (false alarms)} = \frac{b}{b+d}$$

# 实验1—时间距离的影响

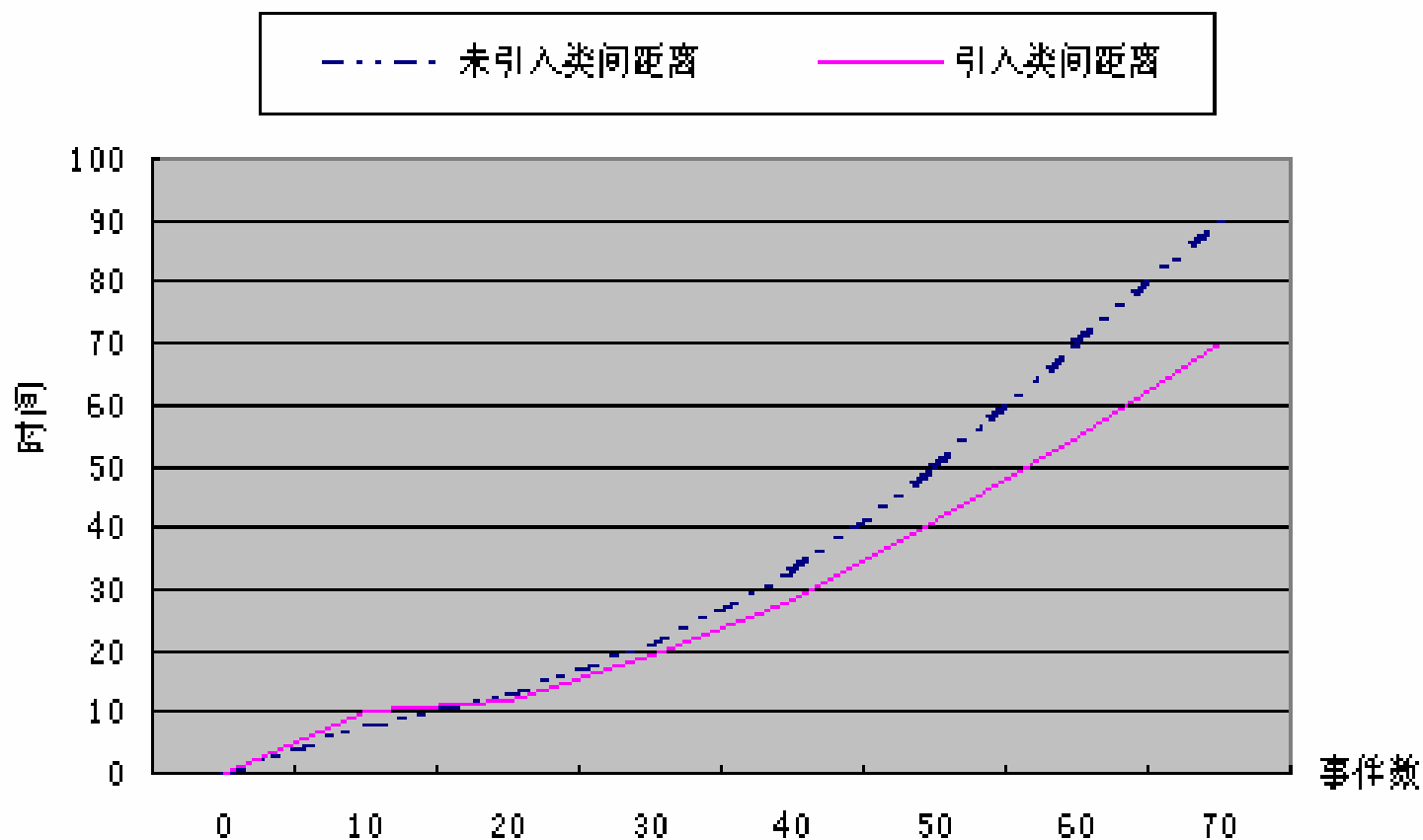


时间距离对事件探测的影响



时间距离对事件追踪效率的影响

# 实验2—“类间距离”对效率的影响



# 问题提出

- 信息社会对能够有效浓缩文本信息的网络自动文摘有着迫切的要求。
- 基于主题的多篇摘要可以有效解决数据冗余
- 事件来龙去脉是对新闻事件深层次知识的挖掘，通过阅读它可以在短时间内了解事件的发生、发展和结束的全过程，以及对人们和社会造成的影响。
- 对基于主题智能化检索的一个有益尝试，提供了一种思路。
- 自动摘要的方法

机械式文摘、基于理解的文摘、复合式文摘

# 新闻文章摘要模型

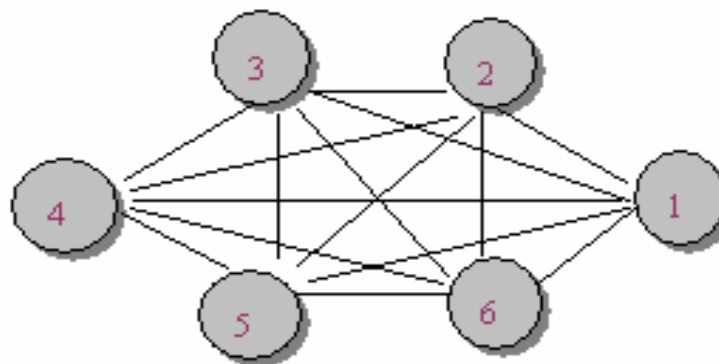
所谓文本的篇章结构信息，主要指文本段落与文章主题的关系以及段落之间的相互联系。

- 各段特征抽取
  - 段落向量空间 $P(w_1, w_2, w_n)$
  - 根据启发信息修正段落向量空间
  - 计算段落之间的语义相似度
  - 根据相似度结果矩阵生成段落关系图
- 频率信息和关键词（词频信息）  
标题词（标题信息）  
提示词（“总而言之”、“由此可见”）  
线索词（“目的”、“关键”、“例如”）  
位置信息

# 段落关系图

0	0.7659	0.9876	0.7756	0.5436	0.8965
	0	0.8769	0.6548	0.6345	0.8423
		0	0.7685	0.7231	0.8634
			0	0.8778	0.6132
				0	0.6874
					0

段落相似度矩阵示例



段落关系图

# 段落重要性评价

## ●评价参数

点的度数  
点的质量  
相关度均值  
相关度方差

## ●评价算法

输入：段落相似度矩阵 M

输出：文摘段(KeyPara)

算法具体描述：

```
{  
    if (NumOfPara <= 2) KeyPara = P1; //段落数小于等于 2，取第一段。  
    if (NumOfPara == 3) {  
        if (DegreeOfPara 不相等) //段落的度数不等  
            KeyPara = 度数大的段；  
        Else  
            KeyPara = 均值大的段；  
    }  
    if (NumOfPara > 3) {  
        if (DegreeOfPara 不相等) //段落的度数不等  
            KeyPara = 度数大的段；  
        Else if (均值相差的范围在 0.01 中)  
            Then KeyPara = 方差小段；  
            else KeyPara = 均值大的段  
    }  
}
```

# 摘要润色

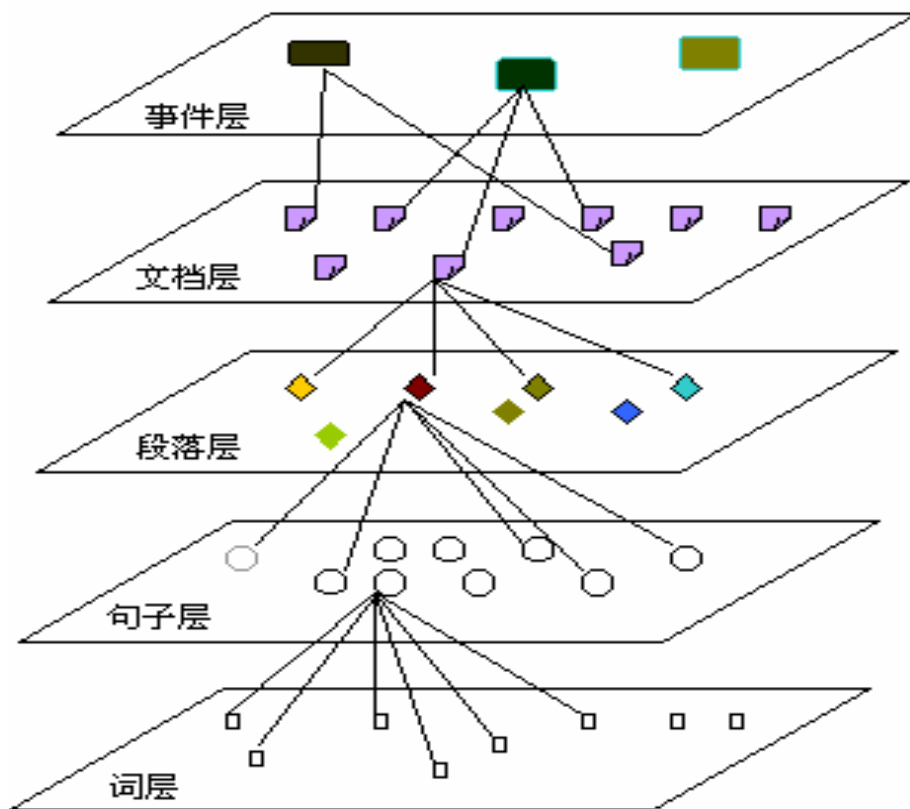
- 非连通图：子主题与事件主题的相关性

$$P(\text{relvance}(s_k, M_E)) = P(s_k | M_E) = \left( \prod_{t \in S_1} \frac{\text{weight}(t, M_E) + 0.01}{1.01 * |M_E|} \right)^{\frac{1}{|S_1|}}$$

- 摘要润色： Title标记、提示词、引领段

```
输入：文本数据集 DataS、文本对应摘要集 SummaryS。  
输出：润色后的摘要集合 SummaryS。  
算法描述：  
REPEAT UNTIL 处理完所有的数据文件{  
  IF 文章是一篇多标题的文章  
  THEN 我们就取子标题的前一段和所有的子标题作为文摘  
  ELSE {  
    由文本中取出文摘段的第一句和最后一句 first,last  
    IF first 含有表示结构的词语  
    THEN 取出这个结构的上一段以及结构成员的第一句  
    ELSE {  
      IF 在 first 中发现了加、减权重的词语  
      THEN 对这段中所有词语进行权重修改  
      ELSE 对段首和段尾中的词语进行润色  
    }  
  }  
}
```

# 概念层次关系



# 事件来龙去脉组织策略

- 好的事件来龙去脉满足的条件

将报道同一事件的所有子事件尽量包含。  
所有子事件按照报道顺序排列。  
不要包含与该事件无关的信息。  
避免冗余信息。  
流畅描述。

- 组织事件来龙去脉的策略

相关性  
新颖性  
次序  
粒度

# 新颖性探测模型

$$\begin{aligned}
 P(\text{novel}_1(s_k)) &= P(H(s_k) \neq H(s_i), \forall i < k) \\
 &= \left[ \prod_{i < k} (1 - p(H(s_k) = H(s_i))) \right]^{\frac{1}{k-1}} \\
 &= \left[ \prod_{i < k} \left( 1 - \left[ \prod_{w \in s_i} \frac{\text{weight}(w, s_i) + 0.01}{1.01 \cdot |s_i|} \right]^{\frac{1}{|s_i|}} \right) \right]^{\frac{1}{k-1}}
 \end{aligned}$$

$$\begin{aligned}
 P(\text{novel}_2(s_k)) &= P(H(s_k) \neq H(s_i), \forall i < k) \\
 &= \left[ \prod_{i < k} (1 - p(H(s_k) = H(s_i))) \right]^{\frac{1}{k-1}} \\
 &= \left[ \prod_{i < k} \left( 1 - \frac{\vec{s}_k \cdot \vec{s}_i}{\|\vec{s}_k\| \times \|\vec{s}_i\|} \right) \right]^{\frac{1}{k-1}}
 \end{aligned}$$

- “当地时间9月11日上午接近9时（北京时间接近21时），美国一架飞机撞上纽约世界贸易中心大楼”
- “当地时间9月11日9:05（北京时间11日21:05），第二架飞机撞上纽约世界贸易中心大楼。”

权重个性化策略： 时间短语、数词等词性权重加强

# 事件来龙去脉算法

输入：数据集 DataS，其中数据元素为 DataModelOfNewsStory；显示次序 Order；  
时间范围[BeginTime~EndTime]；

输出：事件来龙去脉 CauseEvent

算法描述

Begin

```
String summary[]; //临时变量
Int SummaryVal = 0;
Sort (DataS, EventTime); //根据 EventTime 对数据集 DataS 排序
While (DataS[j].EventTime < BeginTime) continue; //选择超过起始时间的数据
    summary[SummaryVal++] = DataS[j].Summary;
For (int j = 1; j <= |DataS|; j++){
    If (DataS[j].EventTime > EndTime) Break; //选择低于终止时间的数据
    Else If (IsNovel (DataS[j].Summary)) //判断当前摘要的新颖性
        summary[SummaryVal++] = DataS[j].Summary; //赋值
}
For (int j = 0; j < SummaryVal; j++)
    CauseEffect = CauseEffect + summary[j] + "\r\n";
```

End

# 事件来龙去脉结果评测

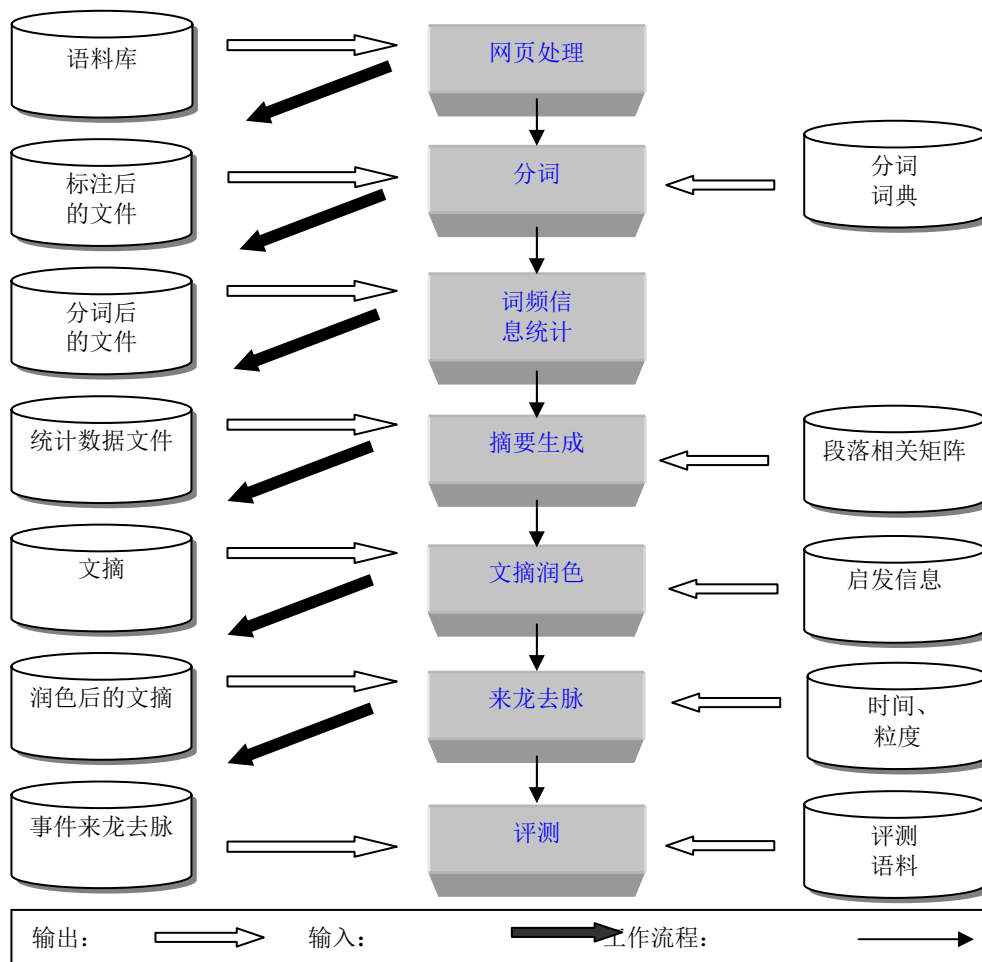
- 常用的评测类型
  - 与经典的摘要生成算法进行比较
  - 人工根据标准进行判断
  - 摘要是否能够替代原文的功能
  - 通过比较计算机生成的摘要和人工摘要的一致性

- 事件来龙去脉评测模型
  - $$recall = \frac{|H(S_T) \cap E|}{|E|}$$
  - $$s - precision = \frac{|H(S_T) \cap E|}{|H(S_T)|}$$
  - $$d - precision = \frac{IsTrue(|H(S_i)| > 0) + \sum_{i=2} IsTrue(|H(S_i)| > |H(S_{i-1})|)}{|S_T|}$$

# 实验评测语料

事件 ID	事件名	起始日期	终止时间	文档数	涉及的子事件数
NE1	厦门远华案	2001-03-18	2001-08-28	59	243 ( 4.1/篇 )
NE2	中美撞机事件	2001-06-27	2001-08-22	56	189 ( 3.4/篇 )
NE3	美国总统布什访华	2002-02-16	2002-02-26	221	834 ( 3.8/篇 )
NE4	美国恐怖事件	2002-09-11	2002-09-20	1042	2010 ( 1.8/篇 )
NE5	国家主席江泽民访美	2002-10-11	2002-11-01	100	425 ( 4.2/篇 )
NE6	伊拉克战争	2003-01-27	2003-07-30	883	2020 ( 2.2/篇 )
NE7	伊朗地震	2003-08-17	2004-01-04	352	527 ( 1.5/篇 )
NE8	神州五号	2003-10-09	2003-11-03	1020	3432 ( 3.3/篇 )
NE9	萨达姆被捕	2003-12-14	2004-01-04	538	1854 ( 3.4/篇 )
...	...	...	...	...	...

# 系统实现

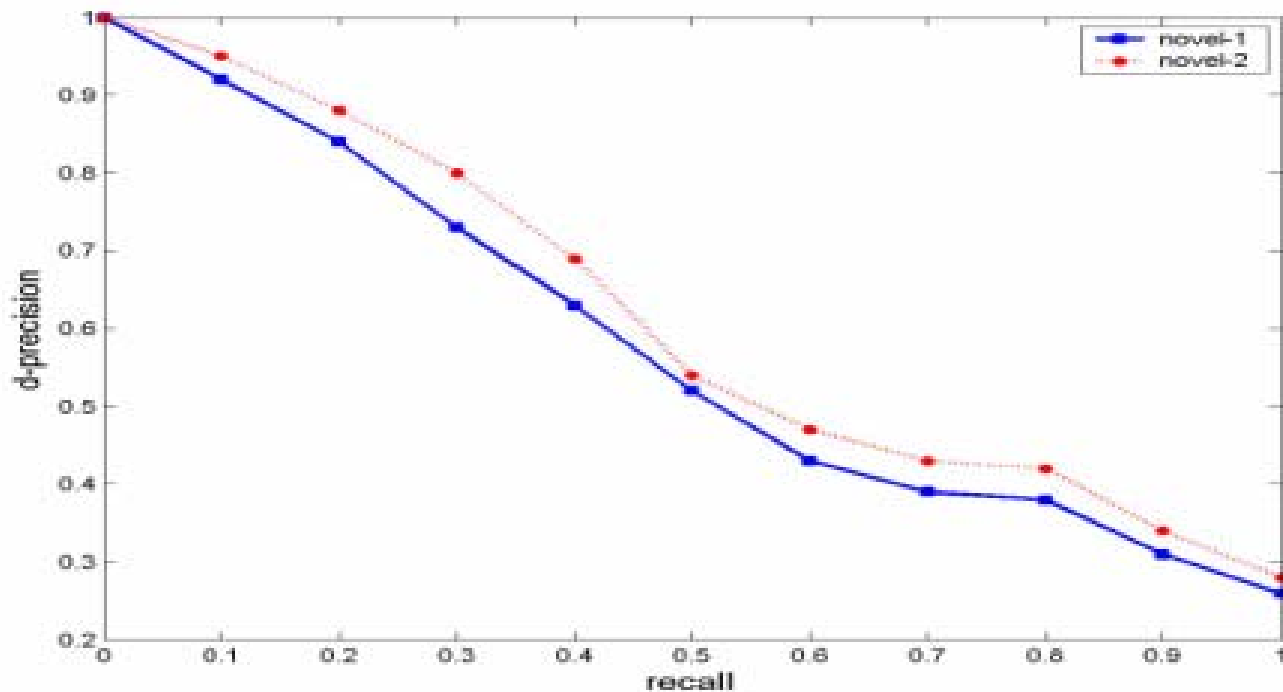


# 单篇摘要的实验结果分析

- 评测策略：
  - 子主题覆盖、与事件主题相关性、以及可读性等方面
- 评价结果：

事件 ID	事件名	文档数	错误文档数	差错率	正确率
NE1	厦门远华案	59	4	6.7%	93.3%
NE2	中美撞机事件	56	3	5.4%	94.6%
NE3	美国总统布什访华	221	17	7.9%	92.1%
NE4	美国恐怖事件	1042	52	4.8%	95.2%
NE5	国家主席江泽民访美	100	11	10.6%	89.4%
NE6	伊拉克战争	883	73	8.3%	91.7%
NE7	伊朗地震	352	18	5.2%	94.8%
NE8	神州五号	1020	115	11.3%	88.7%
NE9	萨达姆被捕	538	41	7.6%	92.4%

# 事件来龙去脉结果分析



# 图像内容理解

---

- 特征抽取
- 语义映射
- 内容理解

# 颜色特征抽取

---

- 颜色直方图
- 累积颜色直方图
- 颜色矩
- 颜色集。

# 颜色特征抽取

---

- 颜色直方图
- 累积颜色直方图
- 颜色矩
- 颜色集。

# 纹理特征抽取

---

- 共生矩阵
- 马尔可夫随机场
- 分形表示
- 小波变换
- ▲ **Gabor**变换
- ▲ 正交和双正交变换
- ▲ 树结构小波变换

# 形状特征提取

---

分类:

- 基于边界
- 基于区域

形状特征方法:

- 傅里叶描述子
- 矩不变量
- 有限元方法
- 小波描述子

# 图像分割

---

- 颜色、纹理和形状特征都依赖好的分割
- 不同特征要求分割的精度不同
- 分类：
  - ★ 监督分割
  - ★ 无监督分割

# 降维

- 图象检索中特征向量的维数一般非常高，但其内含维数（**embedded dimension**）却要低的多
- 降维方法：
  - ◆ **KL**变换
  - ◆ 聚类方法
  - ◆ 主成分分析（**PCA**）
  - ◆ 独立成分分析(**ica**)
  - ◆ 流形学习

# 多维索引技术

---

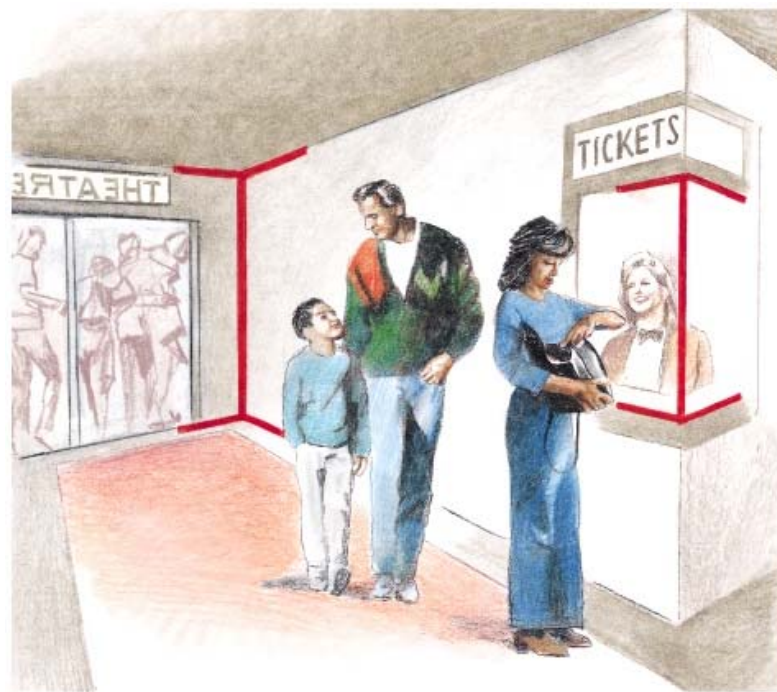
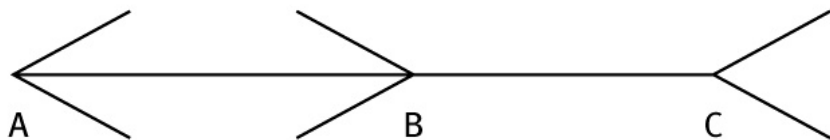
- 桶算法
- k-d树和优先k-d树
- 八叉树
- K-D-B树
- hB树
- R-树、R<sup>+</sup>-树、R\*-树

# 格式塔理论

格式塔（Gestalt）心理学诞生于1912年。它强调经验和行为的整体性，反对当时流行的构造主义元素学说和行为主义“刺激—反应”公式，认为整体不等于部分之和，意识不等于感觉元素的集合，行为不等于反射弧的循环。尽管格式塔原理不只是一种知觉的学说，但它却导源于对知觉的研究，而且一些重要的格式塔原理，大多是由知觉研究所提供的。

# 格式塔心理学

- 知觉的形成是基于先验知识的刺激组织



# 格式塔心理学

- 格式塔理论认为，知觉问题涉及到比较和判断。

Picasso's  
panting

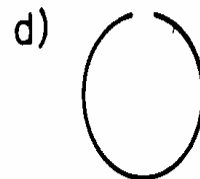
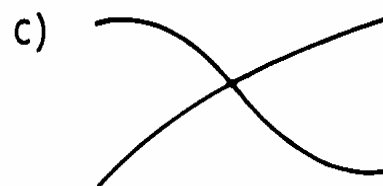
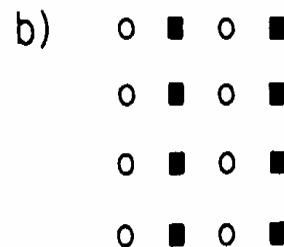
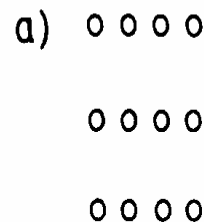


# 知觉组织法则

- Wertheimer 等格式塔心理学家说明了人类知觉组织能归结为几个重要的法则：

## 1) 邻近性。

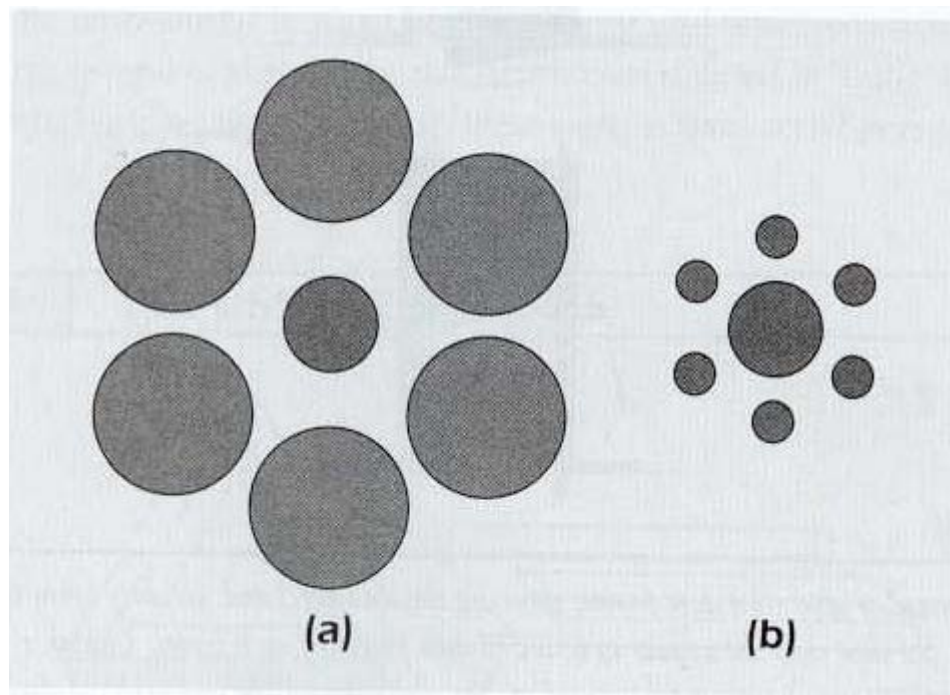
彼此紧密接近的刺激物比相隔较远的刺激物有较大的组合倾向。接近可能是空间的，也可能是时间的。按不规则的时间间隔发生的一系列轻拍响声中，在时间上接近的响声倾向于组合在一起。由于接近而组合成的刺激不必都是同一种感觉形式的；例如，夏天下雨时，雷电交加，我们就把它们知觉为一个整体，即知觉为同一事件的组成部分。



# 知觉组织法则

## ● 2) 相似性

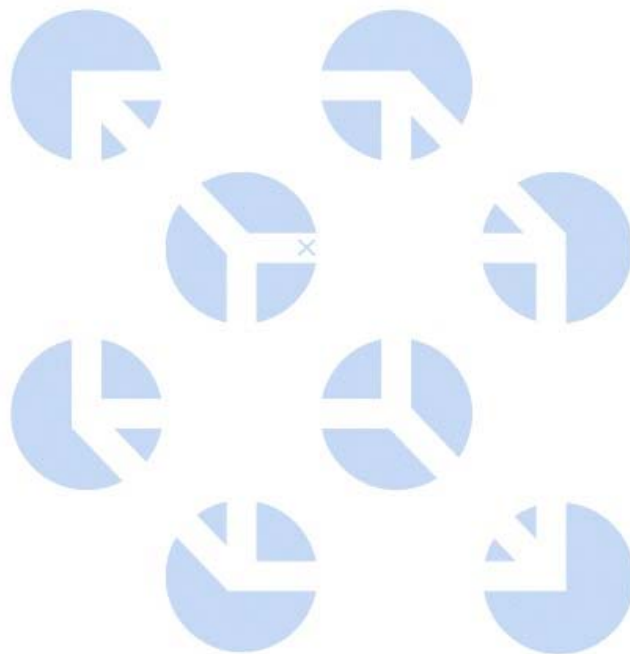
彼此相似的刺激物比不相似的刺激物有较大的组合倾向。相似意味着强度、颜色、大小、形状等这样一些物理属性上的类似。俗话说：“物以类聚，人以群分”，也就包含这种原则。”



# 知觉组织法则

## 3) 好的连续性

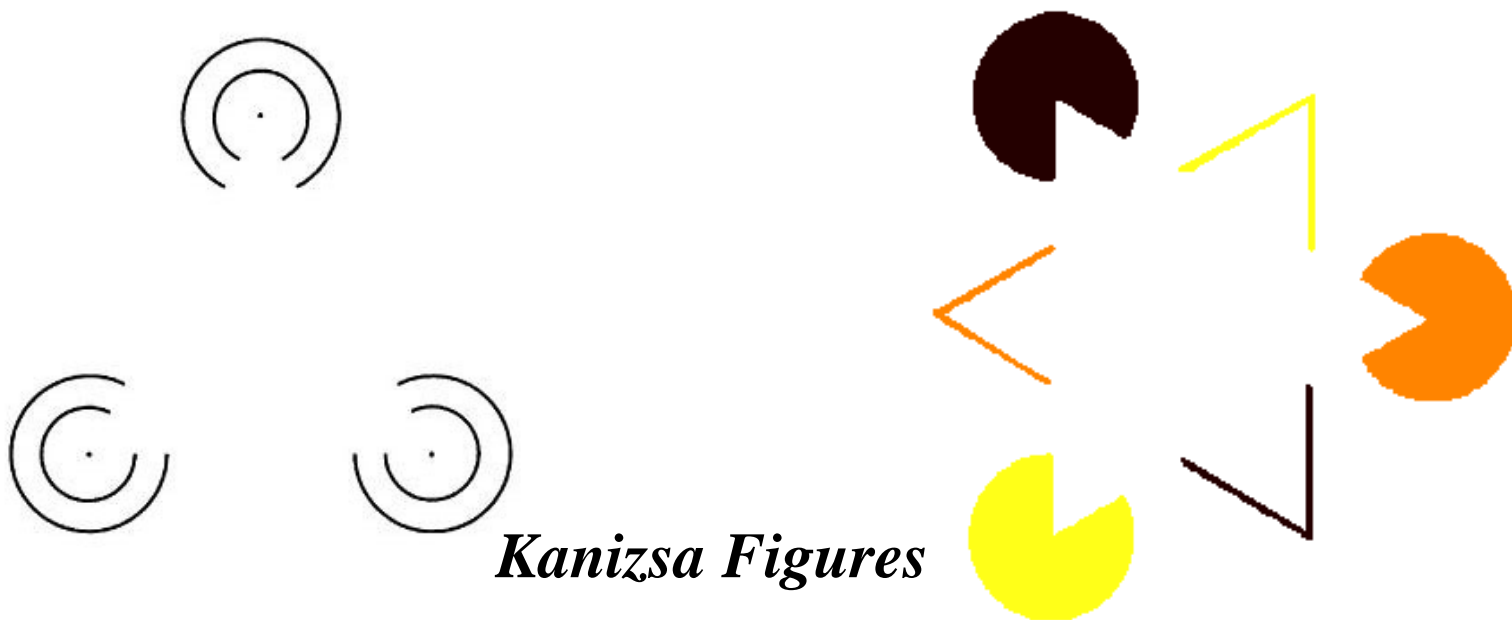
一种倾向，即一些成分和其它成分以这样一种方式连接在一起，以便有可能使一条直线、一条曲线或者一个动作沿着已经确立的方向继续下去。



# 知觉组织法则

## 4) 封闭性

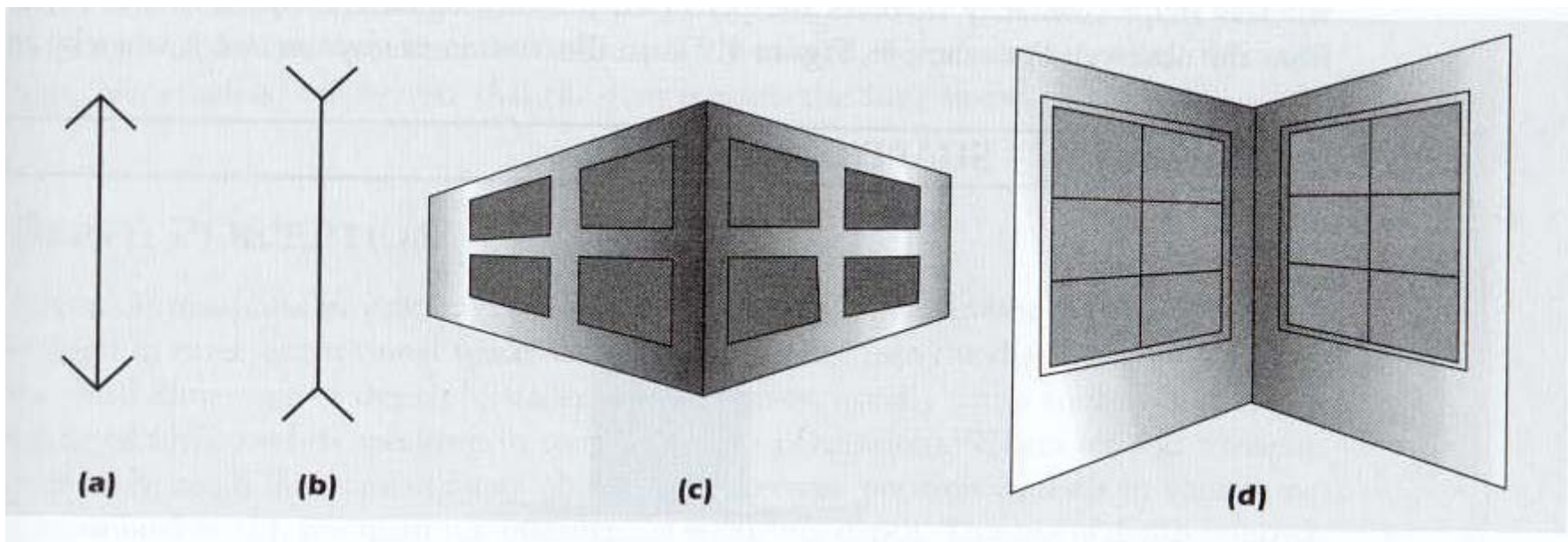
一些成分以这样一种方式组合，以便有助于形成一个更加紧密和更加完整的图形。



# 知觉组织法则

## 5) 对称性

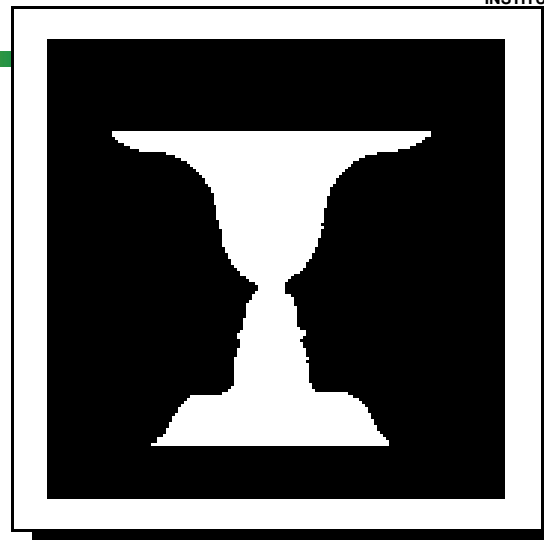
一种条件，有利于这样一些组合，即导致对称或平衡的整体而不是非对称的整体。



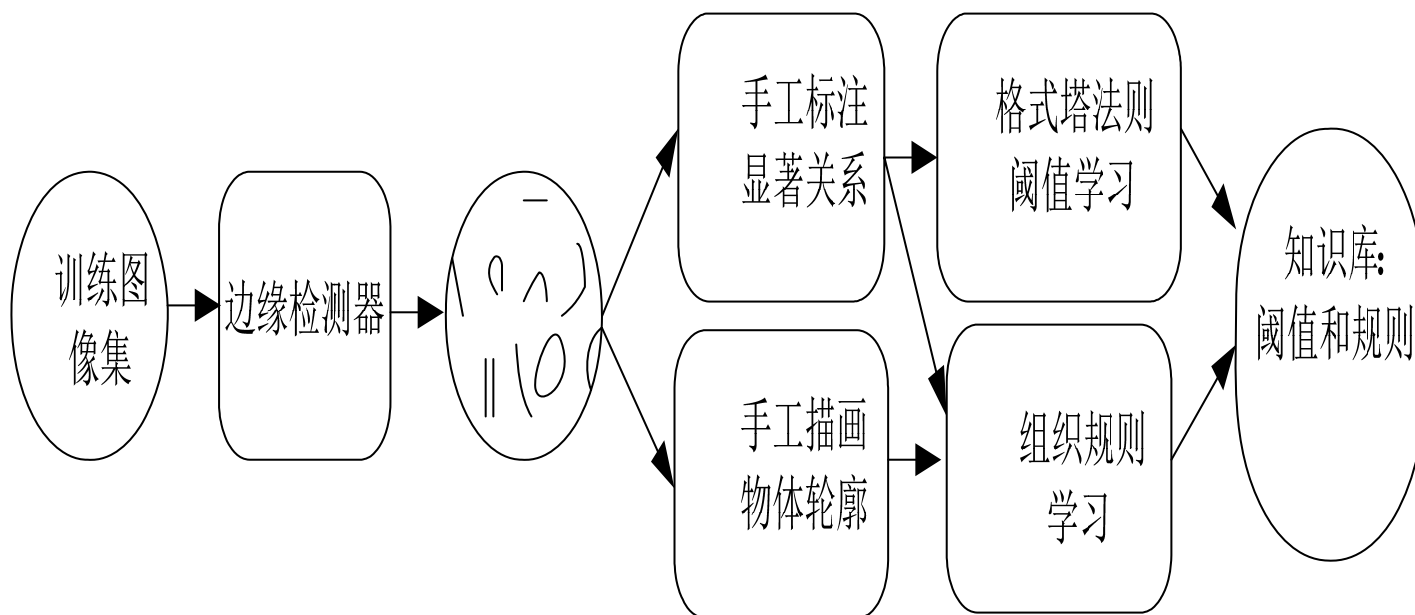
# 知觉组织法则

6) 共同的命运：在一个共同的方向上移动或变化的那些成分的组合，而与视域里朝其它方向运动或变化的成分相区别。这个组合原则本质上是相似组合在运动物体上的应用，它是舞蹈设计中的一个重要手段。

7) 常见的形状



# 格式塔知觉组织法则的量化模型



# 量化过程

- 预处理. 求出图像中的边线段及其各边的属性值;
- 定义边段间的几何距离、亮度差异及方向差异;
- 对边线段的属性值进行属性概化处理,对格式塔法则进行语义模糊化;
- 分析各法则与边线段各差异值间的相关度,以确定三个格式塔法则的影响因子;
- 估计概化域值,使得达到最大的可信水平

# 量化过程

1. 预处理. 求取图像中的边线段。每个边线段是一条直线或是一条常曲率弧，用  $seg$  表示。

2. 定义边段间的几何距离、亮度差异及方向差异：

1) 几何距离

2) 亮度距离

3) 方向距离:  $theta$  ( $\theta$ )

# 量化过程

## 3. 属性概化处理和语义模糊化处理

### ● 图像属性概化结果

图像属性 \ 数字区段	0	1	...	n
<i>mindis</i>	[0,0.03]	(0.03,0.06]	...	(0.72,+∞)
<i>theta</i>	[0,0.087]	(0.087,0.174)	...	(3.045,pi)
<i>grayvar</i>	[0,5]	(5,10]	...	(250,255]
<i>Ymag</i>	[0,0.05]	(0.05,0.10]	...	(1.05,+∞)
<i>Ywidth</i>	[0,0.05]	(0.05,0.10)	...	(1.05,+∞)

# 量化过程

语义模糊化结果

$$P = \{ \text{very near, near, not near} \}$$

$$C = \{ \text{continuous, noncontinuous} \}$$

$$S = \{ \text{similar, not similar} \}$$

# 量化过程

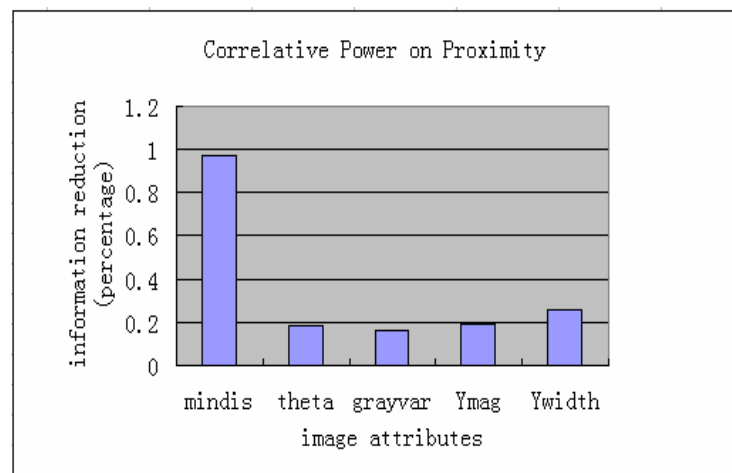
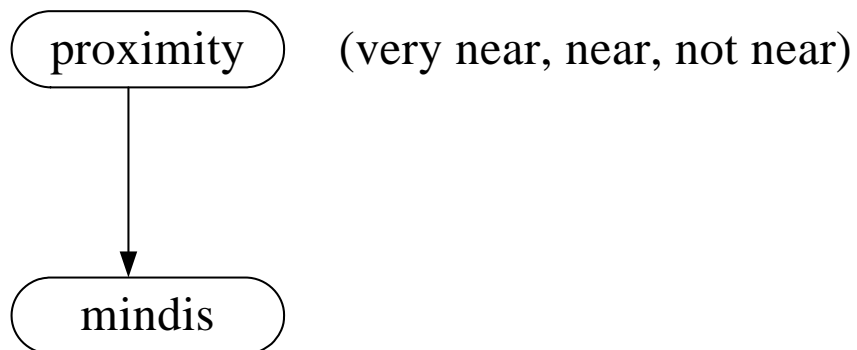
## 4. 确定格式塔法则的影响因子

邻近性

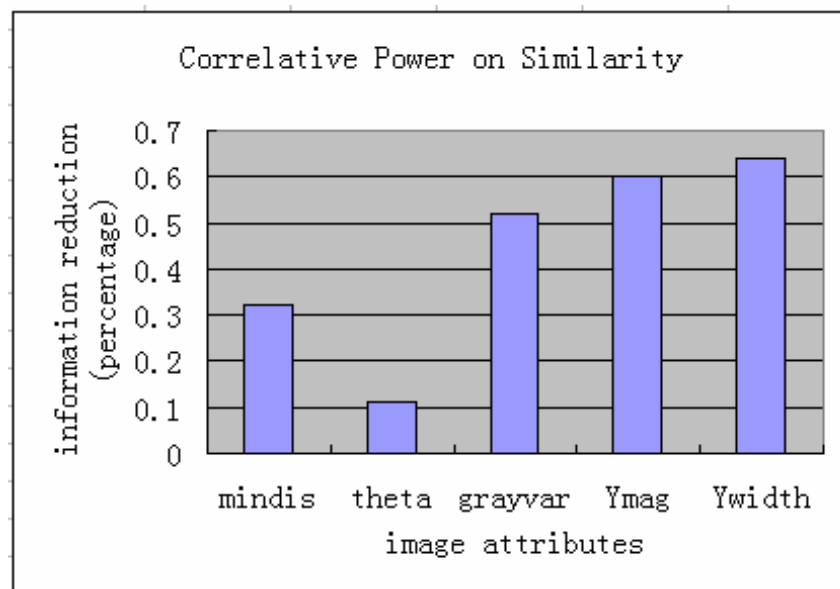
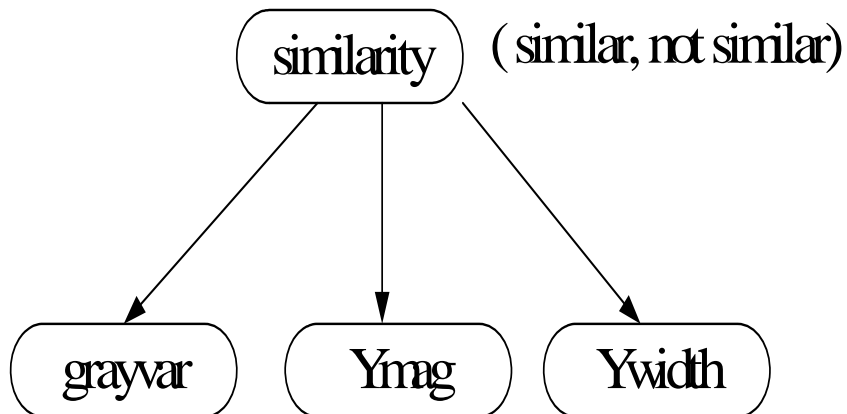
相似性

连续性

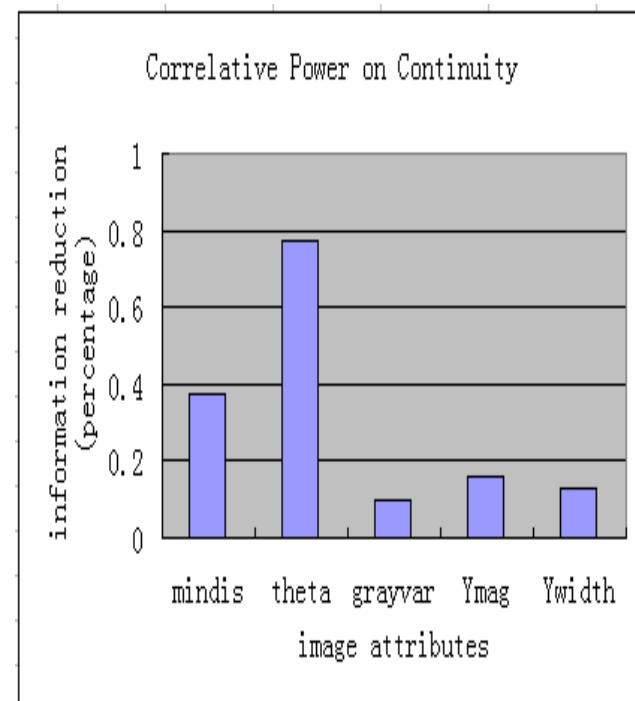
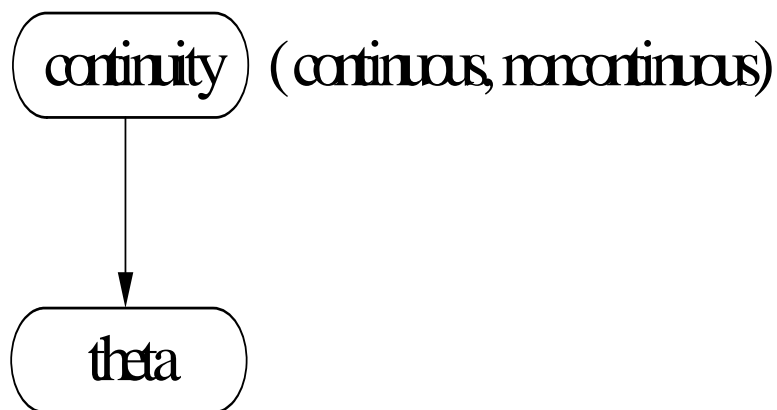
# 相关度分析\_邻近性



# 相关度分析\_相似性



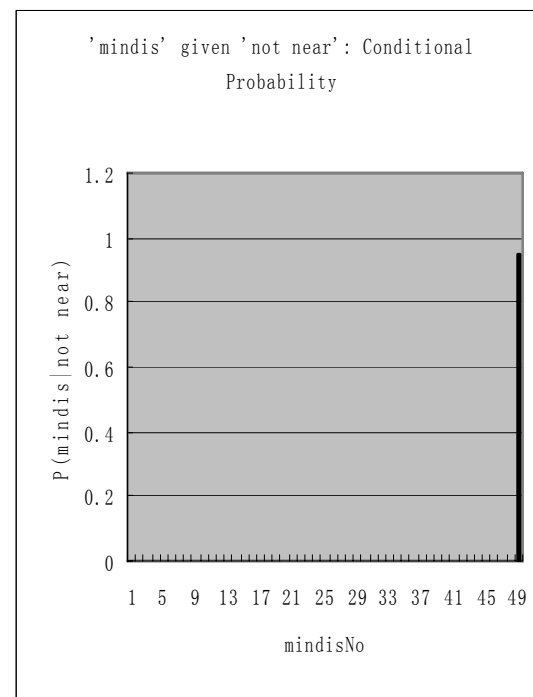
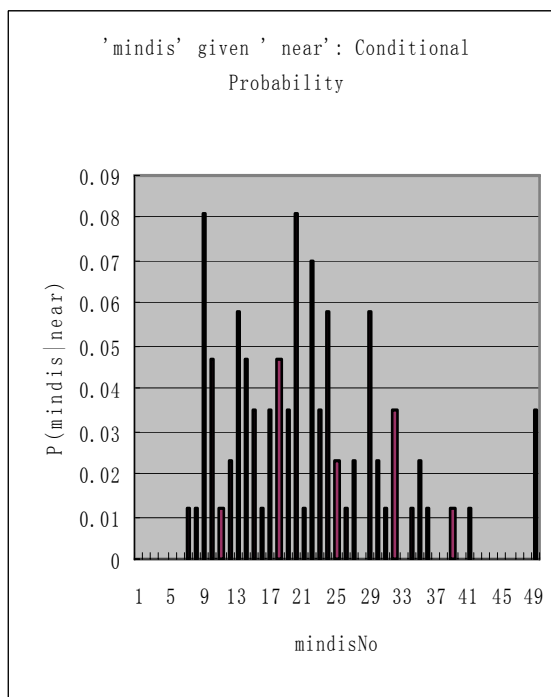
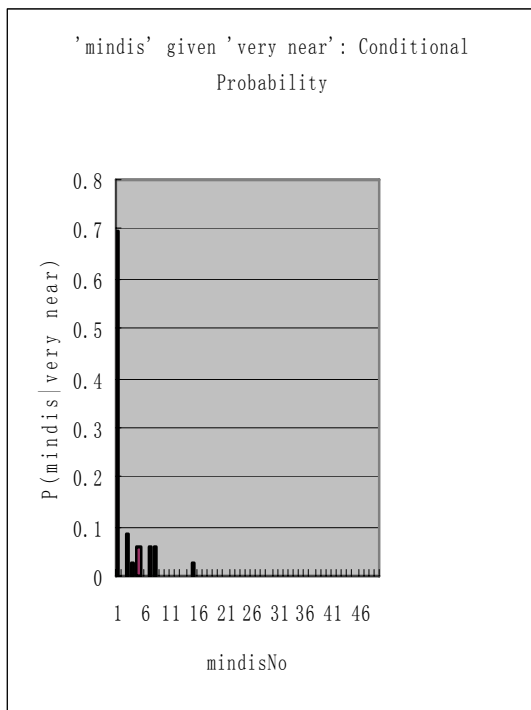
# 相关度分析\_连续性



# 量化过程

## 5. 估计概化域值,使得达到最大的可信水平.

- “很邻近”、“邻近”、“不邻近”时,“mindis”处于各区段的概率



# 量化域值结果\_邻近性

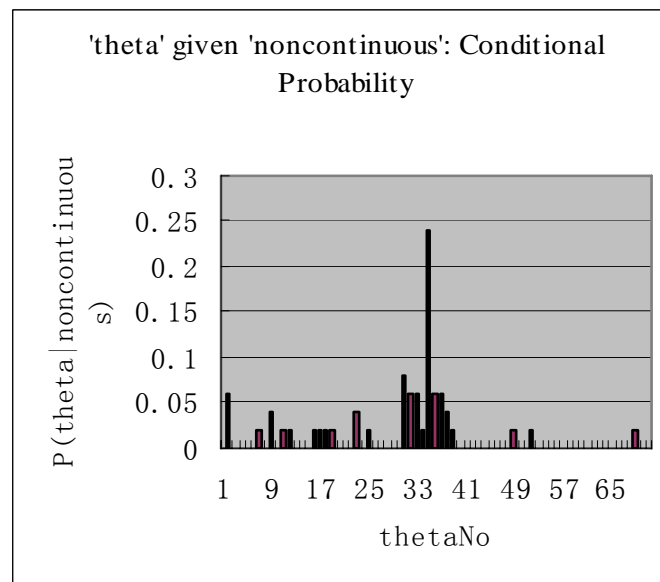
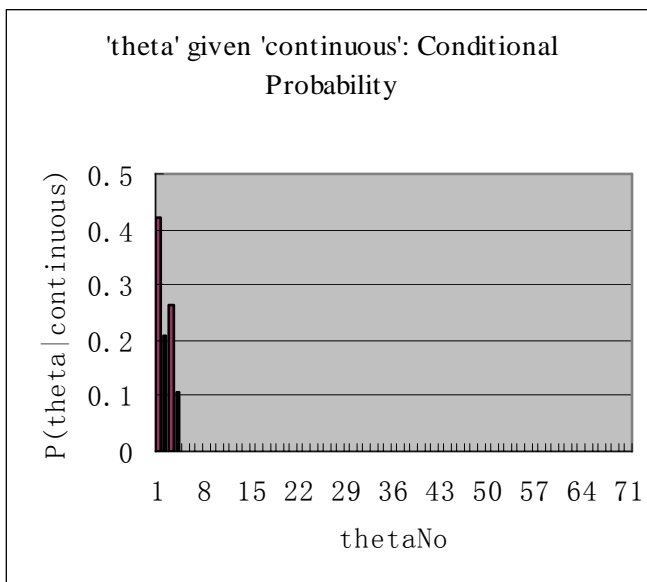
$$p\left(0 \leq \text{mindis} \leq 0.12 \mid \{seg_i, seg_j\} \in \text{very near}\right) \approx 0.97$$

$$p\left(0.12 < \text{mindis} \leq 0.48 \mid \{seg_i, seg_j\} \in \text{near}\right) \approx 0.87$$

$$p\left(\text{mindis} > 0.48 \mid \{seg_i, seg_j\} \in \text{not near}\right) \approx 0.98$$

# 量化过程

- “连续的”、“不连续的”时，“theta”处于各区段的概率



# 量化域值结果\_连续性

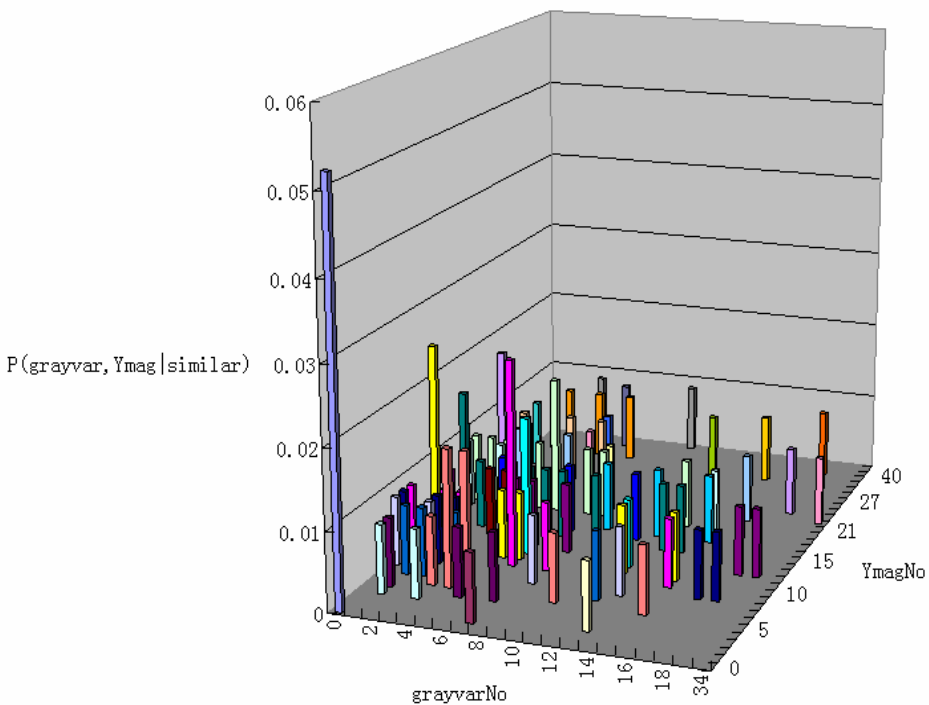
$$p\left(0.0 \leq \theta \leq 0.132 \mid \{seg_i, seg_j\} \in \text{continuous}\right) \approx 1.0$$

$$p\left(0.132 < \theta \leq \pi \mid \{seg_i, seg_j\} \in \text{noncontinuous}\right) \approx 0.95$$

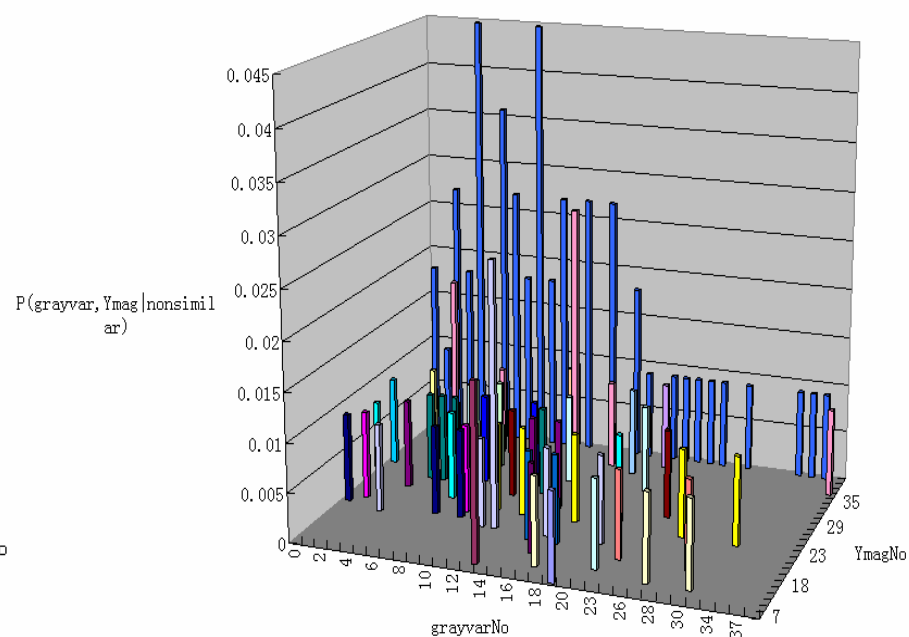
# 量化过程



'grayvar' and 'Ymag' given 'similar': Conditional probability



'grayvar' and 'Ymag' given 'nonsimilar': Conditional Probability



# 量化域值结果\_相似性

$$p\left(0 \leq \text{grayvar} \leq 35 \text{ and } 0 \leq Y_{\text{mag}} \leq 0.775 \mid \{seg_i, seg_j\} \in \text{similar}\right) \approx 0.643$$

$$p\left(35 \leq \text{grayvar} \leq 95 \text{ and } 0 \leq Y_{\text{mag}} \leq 0.5 \mid \{seg_i, seg_j\} \in \text{similar}\right) \approx 0.261$$

$$p\left(0 \leq \text{grayvar} \leq 35 \text{ and } 0 \leq Y_{\text{mag}} \leq 0.775 \mid \{seg_i, seg_j\} \in \text{dissimilar}\right) \approx 0.080$$

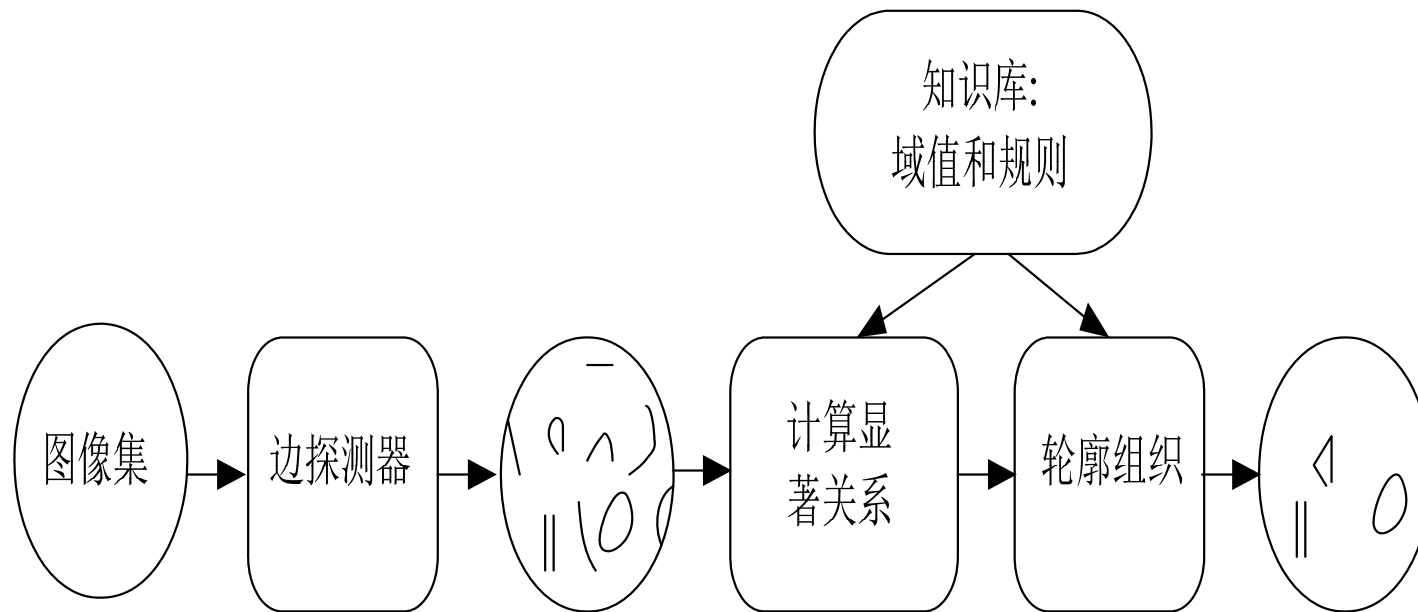
$$p\left(35 \leq \text{grayvar} \leq 95 \text{ and } 0 \leq Y_{\text{mag}} \leq 0.5 \mid \{seg_i, seg_j\} \in \text{dissimilar}\right) \approx 0.054$$

# 学习组织规则

proximity	continuity	similarity	Confidence (Samecontour)
0	0	0	0
0	0	1	0
0	1	0	0
0	1	1	0
1	0	0	0
1	0	1	0.13
1	1	0	0.2
1	1	1	0.81
2	0	0	0.1
2	0	1	0.87
2	1	0	0.44
2	1	1	0.98

如果两个边是邻近的、连续的和相似的，它们最可能处于同一轮廓上；  
如果两个边是强邻近的和相似的，它们最可能处于同一轮廓上。

# 基于量化信息和组织规则的知觉组织模型



## 图结构 $G = (N, A)$

$$p(n_i, n_j) = \text{proximity}(seg_i, seg_j) + \text{similarity}(seg_i, seg_j) + \text{continuity}(seg_i, seg_j)$$

从  $n_1$  到  $n_k$  的一条通路的总收益为

$$P = \sum_{i=2}^k p(n_{i-1}, n_i)$$

$$r(n) = g(n) + p(n, m) \quad \text{式 (1)}$$

# 基于量化阈值和组织规则的知觉组织算法

求图G的各连通分量，并生成其连通子图。对每个连通子图：

Step1. 选择起始结点，这里选度数最大的结点。

Step2. 将起始结点标记为OPEN并置 $g(s)=0$ ，通路长度。

Step3. 如果不存在OPEN结点：若路径长度 $len = 1$ ，找到一条非闭合的最大通路，否则失败退出；

如果存在OPEN结点，继续下一步。

Step4. 将根据上式算得的估计收益 $r(n)$ 为最大的OPEN结点标记为CLOSE。

Step5. 如果 $n$ 是起始结点（目标结点等于起始结点），且路径长度 $len \geq 3$ ，找到通路（可由 $n$ 借助指针上溯至 $s$ ）退出，否则继续。

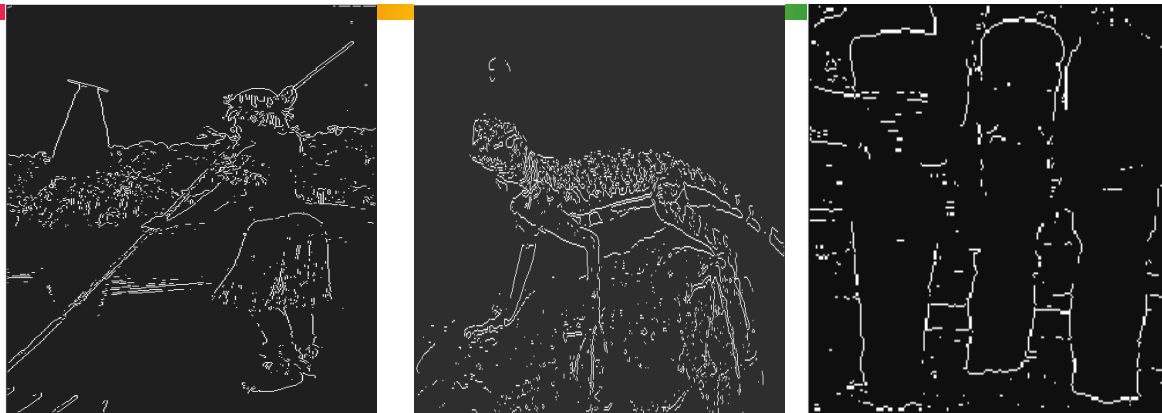
Step6. 展开结点 $n$ ，得到它的所有子结点（如果没有子结点，返回Step3）。

Step7. 如果某个子结点还没有标记，置 $r(n_i) = g(n) + p(n, n_i)$  标记它为OPEN并将指向它的指针返回到结点 $n$ 。

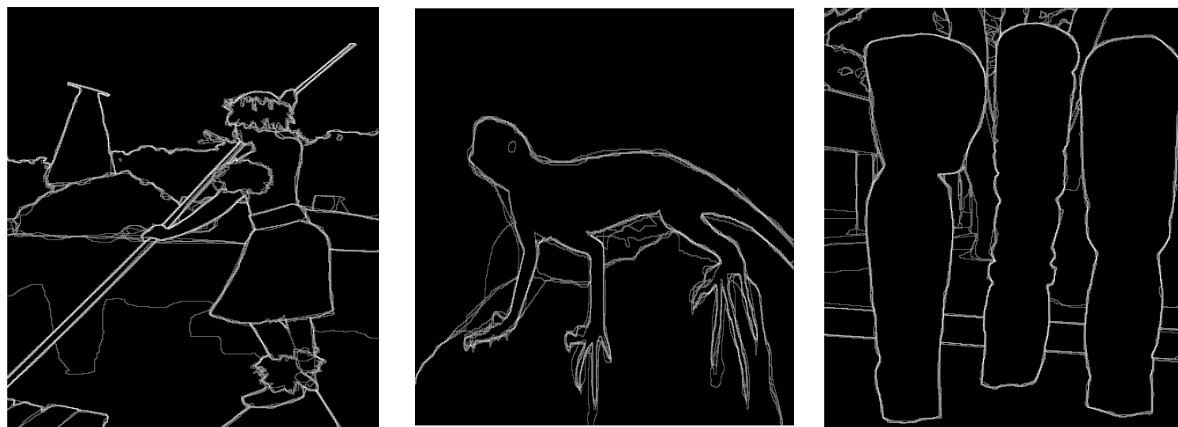
Step8. 如果子结点已标记为OPEN或CLOSE，根据 $g'(n_i) = \max[g(n_i), g(n) + p(n, n_i)]$ 更新它的值。将其 $g'$ 值增大的CLOSE子结点标记为OPEN，并将原指向所有其 $g'$ 值增大的子结点的指针重指向 $n$ 。返回Step3。

# 实验结果





用'sobel' 算子得到的边缘检测结果

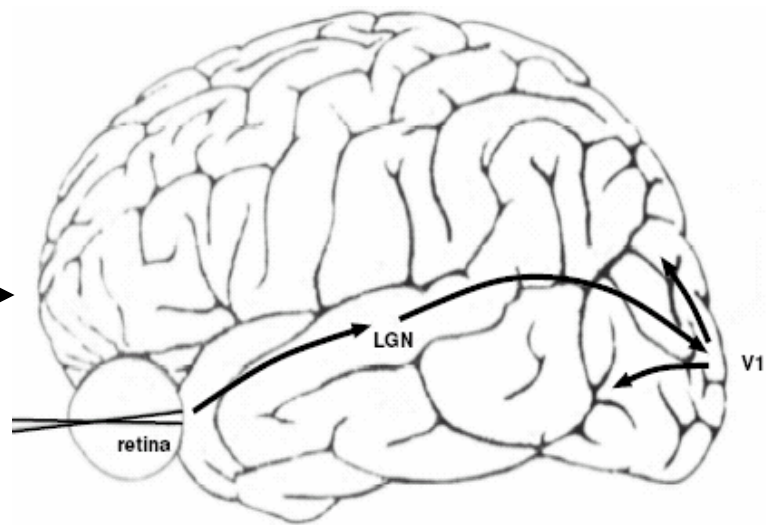


# 有效编码

## 研究动机



怎么联系？



## 稀疏编码发展历程

- 视觉感知的目标就是产生一个外部输入信号的有效表示（Attneave 1954）；
- 有效编码假设：初级视皮层神经细胞的主要功能就是去除输入刺激的统计相关性（Barlow 1961）；
- 稀疏编码理论：初级视皮层神经细胞对自然图像刺激进行稀疏响应（Olhausen & Field, Nature, 1996）。

# 有效编码

## ❖ 机理测试

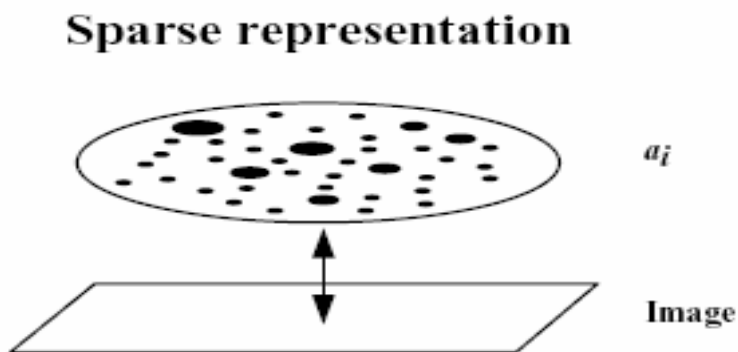
从生物机理上，在自然图像刺激条件下检测神经细胞的响应特性

- Vinje 2000年在science上类似的成果表明：短尾猿视皮层（V1区）用稀疏编码有效表示自然场景
- Nirenberg 等2001年在nature上发表的研究结果表明，在冗余性测度和自然刺激条件下，一组视网膜神经节对外界刺激独立编码。

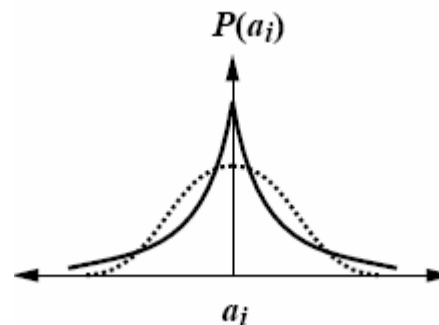
## ❖ 模型仿真

利用自然图像的统计特性，建立模型模拟早期视觉系统的神经细胞编码机制

# 稀疏编码模型



在作用于同一刺激的神经元群中  
仅有少数的神经元被激活；



某一神经元响应值的分布具有稀疏特性：即大部分值都分布在0附近，而激活的概率很少；

# 稀疏编码模型

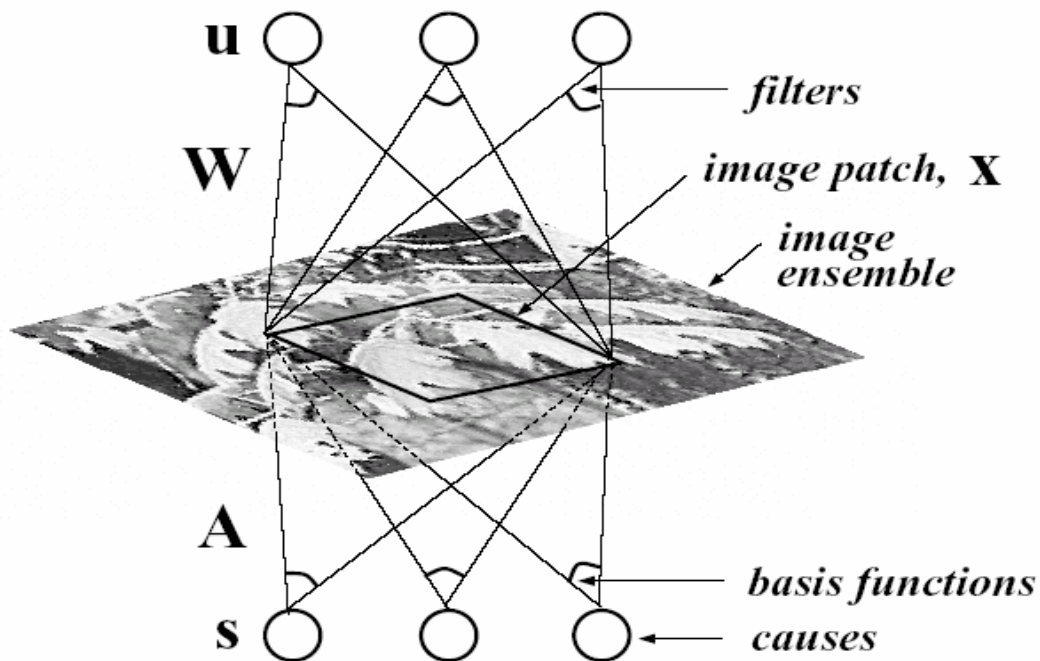
- 神经生物学角度：测试神经元细胞在自然图像刺激下响应的统计特征
  - ✓ *Sparse coding and decorrelation in primary visual cortex during natural vision. Science, 287:1273-1276, Feb 2000.*
  - ✓ *Retinal ganglion cells act largely as independent encoders. Nature, 411:698-701, June 2001.*
- 计算模型角度：利用自然图像的统计特性，构造初级视皮层神经元的  
信息处理模型
  - ✓ *基于稀疏性准则的模型*  
*Emergence of simple-cell receptive field properties by learning a sparse code for natural images. Nature, 381:607-609, 1996.*
  - ✓ *基于独立性条件的模型*  
*The 'independent components' of natural scenes are edge filters. Vision Research, 37(23):3327-3338, 1997.*

# 有效编码研究机制 - 模型仿真

基于外部环境刺激的统计特性，根据特定的优化准则建立响应模型，模拟初级视皮层中神经细胞的响应特性（神经生理学）。

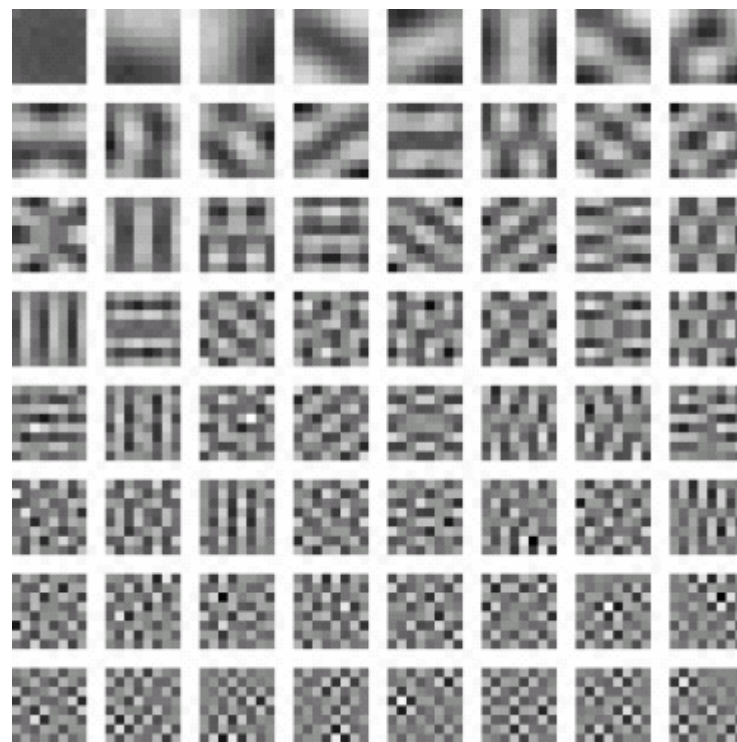
$$X = AS$$

$$S = WX$$



# 基于二阶统计的线性有效编码模型

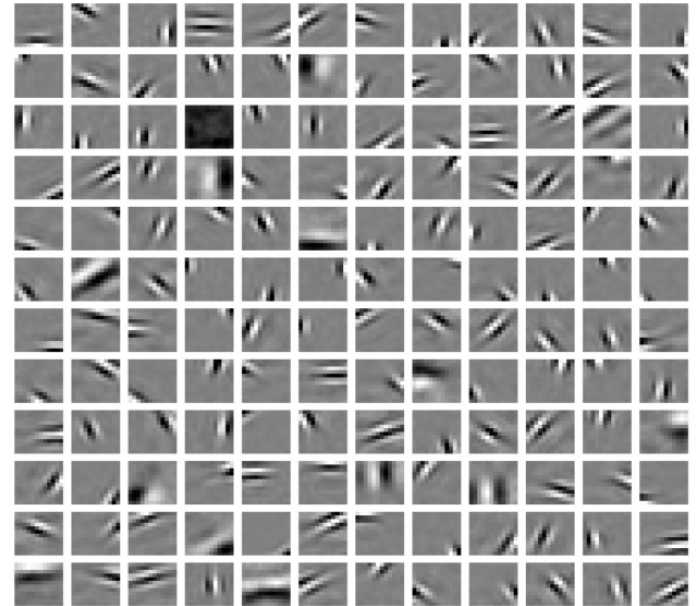
$$\text{Cov}(S) = \text{Diagonal Matrix}$$



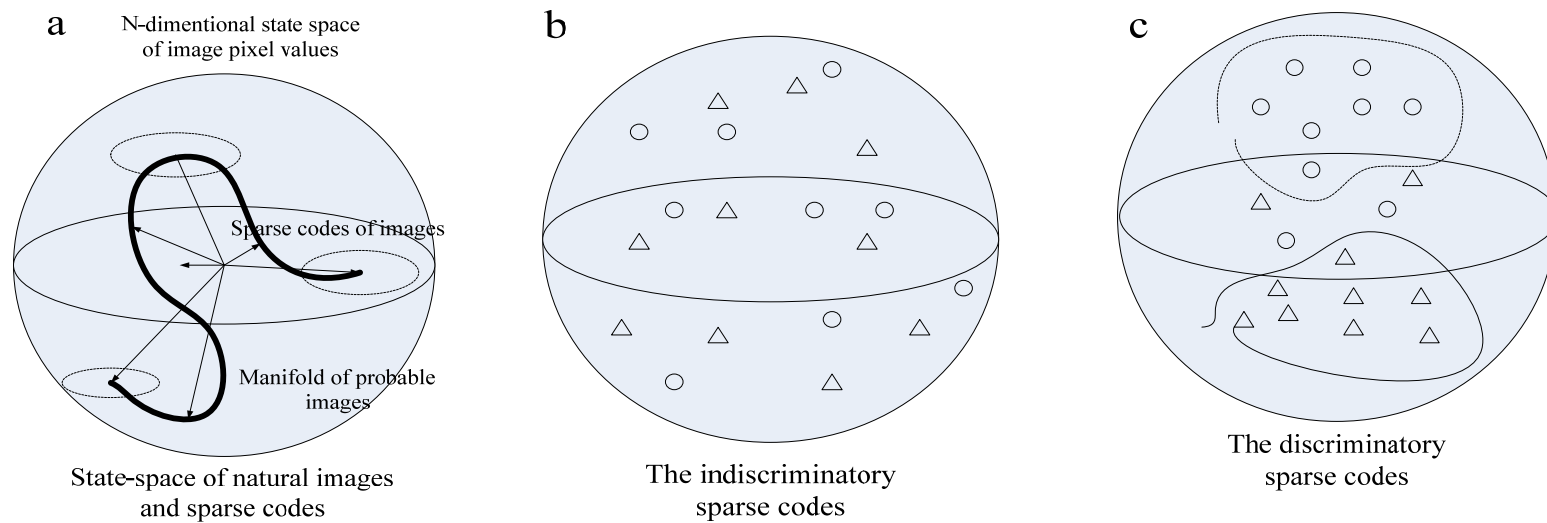
# 基于高阶统计的线性有效编码模型

$$P(S) = P(S_1) * P(S_2) * \dots * P(S_n)$$

$$E(a, \Phi) = \sum_{x,y} \left[ I(x,y) - \sum a_i \Phi_i(x,y) \right] + \lambda_s \sum_i S \left( \frac{a_i}{\sigma_i} \right)$$



# 区分空间编码



- 线性叠加模型

$$I(x, y) = As = \sum_{i=1}^n s_i * a_i(x, y)$$

- 稀疏编码模型

$$E(s, A) = \sum_{x,y} \left[ I(x, y) - \sum s_i a_i(x, y) \right]^2 + \lambda_s \sum_i S \left( \frac{s_i}{\sigma_i} \right)$$

# 面向分类任务的有效编码模型 — 定义

- 训练数据和编码系数

$$X_1 = \{I_1^1, I_2^1, \dots, I_{N_1}^1\} \quad X_2 = \{I_1^2, I_2^2, \dots, I_{N_2}^2\} \quad I_i^j = [s_1, s_2, \dots, s_n]$$

- 距离函数

$$D(I_1, I_2) = \sqrt{\sum_{i=1}^n (s_i^1 - s_i^2)^2}$$

- 类间距离

$$D_w = D(I_i^j, \tilde{m}_j)$$

- 类内距离

$$D_B = D(I_i^j, \tilde{m}_{\hat{j}})$$

# 面向分类任务的有效编码模型 — 定义

- 类别距离

$$DR = \frac{D_B}{D_W} \quad Dis(s) = \ln(DR^2) = \ln\left(\sum_{i=1}^n (s_i^j - \tilde{m}_i^j)^2\right) - \ln\left(\sum_{i=1}^n (s_i^j - \tilde{m}_i^{\hat{j}})^2\right)$$

- 目标函数

$$E(s, A) = Error(s, A) + \lambda_s Sparseness(s, A) + \lambda_d Dis(s)$$

- 优化目标

$$\text{Min}(E)$$

- 两层优化过程:

内层

$$\frac{dE}{ds_i} = -2b_i + \frac{\lambda_s}{\sigma} S'\left(\frac{s_i}{\sigma}\right) + \lambda_d \left( \frac{2 * (s_i^j - \tilde{m}_i^j)}{\sum_{i=1}^n (s_i^j - \tilde{m}_i^j)^2} - \frac{2 * (s_i^{\hat{j}} - \tilde{m}_i^{\hat{j}})}{\sum_{i=1}^n (s_i^{\hat{j}} - \tilde{m}_i^{\hat{j}})^2} \right)$$

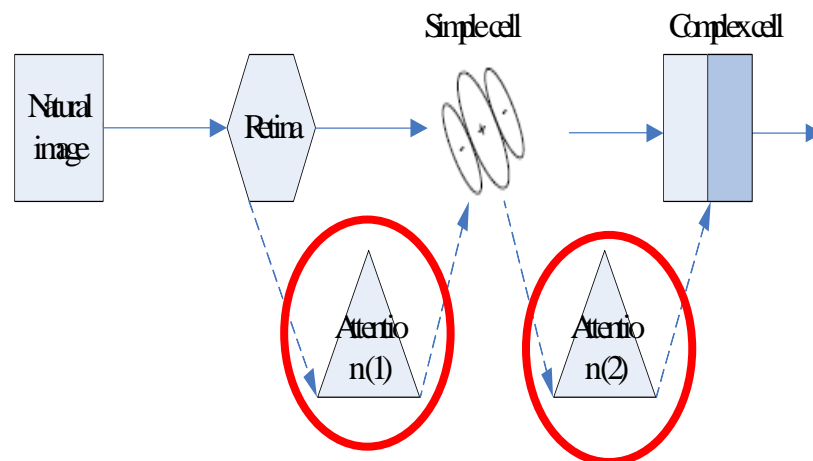
外层

$$\Delta a_i(x, y) = \eta \left\langle s_i \left[ I(x, y) - \sum_j s_j a_j(x, y) \right] \right\rangle$$

# 模型框架

“非均匀采样”注意选择模块

基于“响应显著度”注意选择模块

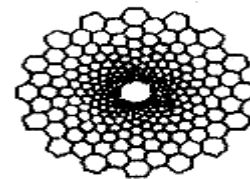


基于注意机制的稀疏编码  
模型框图

# 非均匀采样模块

## 视网膜生理特性

- 视网膜中央凹区感光细胞的密度远远大于周边区域感光细胞的密度；
- 视网膜影像的分辨率在视网膜中央凹区比较高，随着距离的增大分辨率逐渐减少。

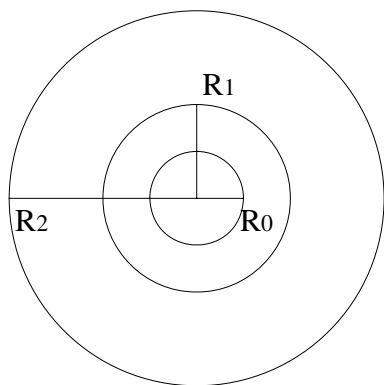


视网膜采样图示

# 非均匀采样模块

输入刺激图像非均匀地划分为不同采样率的圆形区域：

$R_0$  对应最高采样区域，它进行原采样保持与输入刺激的一致， $R_1$ ， $R_2$ 的采样率依次降低， $R_2$ 区域代表最低的采样频率。



$$X_{ij} = \begin{cases} R_{ij}^1 & \text{if } D(i, j) \leq R_0 \\ R_{ij}^2 & \text{if } R_0 \leq D(i, j) \leq R_1 \\ R_{ij}^3 & \text{if } R_1 \leq D(i, j) \leq R_2 \end{cases}$$

# 非均匀采样模块

## 递归高斯卷积模型模拟非均匀采样算法

$$\begin{aligned}
 R_{ij}^{(1)} &= X_{ij} \\
 R_{ij}^{(2)} &= \sum_{p=-2}^{p=2} \sum_{q=-2}^{q=2} G_{pq} R_{i-p, j-q}^{(1)} \\
 R_{ij}^{(l)} &= \sum_{p=-2}^{p=2} \sum_{q=-2}^{q=2} G_{pq} R_{i-p*2^{l-1}, j-q*2^{l-1}}^{(l-1)}
 \end{aligned}
 \quad [G_{pq}] = \begin{bmatrix} 1 & 4 & 6 & 4 & 1 \\ 4 & 16 & 24 & 16 & 4 \\ 6 & 24 & 36 & 24 & 6 \\ 4 & 16 & 24 & 16 & 4 \\ 1 & 4 & 6 & 4 & 1 \end{bmatrix} * \frac{1}{256}$$

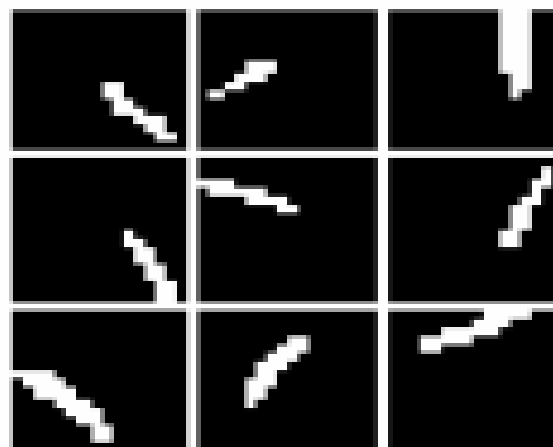
# 基于响应显著度的注意选择模块

响应显著度 (Response Saliency) 是度量一个神经细胞相对一组神经细胞而言对输入刺激产生响应的显著程度。

- 响应显著性是一个标量；
- 被选择作为下一级神经细胞输入的参考量  
响应显著性值大的神经细胞输出作为下一级神经细胞的输入被进一步处理，相反，显著性值小的神经细胞输出则被忽略。

# 模式距离

- ❖ **模式距离 (Pattern Distance)**表示两个简单细胞表征的响应模式之间的距离，记为 $D$ 。
- ❖ 它是简单细胞的三个响应特性：位置（L），方向（O）和空间频率（F）的函数。
  - 兴奋区域的中心表示位置（L）；
  - 兴奋区域主轴与x轴的夹角表示方向（O）；
  - 兴趣区域的面积大小表示空间频率（F）。



$$D(A_i, A_j) = W_1 * N \left( \sqrt{(L_{ix} - L_{jx})^2 + (L_{iy} - L_{jy})^2} \right) + W_2 * N \left( |O_i - O_j| \right) + W_3 * N \left( |F_i - F_j| \right)$$

# 区分距离

- ❖ 区分距离 (*Discrepancy Distance*) 表示一个简单细胞与作用于同一刺激的简单细胞群之间响应模式的区别，记为 *Diff*。
- ❖ 它不仅与模式距离相关，还与单个细胞的响应值相关。

$$Diff(A_i, S) = \left( \sum_{A_j \in S} N(D(i, j)) * \frac{|R_j|}{\sum_{A_k \in S} R_k} \right)$$

# 响应显著度值 (RS)

- 内因—简单细胞自身的响应值 ( $R_i$ )  
响应值越大则响应显著性值也越大
- 外因—作用于同一刺激的简单细胞集合 (S) 的影响  
按照格式塔知觉原理，具有相似特性的神经细胞之间的响应相互增强，反之，相互抑制。我们把这种影响用区分距离来度量。

$$RS(A_i) = N(|R_i|) + \lambda * (1 - Diff(A_i, S))$$

# 选择策略

## ■ 阈值选择机制 (TSM)

TSM 是一种阈值过滤算法。当响应显著性值大于阈值时，这个神经元的响应被选择进入下一级神经元的处理，否则被忽略。

$$Output(A_i) = \begin{cases} 0 & \text{if } RS(A_i) \leq T_i \\ R_i & \text{if } RS(A_i) > T_i \end{cases}$$

## ■ 比例选择机制 (PSM)

PSM 是一种瓶颈过滤算法。根据响应显著性值进行排序，选择前面的一定比例（假设为p）的神经元进入下一级处理。

# 阈值选择算法

算法 1: 阈值选择算法

**Input:** 重构误差上限 (ET), 基函数集 (A), 训练图像集 (I), 训练图像集对应的稀疏编码系数集合 (R) 和响应显著度集合 (RS)。

**Output:** 阈值 T

**Method:** 算法步骤如下。

- 1、选择一个 T 的初始估计值;
- 2、用 T 过滤编码系数。如果  $RS_i$  大于 T, 则  $R'_i = R_i$ , 否则  $R'_i = 0$ ;
- 3、对于训练集计算重构误差值:

$$\text{Error}(R'_i, A) = \sum_{I_i} \sum_{x,y} \left[ I_i(x,y) - \sum_i R'_i A_i(x,y) \right]^2$$

- 4、如果  $\text{Error} \geq ET$ , 则  $T = \eta T$  其中  $0 < \eta < 1$ , 重复步骤 2 到 4; 否则算法终止, 输出 T。

# 比例因子选择算法

算法 2: 比例因子选择算法

**Input:** 重构误差上限 (ET), 基函数集 (A), 训练图像集 (I), 训练图像集对应的稀疏编码系数集合 (R) 和响应显著度集合 (RS)。

**Output:** 比例因子 P

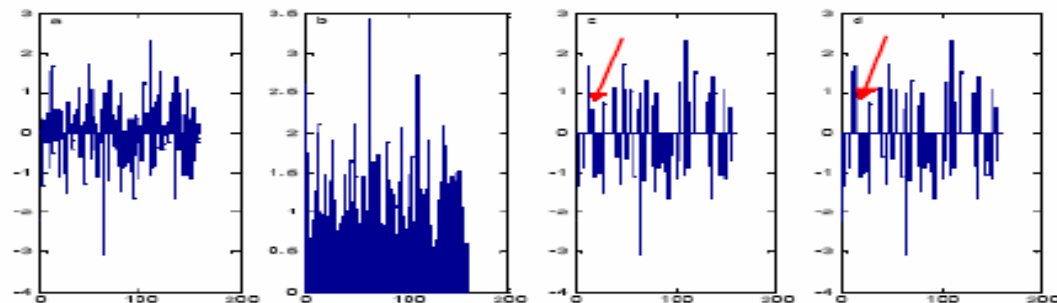
**Method:** 算法步骤如下。

- 1、选择一个 P 的初始估计值;
- 2、用 P 过滤编码系数。如果  $RS_i$  属于前  $P \times$  简单细胞总个数, 则  $R'_i = R_i$ ,  
否则  $R'_i = 0$ ;
- 3、对于训练集计算重构误差值:

$$\text{Error}(R'_i, A) = \sum_{I_i} \sum_{x,y} \left[ I_i(x,y) - \sum_i R'_i A_i(x,y) \right]^2$$

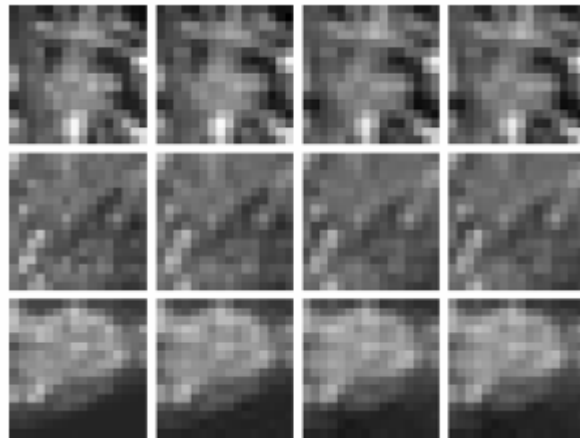
- 4、如果  $\text{Error} \geq ET$ , 则  $P = (1+\eta)P$  其中  $0 < \eta < 1$ , 重复步骤 2 到 4; 否则  
算法终止, 输出 P。

# 仿真结果 (1)



输入图像刺激的各种响应系数图。a) 稀疏编码模型产生的系数；b) 各个简单细胞响应显著度值；c) AGSC-P模型产生的系数 ( $p=45\%$ )；d) AGSC-T模型产生的系数 ( $T=1.372$ )。

# 仿真结果 (2)



输入图像和重构图像。第一列是原始的输入图像；第二列是通过完整稀疏编码系数重构的图像；第三列是通过AGSC-P模型系数重构的图像；第四列是通过AGSC-T模型系数重构的图像。

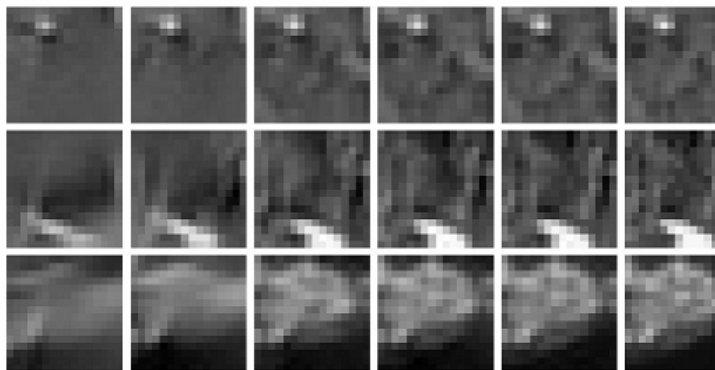
# 仿真结果 (3)

$$\text{Error}(s, A) = \sum_{x,y} \left[ I(x, y) - \sum_i s_i a_i(x, y) \right]^2$$

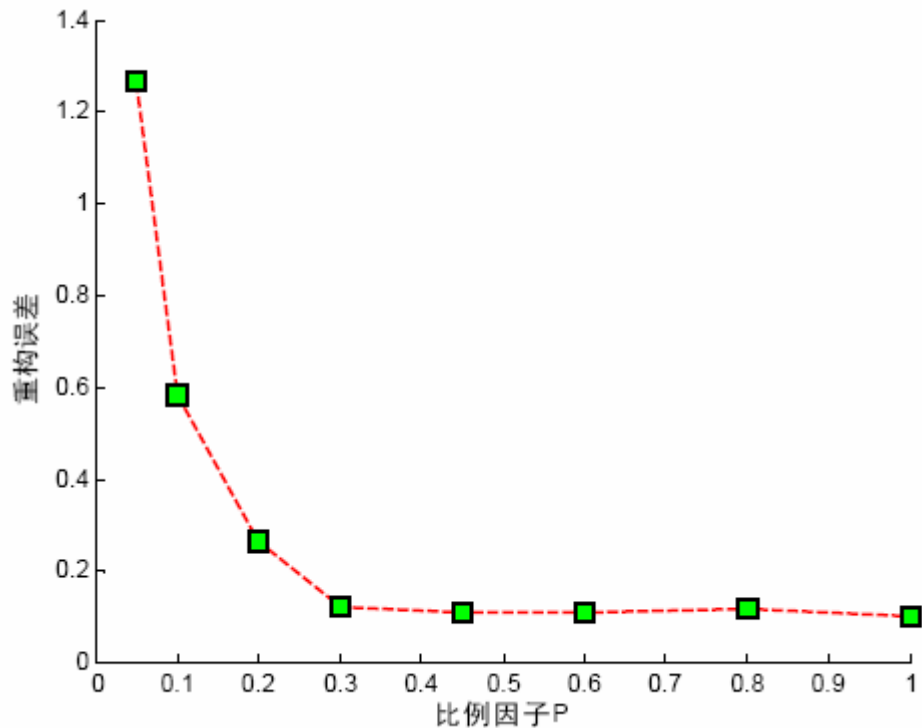
	Mean error	Maximum error	Minimum error
AGSC-T	0.1113	0.3929	0.0078
AGSC-P	0.1105	0.3913	0.0060
SC	0.1002	0.3209	0.0032

稀疏编码模型（简称SC）和基于注意机制的稀疏编码模型（AGSC）的重构误差比较图。

# 仿真结果



AGSC模型在比例因子 $P$ 取不同值的情况下的重构图像，从左到右 $P$ 的取值分别为 0.1, 0.20, 0.30, 0.45, 0.80, 最右边的图像为原始图像。



# 结束语

- 本体语义表示
- 事件来龙去脉
- 提出了一个格式塔知觉组织法则的量化模型和一个基于量化域值和组织规则的知觉组织模型。提高了知觉组织的适用性，具有较好的鲁棒性。
- 知觉有效编码

# Thank You

Question!

Intelligence Science  
<http://www.intsci.ac.cn/>

