

# 主动轮廓线模型 (Snake模型)



# 1.引言



- Marr视觉计算理论的不足

- 三个独立的层次, 底层缺乏约束导致病态问题
- 自下而上, 底层的错误将被带给高层无法修正

- Snakes: active contour models

- Kass, 1987, ICCV

- 对传统的视觉计算理论的挑战

- 设计这样一个能量函数: 其局部极值组成了可供高层视觉处理进行选择的方案, 高层机制可能通过将图像特征推向一个适当的局部极值点从该组方案中选择最优的一种



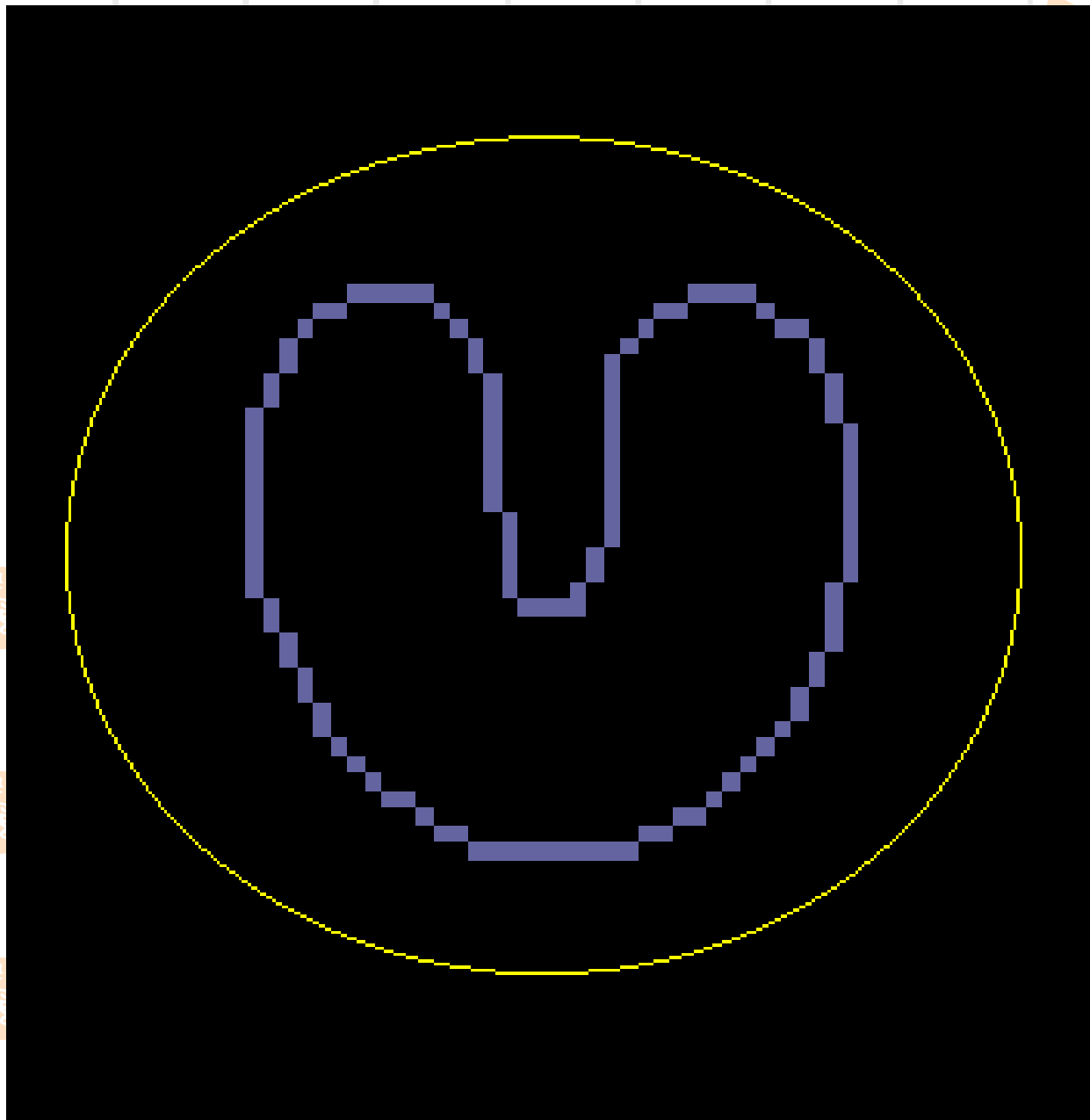
# 1.1 Snake模型的基本原理

- 基本原理是表征拟合误差的“能量”为最小化的曲线。
- 设对于拟合目标有一个待选曲线集, 定义能量函数与待选集中每一条曲线相关联, 能量函数的设计原则就是: 有利属性要能导致能量缩小。
- 有利属性包括: 曲线连续、平滑、曲线与高梯度区域接近以及其他一些具体的先验知识。
- 活动轮廓在取值范围内移动时, 就能在能量函数的指导下收敛到局部边界, 且能保持曲线的连续和平滑。

# 1.1 Snake模型的基本原理

- 蛇模型是在曲线本身的内力和图像数据的外部约束力作用下的移动的变形轮廓线。
- 作用在蛇模型上的力依据轮廓所在的位置及其形状决定如何在空间局部的变化。
- 内力和外力的作用是不同的:内力起平滑约束作用,外力则引导蛇模型向图像特征移动。
- 施加在蛇模型上的外力来自于图像或更高层的处理外力,将蛇模型推离不期望的特性。
- 蛇模型的内力包含两项. 形象的说,可以认为蛇模型是由两种抽象的弹性材料构成:弦与杆。前者使轮廓抵抗韧性,而后者使轮廓抵抗弯曲。

吉祥慶



吉祥慶

吉祥慶

吉祥慶

吉祥慶

吉祥慶

吉祥慶

## 1.2 Snake模型的特点



### ■ Snake模型的优点

- 图像数据、初始估计、目标轮廓及基于知识的约束统一于一个特征提取过程中；
- 经适当地初始化后，它能够自主地收敛于能量极小值状态；
- 尺度空间中由粗到精地极小化能量可以极大地扩展捕获区域和降低计算复杂性

### ■ Snake模型的缺点

- 对初始位置敏感，需要依赖其他机制将Snake 放置在感兴趣的图像特征附近；
- 它有可能收敛到局部极值点，甚至发散。



## 2. Snake模型的数学模型

定义Snake模型为一可变形曲线

$$v(s) = \{ x(s), y(s) \}$$

**S**为归一化的曲线长度,变化范围(0,1)

蛇模型的总能量函数是

$$E_{snake} = \int [ E_{int}(v(s)) + E_{ext}(v(s)) ] ds$$

## 2. Snake模型的数学模型

$E_{int}$  是内部能量, 控制蛇模型特性, 定义为

$$E_{int} = 1/2 (\alpha | v'(s) |^2 + \beta | v''(s) |^2)$$

$v'(s)$   $v''(s)$  分别是  $v$  对  $s$  的一阶和二阶导数, 系数  $\alpha$ 、 $\beta$  分别是控制蛇模型的弹性和刚性, 这些参数操纵着模型的物理行为和局部连续性

外部能量  $E_{ext}$  决定着向某种固定的特征移动蛇模型, 吸引蛇模型到显著的图像特征。因为这些特征只能根据特定的问题而定义, 所以一般的外部能量函数不易确定。因此,  $E_{ext}$  没有统一的数学表达式, 必须从问题本身的特性出发, 根据实际情况处理



## 外部能量 ( $E_{ext}$ )

- 图像能量
- 定义函数  $E_{image}(x,y)$ ，反映的是对图像特征（如边界）的兴趣程度

$$E_{ext} = \int_s E_{image}(v(s)) ds$$

$E_{image}(x, y)$ 函数的定义是一个关键问题.

典型的例子为

$$E_{image}(x, y) = -|\nabla I(x, y)|^2$$

$$E_{image}(x, y) = -|\nabla(G_\sigma(x, y) * I(x, y))|^2$$

# 能量与内外力平衡方程

- 目标轮廓的确定就被转化成了极小化如下的能量泛函的问题

$$E_{snake} = \int_s \frac{1}{2} (\alpha(s) |v_s|^2 + \beta(s) |v_{ss}|^2) + E_{image}(v(s)) ds$$

- 由变分法的原理出发, 可以将其转化为Euler 方程

$$\alpha v_{ss} - \beta v_{ssss} - \nabla E_{image} = 0$$

- 这一方程可以被看作是轮廓内外力的平衡公式.

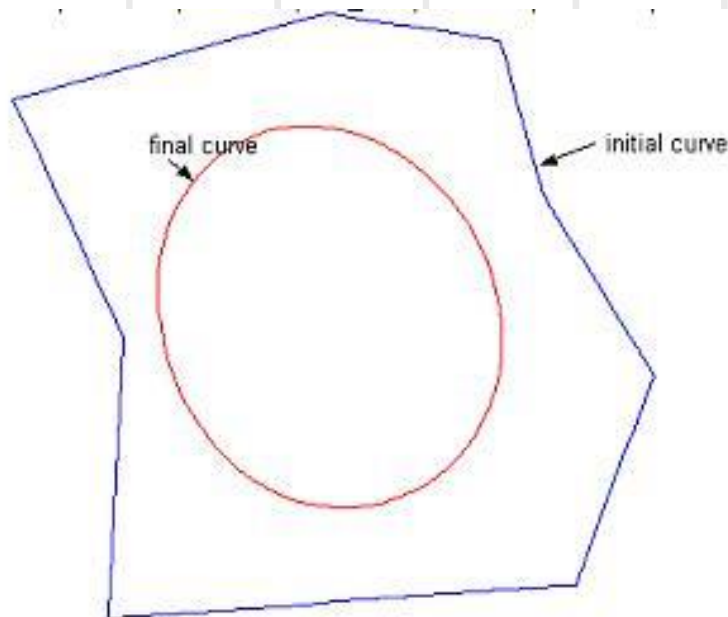
- 每个力都有对应的意义, 在这些力的作用下轮廓发生形变。

# 弹性力

- 由轮廓的弹性能量产生

$$F_{elastic} = \alpha v_{ss}$$

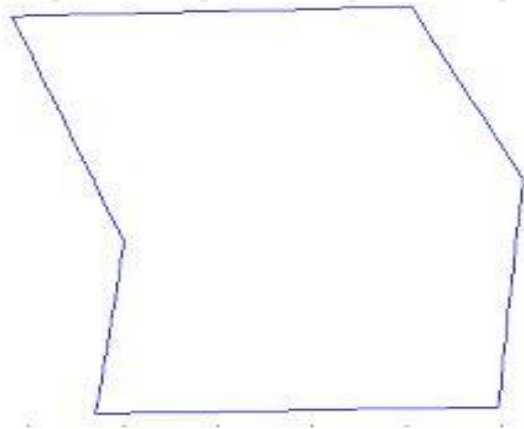
- 特性



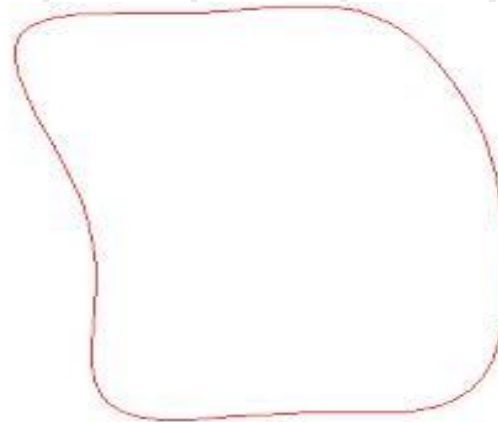
- 这个力使得轮廓连续。

# 刚性力

- 对应着轮廓的刚性能量，也就是曲率
- 特性



Initial curve  
(High bending energy)



Final curve deformed by  
bending force. (low  
bending energy)

这个力使得轮廓尽量平滑。

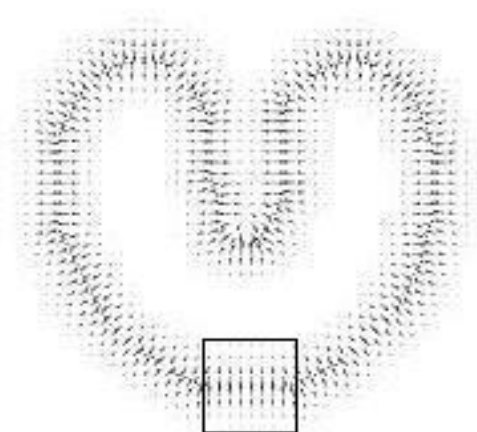
# 外部力

$$F_{ext} = -\nabla E_{image}$$

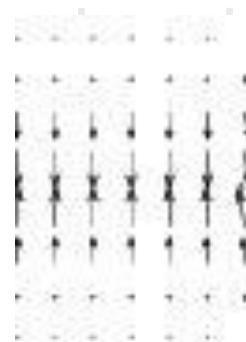
- 外部力作用在使得外部能量减小的方向上



Image



External force



Zoomed in

# 离散化

- 轮廓  $v(s)$  由一系列控制点组成  $v_0, v_1, \dots, v_{n-1}$
- 轮廓通过依次连接各个控制点并分段线性化得到。
- 平衡力方程独立作用于各个控制点
- 每个控制点在内外力的作用下是可以移动的。
- 能量以及平衡力的方程均作离散化处理。

### 3.Snake模型用于轮廓提取的实例

在实际应用中,我们需要对**Snake**模型离散化,计算的是曲线的各个控制点的能量值,定义的能量函数如

$$\sum_{i=1}^n E_i = \sum_{i=1}^n (\alpha_i E_{continuity, i} + \beta_i E_{curvature, i} + \gamma_i E_{image, i})$$

#### (1) 内部能量的连续性项能量

$$E_{continuity, i} = \frac{|d_{mean} - \|v_i - v_{i-1}\| - Smallest_{i\_continuity}}{Largest_{i\_continuity} - Smallest_{i\_continuity}}$$

$|d_{mean} - \|v_i - v_{i-1}\|$  是待考察点的 $3 \times 3$ 邻域

$d_{mean}$  表示曲线上相邻点的平均距离,相邻点间的间距与平均值越接近,其能量值越小,这样即保证了平滑,又避免了堆积。

## (2) 内部能量的曲率项能量

$$E_{curvature, i} = \frac{(1 - \frac{u_i^\omega \cdot u_{i+1}^\omega}{|u_i^\omega| \cdot |u_{i+1}^\omega|}) - Smallest_{i\_curvature}}{Largest_{i\_curvature} - Smallest_{i\_curvature}}$$

$$\frac{u_i^\omega \cdot u_{i+1}^\omega}{|u_i^\omega| \cdot |u_{i+1}^\omega|} = \frac{(x_i - x_{i-1})(x_{i+1} - x_i) + (y_i - y_{i-1})(y_{i+1} - y_i)}{\sqrt{[(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2][(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2]}}$$

$\frac{u_i^\omega \cdot u_{i+1}^\omega}{|u_i^\omega| \cdot |u_{i+1}^\omega|}$  是向量  $u_i^\omega$  和  $u_{i+1}^\omega$  之间的夹角  $\Delta\theta$  的余弦值

夹角越小,  $1 - \frac{u_i^\omega \cdot u_{i+1}^\omega}{|u_i^\omega| \cdot |u_{i+1}^\omega|}$  越小, 用来估计曲线上各点的曲率





### (3) 图像能量

这一项表示图像的约束条件,根据有利边界点的原则,边界点应具有较小的值

$$E_{image,i} = \frac{Smallest_{i\_image} - \phi_{edge}(v_i)}{Largest_{i\_image} - Smallest_{i\_image}}$$

$\phi_{edge}(v_i)$  是边缘检测算子,  $Largest_{i\_image}$  是待考察点的  $3 \times 3$  邻域内  $\phi_{edge}(v_i)$  的最大值,  $Smallest_{i\_image}$  是最小值. 这样的计算用于归一化

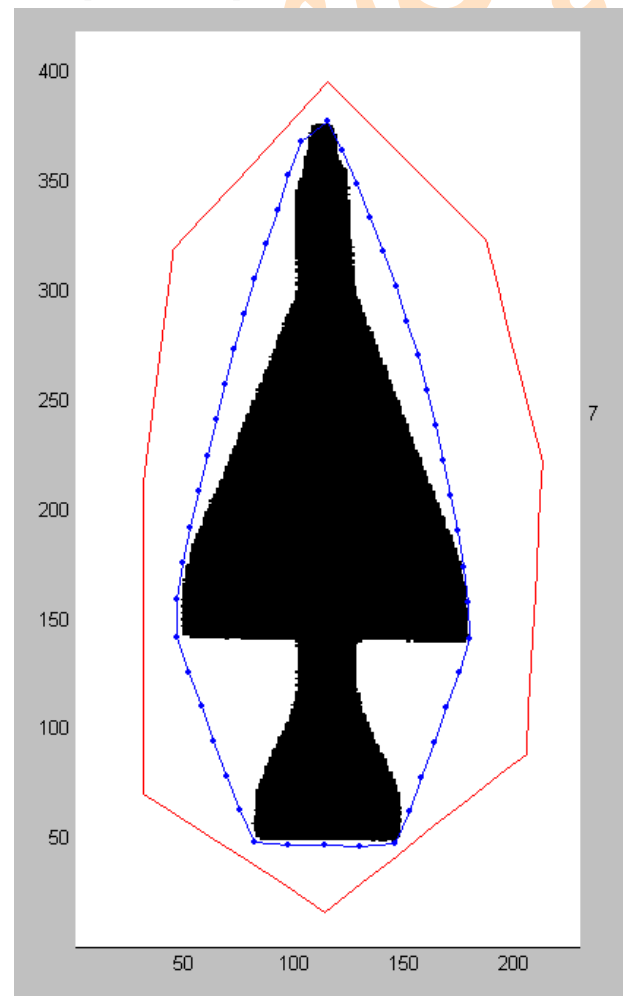
(4) 在确定能量函数后, 对曲线按照能量最小进行迭代.

### 3. Snake模型用于轮廓提取的实例



## 4. 传统Snake方法的不足

- 参数敏感,对初始轮廓要求高
- 搜索范围小
- 容易陷入局部极小点
- 对于边界上的凹点无法有效跟踪



## 4.Snake模型的改进



- 改善Snake 对初始化轮廓的敏感性;
- 保证Snake 能够收敛到全局极值;



■ 改善Snake 在能量极小化过程中的收敛速度或数值稳定性.



# 气球力—balloon force

- Cohen L D, On active contour models and balloons.1991, Image Understanding
- 在轮廓线上施加另一外部约束力,使轮廓线向目标靠拢。在该力的作用下轮廓线不断的向外膨胀,最终进化到目标轮廓,可以形象的称之为气球力
- 由气球力所构造的能量项,在能量函数中的数学形式可以表达为:

$$E_{bat} = \min_{\substack{-n \leq j \leq n \\ -m \leq k \leq m}} e_{jk}(v_i) \quad e_{ik}(v_i) = n_i \cdot (v_i - p_{jk}(v_i))$$

$p_{jk}(v_i)$  为以控制点  $v_i$  为中心的大小为  $n \times m$  的领域内的第(j,k)个邻点,  $\cdot$  代表矢量间的点乘运算.  $n_i$  是轮廓线上控制点  $v_i$  处的单位法线矢量,这样在规定的领域内,在法线矢量  $n_i$  方向上离控制点  $v_i$  最远的点将拥有最小的能量值

# 气球力—balloon force

在引入气球力能量项之后, **Snake**模型的外部能量项可以描述为

$$E_{ext} = E_{image} + E_{con} = k E_{bat} - l || \nabla I(v) ||$$

其中参数**k** 用来控制气球力的方向,当**k** 为负数时,气球力使轮廓线向内收缩,相反当**k**为正数时,气球力使轮廓向外膨胀;  
在选择参数**k** 和**l** 的大小时,一般将它们置于同一数量级,且**l** 稍大于**k**,这是为了在边缘点时轮廓线能够停止运动。  
这样,原始模型的缺点得到改善,对轮廓线的初始化位置要求明显降低,即使在初始位置离希望提取的边缘相当远时,**Snake**照样能够进化到目标轮廓。  
该模型改善了蛇模型对初始轮廓的敏感性,并且能够跨越图像中的伪边缘点。

# 梯度矢量流-Gradient Vector Flow(GVF)

- Xu C. and JL Prince. 1998. Snakes, shapes, and gradient vector flow. IEEE Trans Image Processing. 7(3): 359-363
- 它的数学基础来源于电磁场理论中的亥姆霍兹理论, 这种理论阐明了可以将一种普通的静态矢量场分解为两个组成部分, 即无旋场部分和有旋场部分。
- 在传统的主动轮廓模型中, 图像梯度信息仅仅是作为一个静态的无旋场来平衡方程。但是实际上我们能得到一个更加一般化的静态矢量场, 它不仅包含无旋场部分, 还包含有旋场部分。
- GVF 的提取可以有效的解决曲率变化很大的控制点的收敛效果, 但是相对的计算会很慢。

# 梯度矢量流-Gradient Vector Flow(GVF)

- GVF 定义一个力的向量场

$$\mathbf{V}(\mathbf{x}, \mathbf{y}) = (u(x, y), v(x, y))$$

- GVF snake的内外力平衡方程为

$$\alpha v_{ss} - \beta v_{ssss} + V = 0$$

- GVF snake定义的能量泛函为

$$E = \iint \mu(u_x^2 + u_y^2 + v_x^2 + v_y^2) + |\nabla f|^2 |V - \nabla f|^2 dx dy$$



- GVF场可以通过求解下述方程得到

$$\mu \nabla^2 u - (u - f_x)(f_x^2 + f_y^2) = 0$$

$$\mu \nabla^2 v - (v - f_y)(f_x^2 + f_y^2) = 0$$

$\nabla^2$  是拉普拉斯算子.

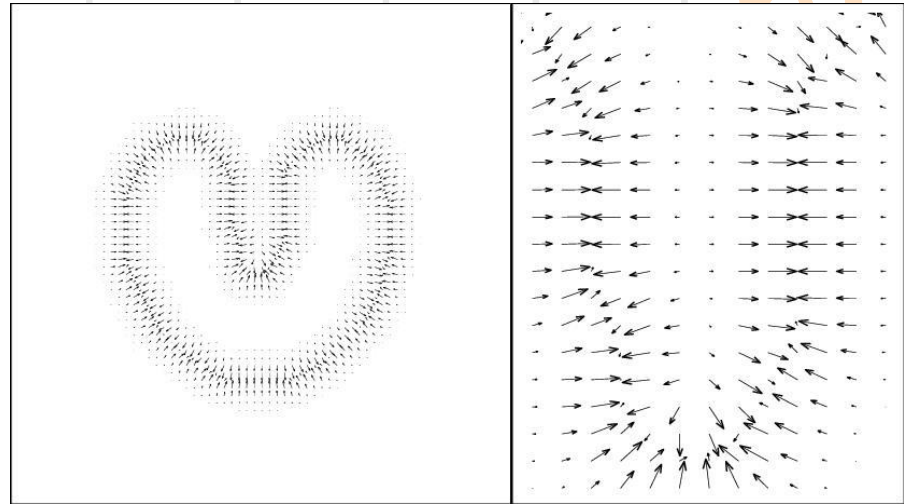
- 上述方程的求解是通过顺序迭代u和v实现的

- 能够检测边界上凹点的原因.

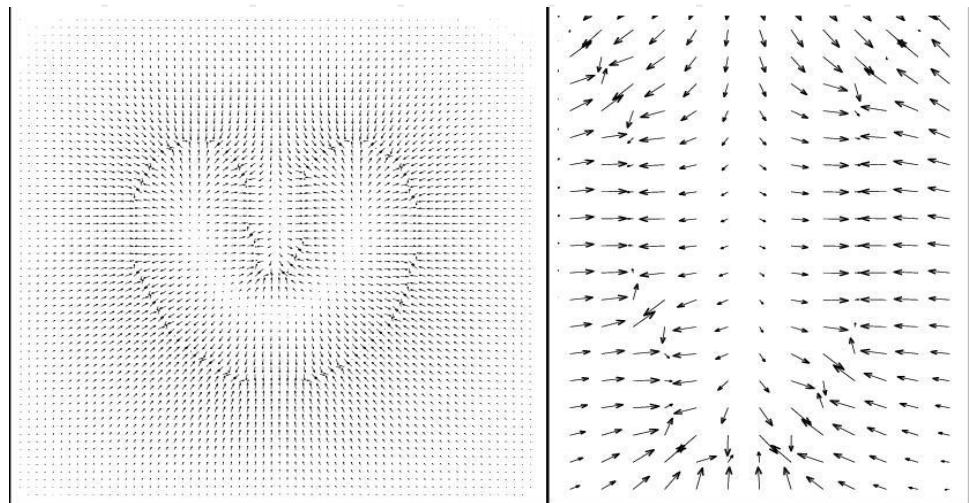
使得 $\nabla f$  由 $|\nabla f|$ 大的地方向 $|\nabla f|$ 小的地方扩散,  
因而扩大了Snake 模型的捕捉范围,也能较好地进入深度凹陷区域

# GVF 方法与传统Snake方法的比较

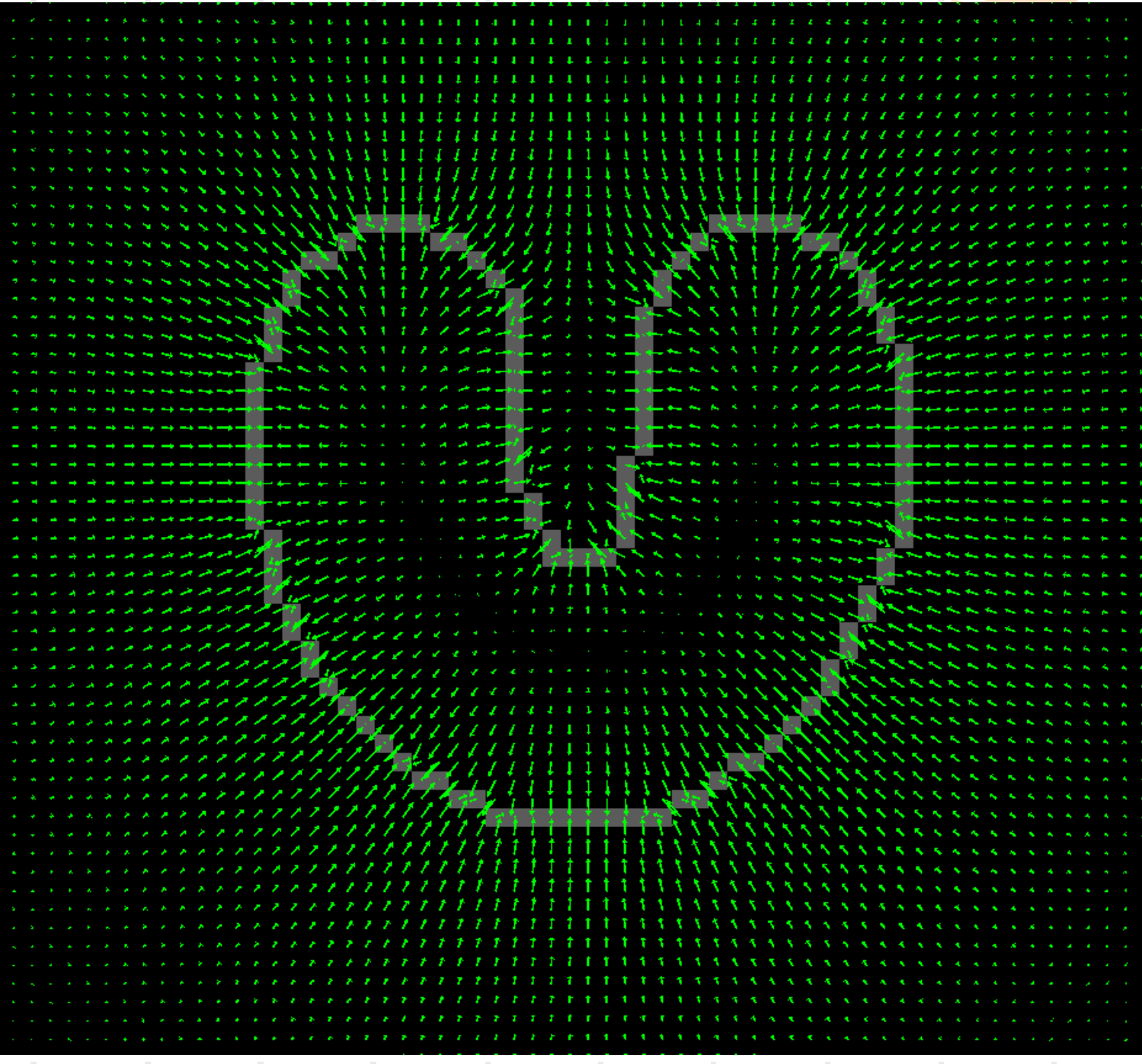
Traditional force



GVF force



福

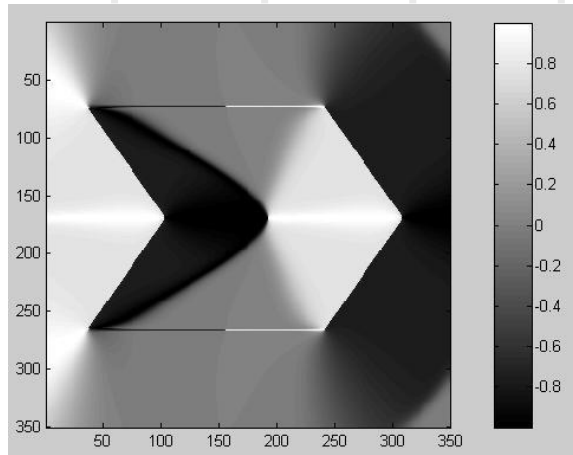
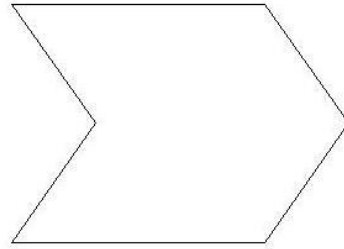


吉祥

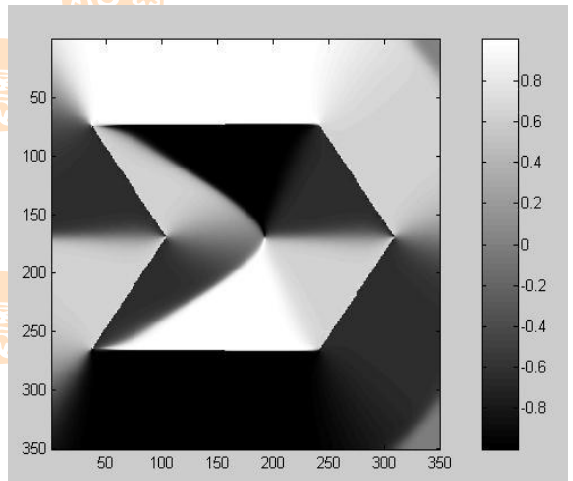
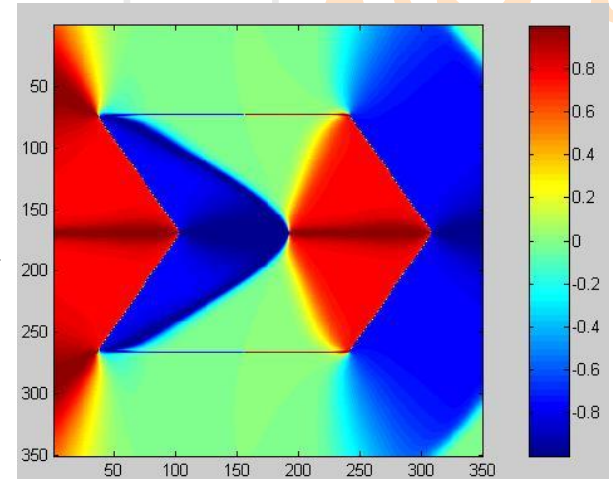
吉祥

吉祥

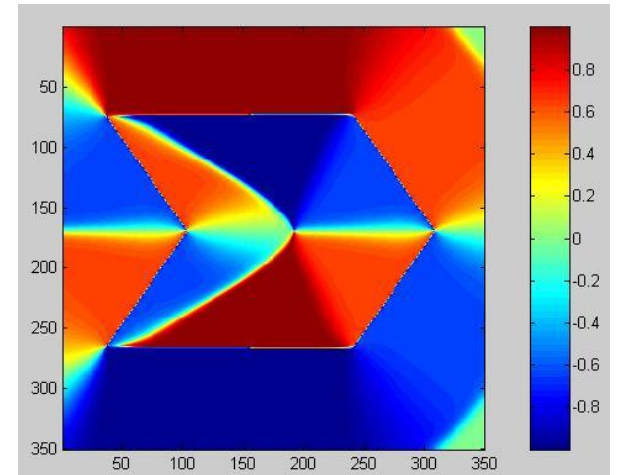
# 内外力的示意图



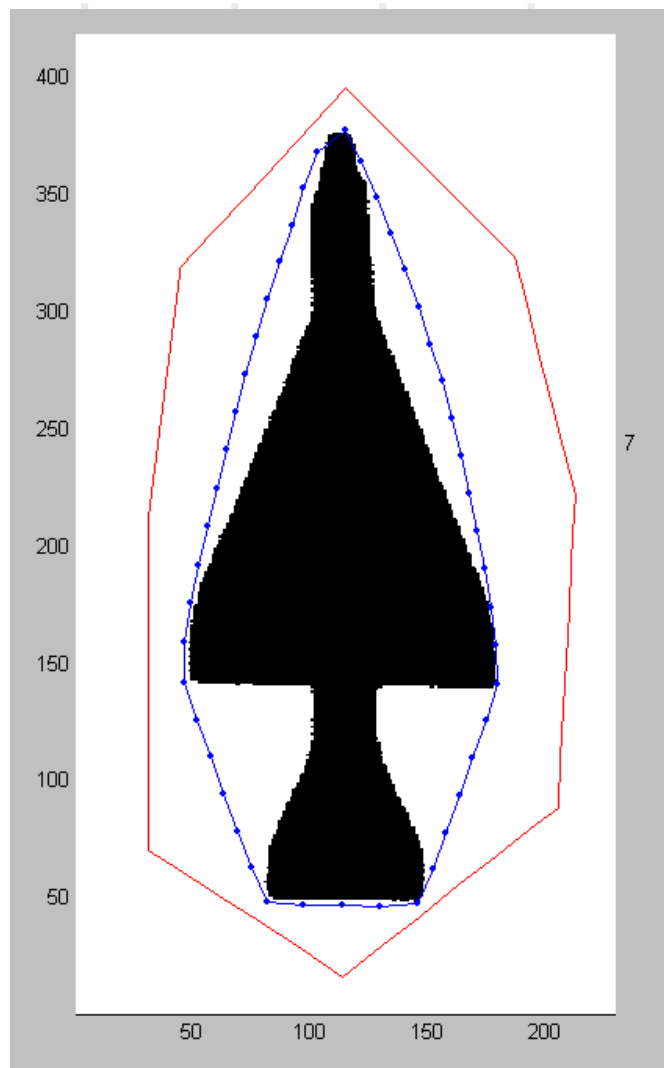
$$\longleftrightarrow u(x,y) \longrightarrow$$



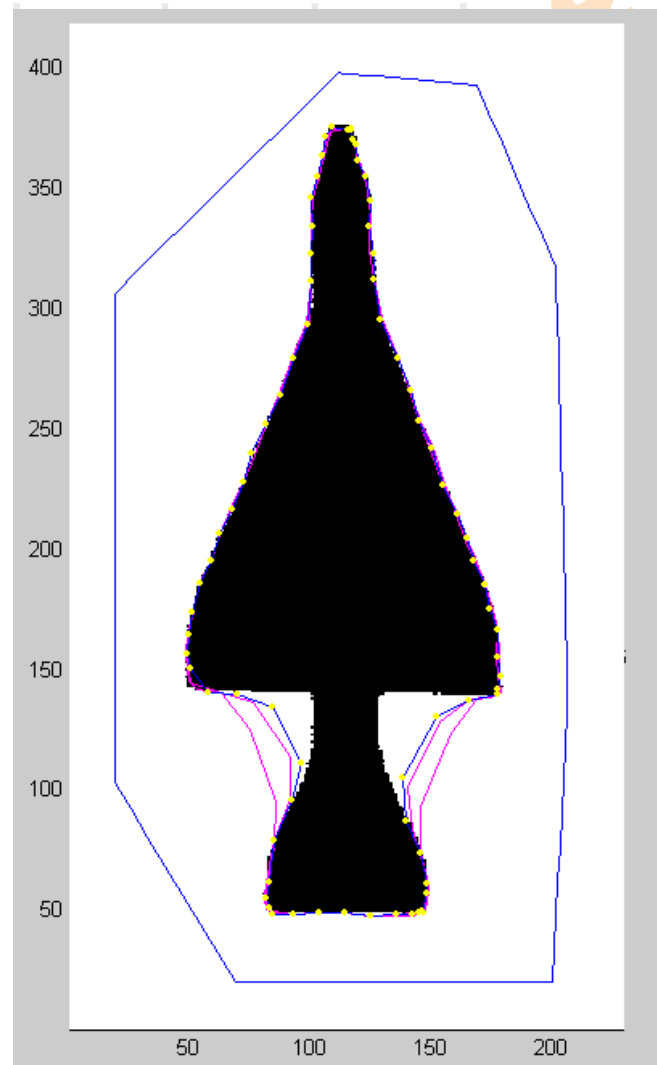
$$\longleftrightarrow v(x,y) \longrightarrow$$



# 实验结果

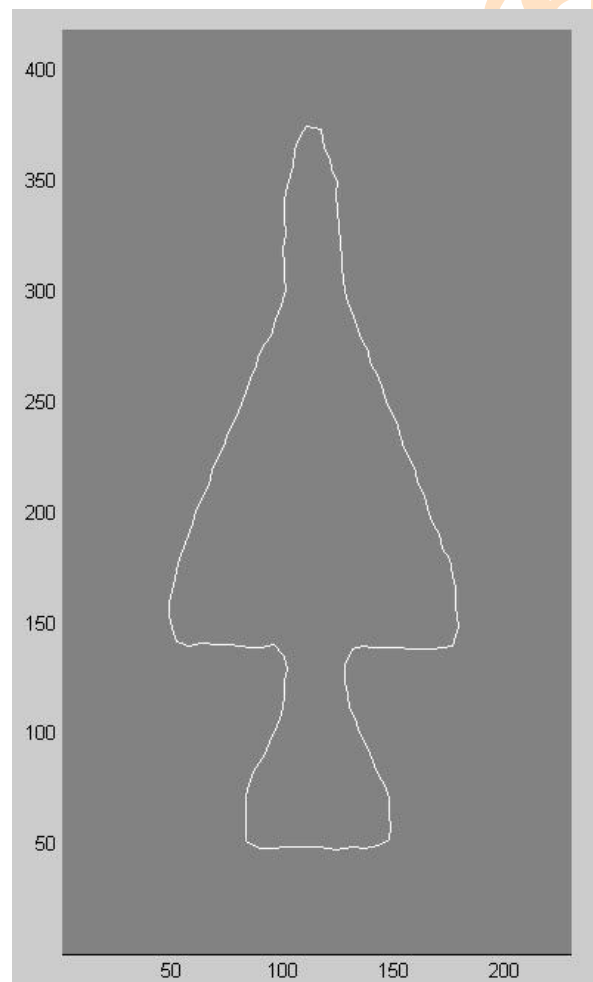
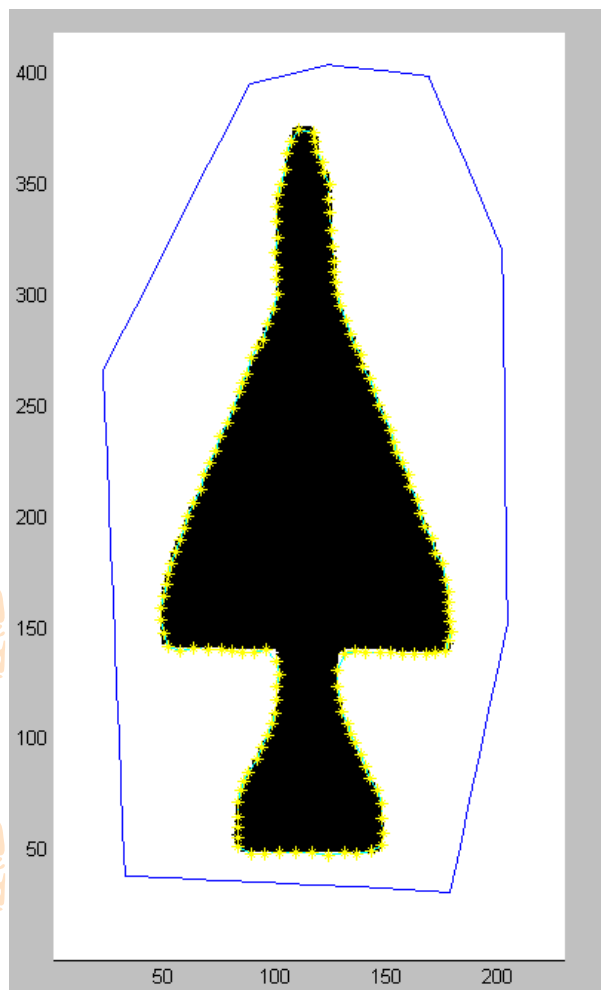


Traditional snake



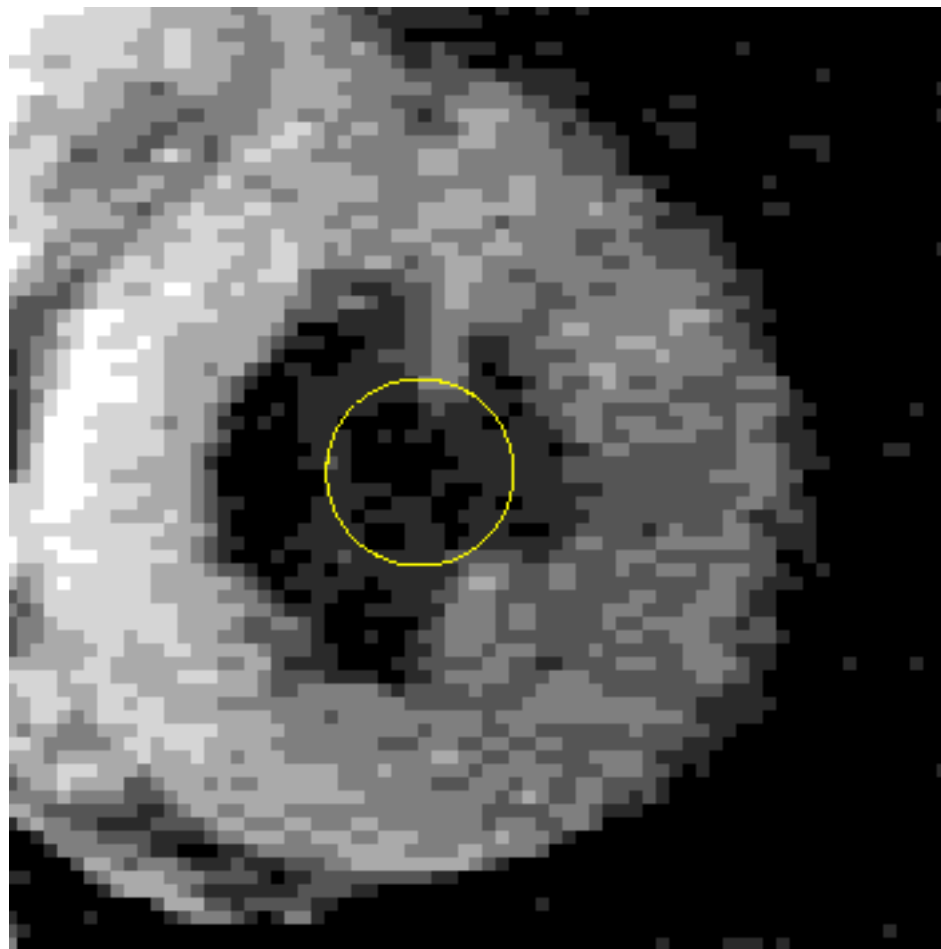
GVF snake

经过动态参数修正以后



Final shape detected

## Medical Imaging



左心房的磁共振图片

即使在原始图像质量很差时,也能很好的提取轮廓

吉祥如意

吉祥如意

吉祥如意

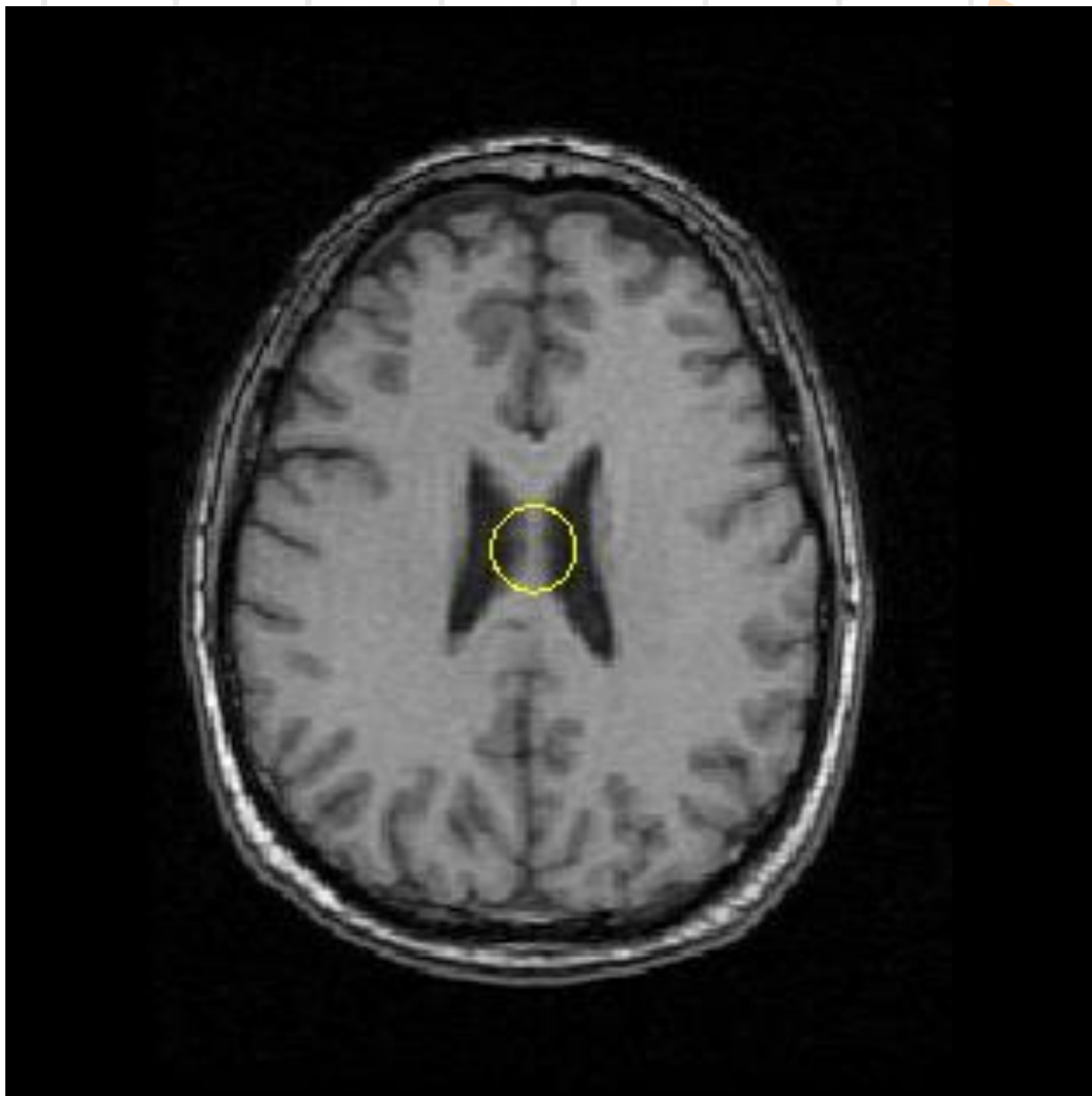
吉祥如意

吉祥如意

吉祥如意

吉祥如意

王齊慶



王齊慶

王齊慶

王齊慶

王齊慶

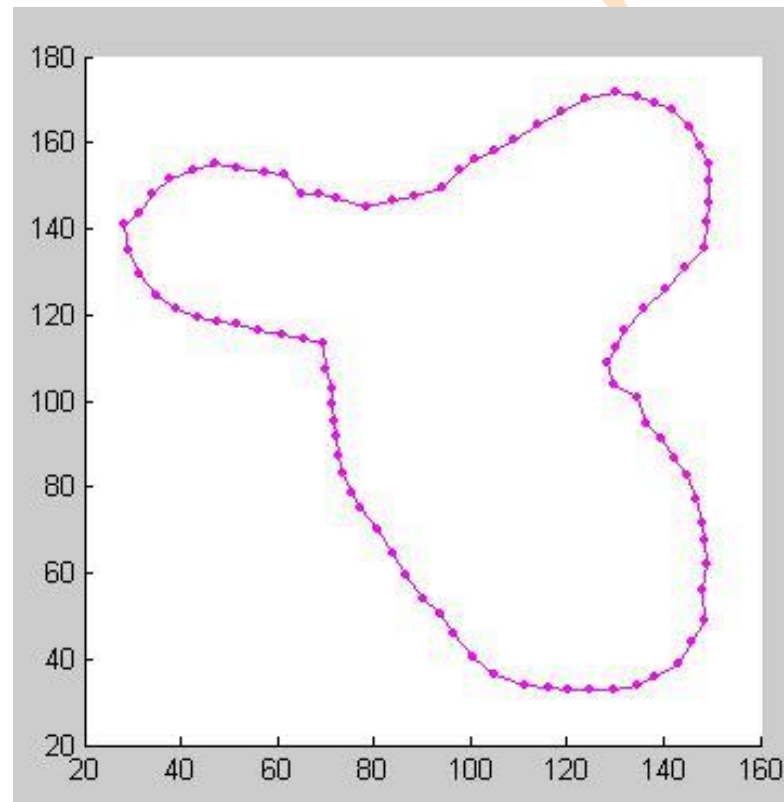
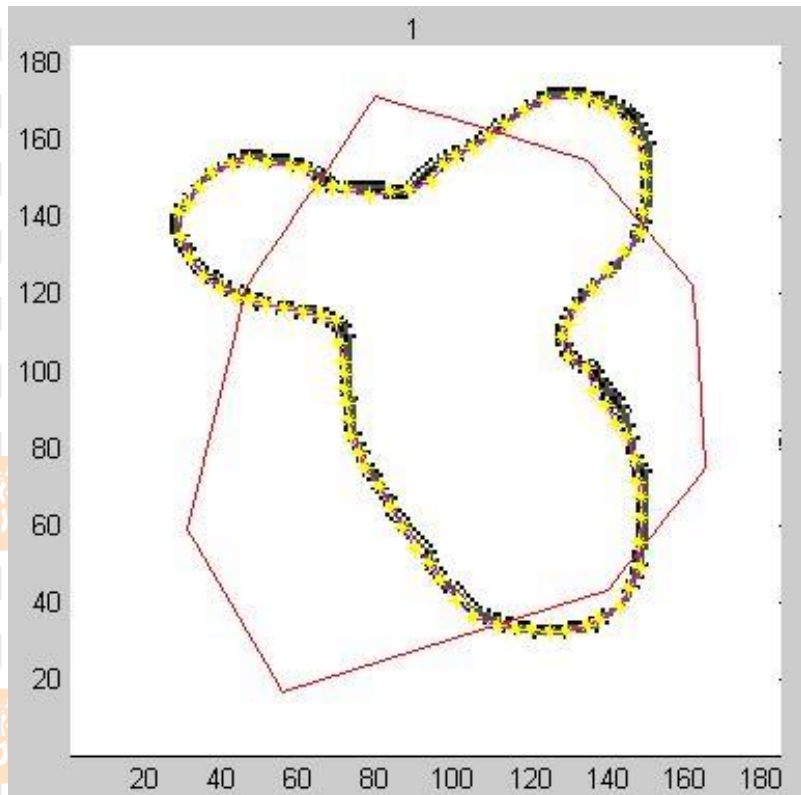
王齊慶

王齊慶



初始轮廓甚至可以与真实轮廓相交  
传统的Snake方法是无法实现的

吉祥如意



初始轮廓的选取是任意的吗？

吉祥

如意

吉祥如意

吉祥如意

# GVF snake的问题

- 参数敏感
- 计算代价高，速度慢
- 初始轮廓的选取有临界点

# 其它的改进方法



## ■ B - Snake 模型

- 目标轮廓用B - 样条来表达
- 利用参数B 样条的局部控制能力及参数连续性等特点, 改进了原模型中存在的一些缺陷, 例如不稳定地数据特性、收敛速度慢、很难处理不连贯性及对噪声的敏感性。



# 其它的改进方法



## ■ 双主动轮廓模型

- 所谓双主动轮廓模型,就是指在所需要提取的目标轮廓的内部和外部各放置一个主动轮廓线,分别为内轮廓线和外轮廓线。
- 在初始情况下,两个轮廓线独立进行,分别向目标轮廓收敛,互不影响;当两个轮廓线都静止不动时,则分别对内外轮廓进行能量计算并比较,选择能量较大的一个轮廓线,对其施加一个外部作用力,该力的方向指向另外一个轮廓线,强制使当前轮廓线从当前平衡位置离开,然后重新启动进化过程,当达到平衡时,再次进行能量比较并重复以上步骤,直至最终内外轮廓线的能量之差减小到允许的范围。
- 可有效解决由于被噪声所引起的奇异点吸引而陷入能量局部最小的问题。



# Snake模型的应用

## ■ 轮廓跟踪

### ➤ 目标识别

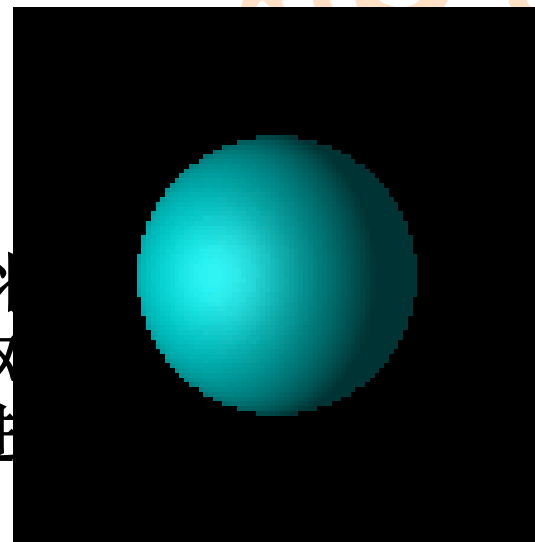
使目标提取从一开始就处于受控的状态，从而可以快速的对目标进行识别。同时还给出了对目标的约束，从而可以快速准确的对目标进行识别。

### ➤ 三维重建

三维表面用序列图像中的轮廓曲线来表达和重建。

### ➤ 立体匹配

如果两幅图像中的轮廓线(基元)互相对应，那么沿着3D 轮廓的其他几何形状应具有相似性。



# Snake模型的应用

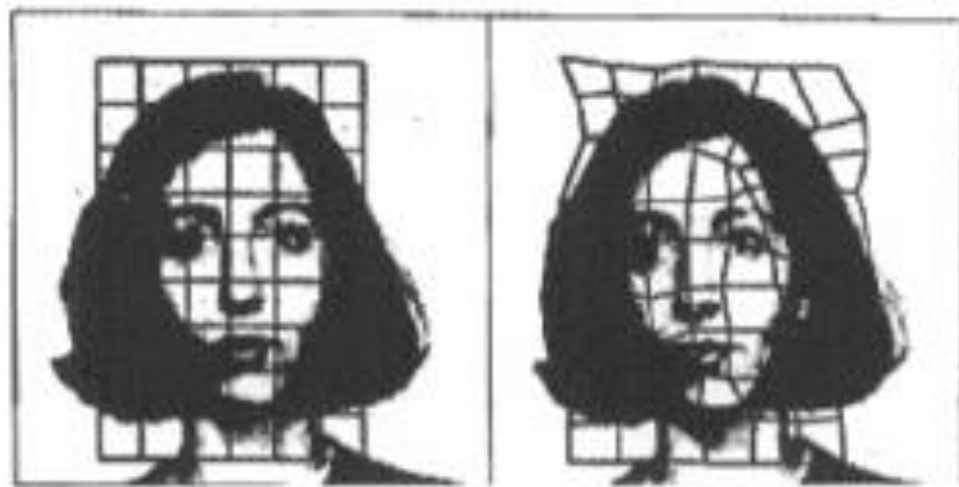
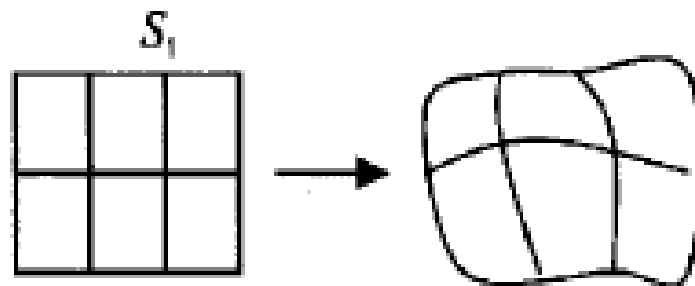


## ■ 变形模板

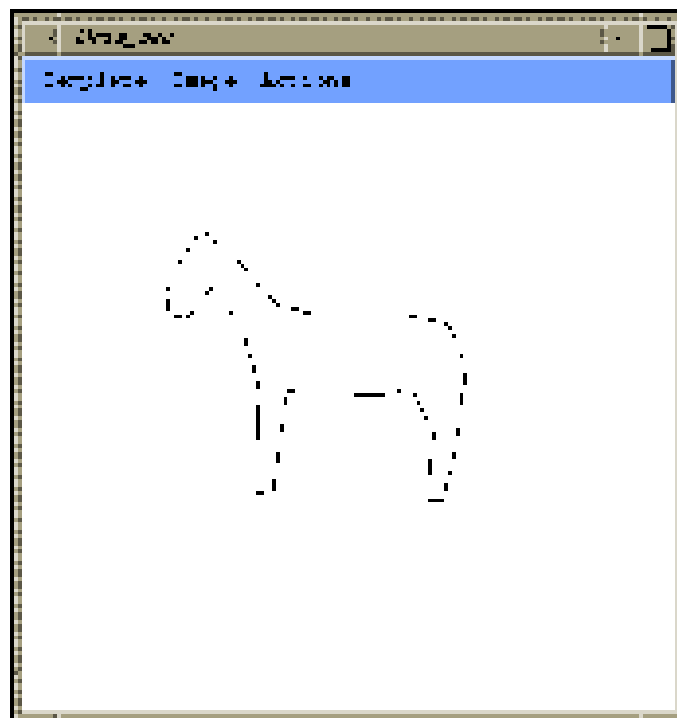
- 由于自然的物体形状的多样性和复杂性以及图像的噪声等复杂因素的影响，在使用刚性模型提取图像轮廓时遇到了极大的困难. 基于Snake 模型的变形模板方法，为解决这一问题展现了令人鼓舞的前景.
- 相对于刚性模板来说，变形模板使自己发生变形以匹配到显著的图像特征.
- 从数学角度来看，这可以解释为有两项组成的目标函数，其中一项测量变形模板与理想轮廓的偏差，另一项衡量变形模板与相应的图像特征的匹配程度.
- 于是，轮廓提取变成了使目标函数最大化或者最小化的最优化问题. 变形模板不但能够检测出目标特征，而且能够给出目标的描述，从而将其用于分类和匹配中.



# Snake模型的应用---变形模板

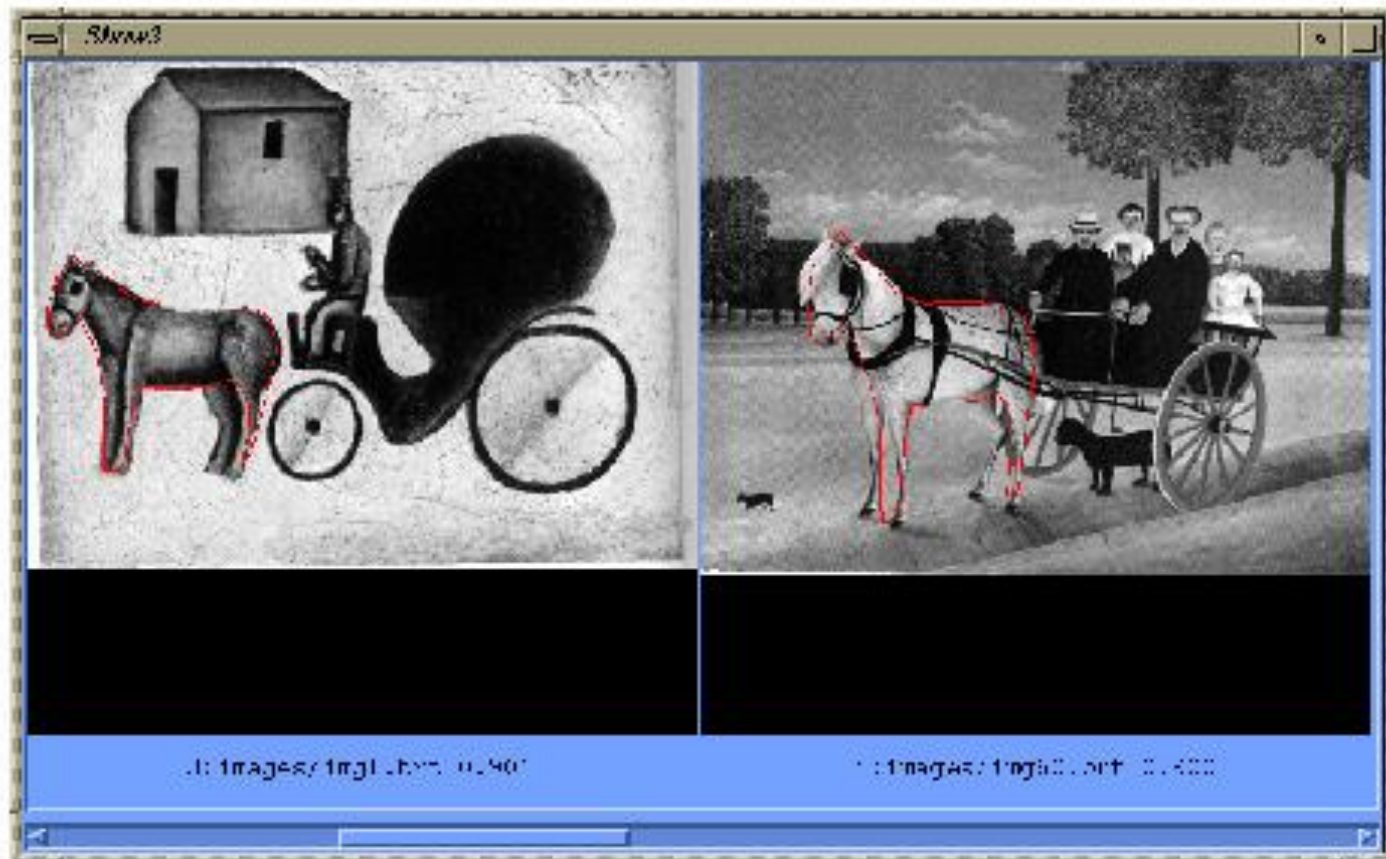


# 基于变形模板的弹性匹配



a) The user's query shape





b) Two of the retrieved images.



c) Another retrieved image in which two horses were found.

# Snake模型的应用

## ■ 运动跟踪

- 在时变图像中使用蛇模型进行跟踪地思想最初由 Kass 等人提出来, 他们使用蛇模型跟踪说话地唇动。当从动力学的角度研究蛇模型时, 研究工作取得了很大的进展。
- Szeliski 等人将蛇模型的动态模型作为卡尔曼 (Kalman) 滤波器的系统模型, 同时考虑了系统噪声和观测噪声模型, 提出了卡尔曼蛇 (kalmanSnake) 跟踪模型。由于该模型受到弹性动力学模型的制约, 因此可以有效地跟踪刚体和非刚体地复杂运动。

# 基于学习的计算机视觉方法

- 经典的模式识别技术：手工设计的特征提取+分类器



- 现代主流模式识别技术：无监督中层特征学习



- 深度学习技术：端到端的层级特征学习



## Markov随机场

- 图像建模的重要工具，应用广泛.

(J. Besag, 1974)

- 预备知识（标注问题，labeling）

位(site)集合:

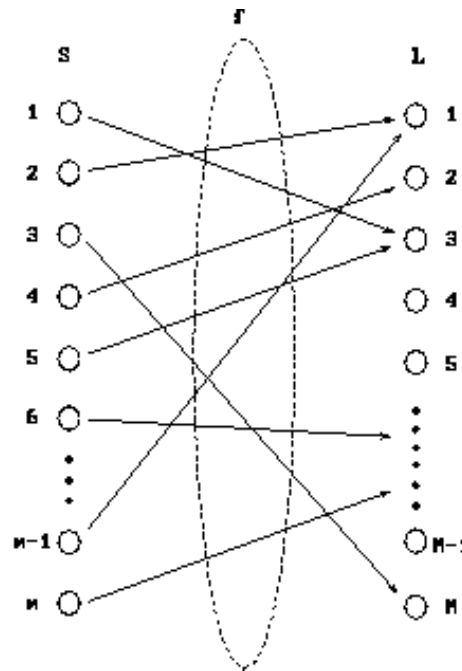
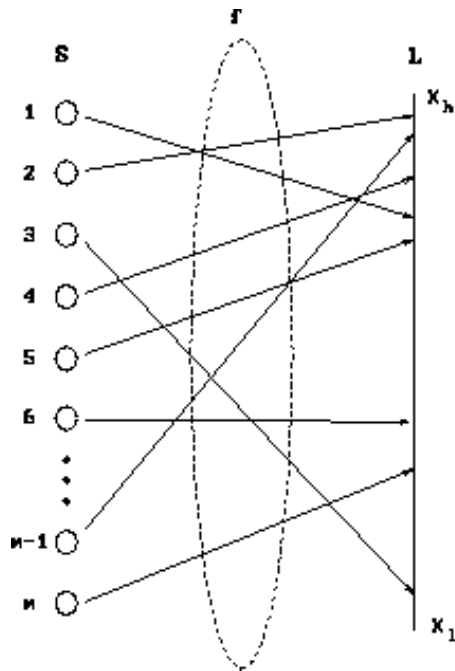
$$S = \{1, 2, \dots, m\}$$

标志(label)集合，位上可能发生事件的集合，可以是连续的，也可以是离散的:

$$L_c = [X_l, X_h] \subset R \quad , \quad L_d = [l_1, l_2, \dots, l_n]$$

# Markov随机场

- 标注：为位集合中每个位指定一个标志的过程，位集合到标志集合的映射：



$$f = (f_1, f_2, \dots, f_m)$$

$$f : S \rightarrow L$$



## Markov随机场

- **标注：** 从如下  $F$  空间中导出  $f$  的过程：

$$F = L_1 \times L_2 \cdots L_m, \text{ 当 } L_1 = L_2 = \cdots = L_m \text{ 时, } F = L^m$$

在图象领域，可将  $f$  理解为一幅图象， $F$  则是全部可允许图像的集合.

- 标注也被称为着色 (coloring, 数学规划) 或配置 (configuration, 随机场)
- 如果各个位为随机变量，则位集合  $S$  称为随机场.

## Markov随机场

- 在随机场中，从  $F$  导出  $f$  的过程就是确定  $f$  出现的概率。
- 假设各个位的标注是彼此无关的，则有

$$P(f) = \prod P(f_i) \quad , \quad P(f_i | \{f_{i'}\}) = P(f_i)$$

只需单独考虑每个位，问题简单（理想）

- 实际应用时，需要考虑上下文约束  
(contextual constraints)  $\rightarrow$  Markov随机场



## Markov随机场

- 当且仅当以下两个条件满足时，随机场为Markov随机场：

$$P(f) > 0$$

正性 (Positivity)

$$P(f_i | f_{S-\{i\}}) = P(f_i | f_{N_i})$$

Markov性 (Markovianity)

- 若 $f_i$ 能够独立发生，那么 $f$ 就能够发生
- 一个像素点的随机概率只与它邻域的像素有关

根据矩阵中各位置与位置*i*的距离，可以将邻域系统表达为等级形式

|   |   |   |          |   |   |   |
|---|---|---|----------|---|---|---|
|   |   |   | 6        |   |   |   |
|   | 5 | 4 | 3        | 4 | 5 |   |
|   | 4 | 2 | 1        | 2 | 4 |   |
| 6 | 3 | 1 | <i>i</i> | 1 | 3 | 6 |
|   | 4 | 2 | 1        | 2 | 4 |   |
|   | 5 | 4 | 3        | 4 | 5 |   |
|   |   |   | 6        |   |   |   |

### 邻域系统的等级划分

一个像素点和图像中其他各像素点的相关性就可以通过条件概率和邻域系统来描述

## Gibbs随机场

- 邻域系统 (neighboring system)

邻域集 (neighbor set):

一阶邻域 (四连通), 二阶邻域 (八连通) 等

团 (cliques):

由邻域关系限定的位子集

单位团 (single-site), 双位团 (pair-site), 三位团 (triple-site) 等

$C_1 = \{i\}, C_2 = \{\{i, i'\}\}, C_3 = \{\{i, i', i''\}\}, i, i', i''$  互为邻居

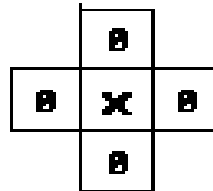
团是有序的:  $\{i, i'\} \neq \{i', i\}$

# Gibbs随机场

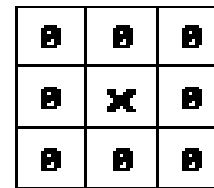
- 邻域

- 团

团具有尺寸,  
形状和方向



(a)



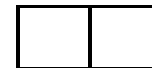
(b)

|   |   |   |   |   |
|---|---|---|---|---|
| 5 | 4 | 3 | 4 | 5 |
| 4 | 2 | 1 | 2 | 4 |
| 3 | 1 | X | 1 | 3 |
| 4 | 2 | 1 | 2 | 4 |
| 5 | 4 | 3 | 4 | 5 |

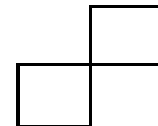
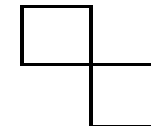
(c)



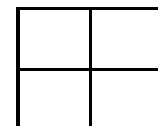
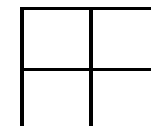
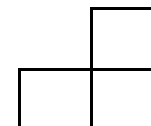
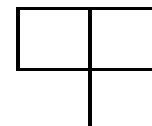
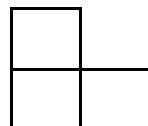
(d)



(e)



(f)



(g)

(h)

(i)

(j)

(k)

(l)

(m)

(n)

(o)

(p)

(q)

(r)

## Gibbs随机场

- 当且仅当随机场的配置服从Gibbs分布时，称为Gibbs随机场：

$$P(f) = \frac{1}{Z} e^{-\frac{1}{T}U(f)} \quad T : \text{温度常量, 常取1}$$

$$Z = \sum_{f \in F} e^{-\frac{1}{T}U(f)} \quad \text{规范化常量, 称为划分函数 (partition function)}$$

$$U(f) = \sum_{c \in C} V_c(f) \quad \text{所有团势能之和, 称为能量函数 (energy function)}$$

$$V_c(f) : \text{团势能 (clique potential)}$$

## Gibbs随机场

- 物理意义

配置的能量越小，其概率越大

- 均匀性 (homogeneity):

$V_c(f)$  与团在随机场中的位置无关

或者说,  $P(f_i | f_{N_i})$  与位  $i$  无关

- 各向同性 (isotropic):

$V_c(f)$  与团的方向无关

- 在纹理领域, Markov (Gibbs) 随机场具有均匀性

# Gibbs随机场

- Hammersley-Clifford定理

Markov随机场与Gibbs随机场等价

- 意义:

既可以用局部成分的相互影响来建模, 也可以用全局能量来建模.

- 如何确定团势能的形式和参数是Markov (Gibbs) 随机场的主要工作.

- 划分函数的计算复杂度很高, 是一个难题, 实际多做一定简化.

## A Typical Example of MRF

### 3-D Depth Reconstruction from a Single Still Image

Ashutosh Saxena, Sung H. Chung, Andrew Y. Ng , ICCV 2007

2006年，斯坦福大学的Andrew Ng等人应用机器学习的方法，采用分层的多尺度马尔科夫随机场模型，通过对大量的训练集数据的学习，建立室外场景图像中各目标间的空间关系模型，并将取得的结果应用在无人汽车的自主导航中。

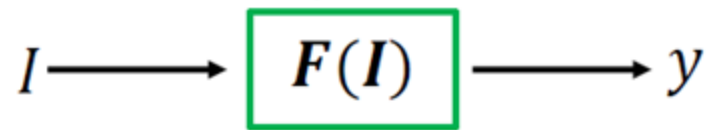
这种利用全局的三维几何纹理信息的方法，实际上就是在利用了图像“质地变化率”的同时也利用了图像中的遮挡信息（空间关系）。



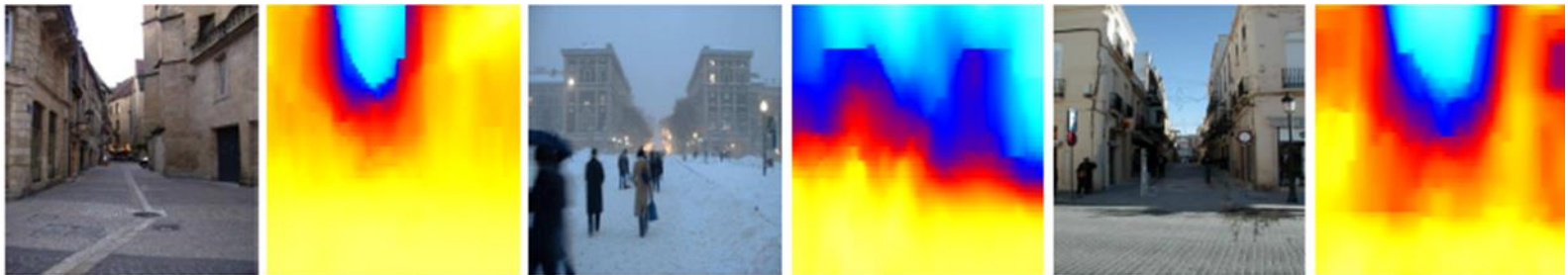
## A Typical Example of MRF

### 3-D Depth Reconstruction from a Single Still Image

Ashutosh Saxena, Sung H. Chung, Andrew Y. Ng, 2007



基本思路：特征提取，图像建模，机器学习



# A Typical Example of MRF

## Visual Cues for Depth Perception

- **Monocular Cues**
  - texture variations, texture gradients, interposition, occlusion, known object sizes, light and shading, haze, defocus, etc.
  - global information, local information.
- **Stereo Cues**
- **Motion Parallax and Focus Cues**

# A Typical Example of MRF

## Feature Vector

- The image is divided into small rectangular patches, and estimate a single depth value for each patch.
- Two types of features: **absolute depth features**—used to estimate the absolute depth at a particular patch—and **relative features**, used to estimate relative depths.

# A Typical Example of MRF

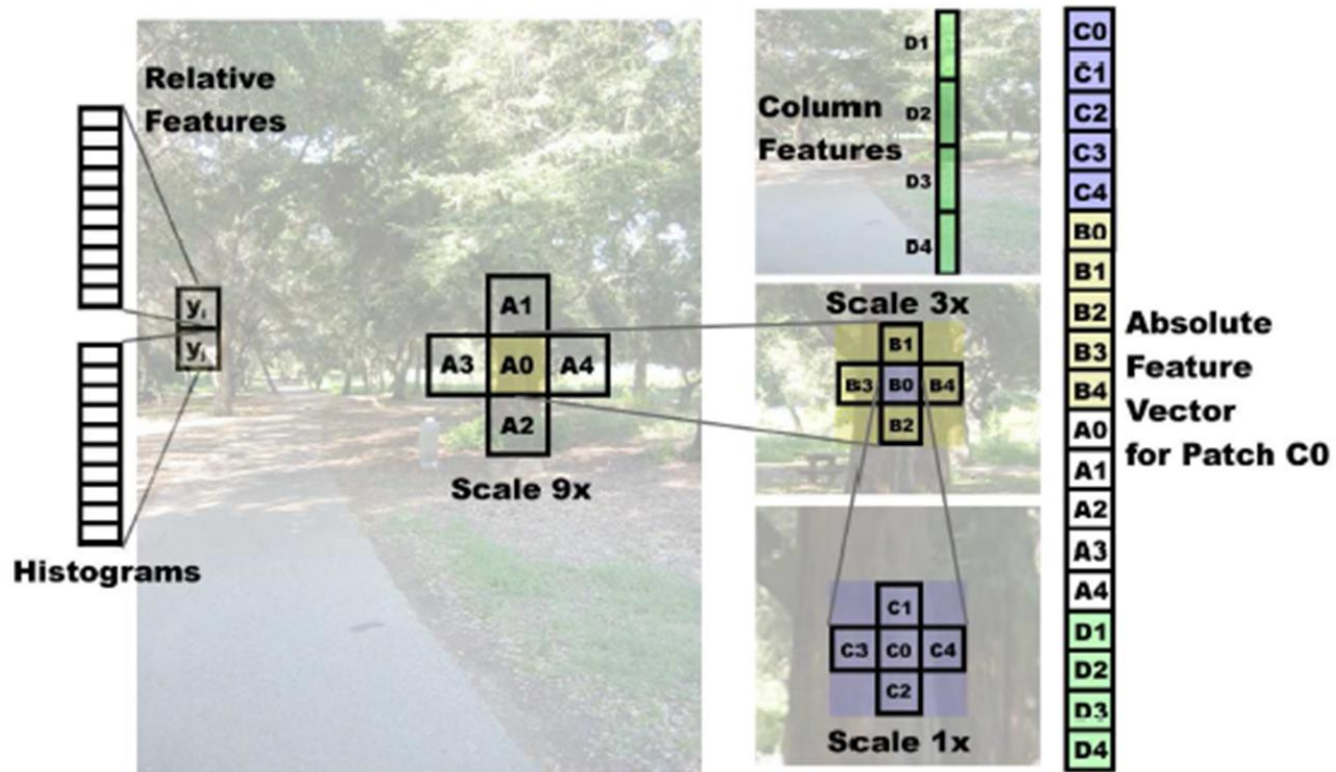
## Feature Vector

- three types of local cues:  
texture variations, texture gradients, color.
- Capture cues
  - Laws' masks(3\*3)----- texture variations
  - Oriented edge filters(6)----- texture variations
  - Color channel(2)----- color



# A Typical Example of MRF

## Feature for absolute depth



## A Typical Example of MRF

### Feature for absolute depth

- **initial feature vector**---sum absolute energy and sum squared energy,
- **multiple spatial scales**---capture more global properties of the image,
- **immediate neighbors**---capture occlusion relationships,
- **vertically column**---show vertical structure.
- absolute depth feature vector  $x$  is  $19 * 34 = 646$  dimensional.

## A Typical Example of MRF

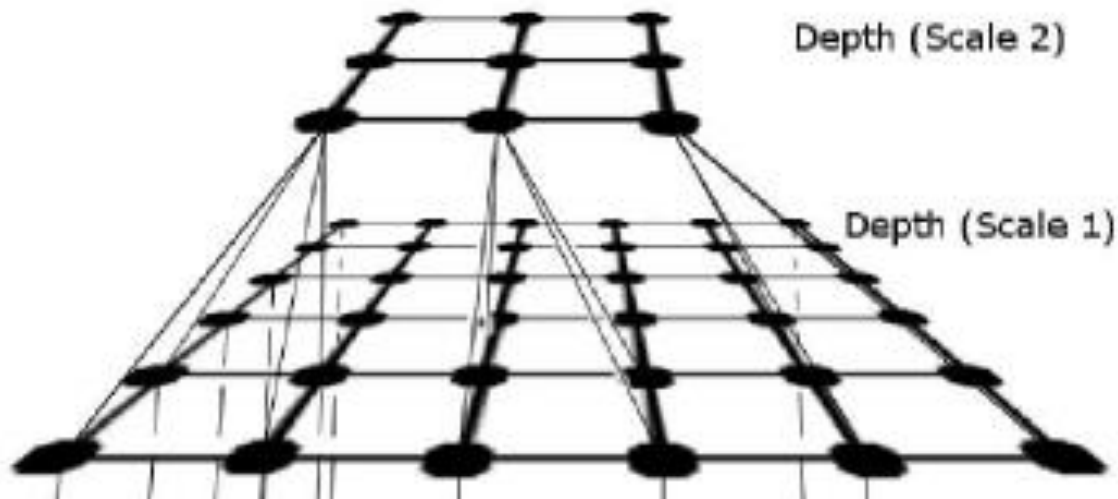
### Feature for relative depth

- A different feature vector is used to learn the dependencies between two neighboring patches.
- A 10-bin histogram of each of the 17 filter outputs  $|I * F_n|$ , giving us a total of 170 features  $y_{is}$  for each patch  $i$  at scale  $s$ .
- Relative depth features  $y_{ijs}$  for two neighboring patches  $i$  and  $j$  at scale  $s$  will be the differences between their histograms.

$$y_{ijs} = y_{is} - y_{js}.$$

# A Typical Example of MRF

## Probabilistic Model



## Hierarchical Multi-scale MRF Model



# A Typical Example of MRF

## Probabilistic Model

- Gaussian Model

$$P_G(d|X; \theta, \sigma) = \frac{1}{Z_G} \exp \left( - \sum_{i=1}^M \frac{(d_i(1) - x_i^T \theta_r)^2}{2\sigma_{1r}^2} - \sum_{s=1}^3 \sum_{i=1}^M \sum_{j \in N_s(i)} \frac{(d_i(s) - d_j(s))^2}{2\sigma_{2rs}^2} \right)$$

- Laplacian model

$$P_L(d|X; \theta, \lambda) = \frac{1}{Z_L} \exp \left( - \sum_{i=1}^M \frac{|d_i(1) - x_i^T \theta_r|}{\lambda_{1r}} - \sum_{s=1}^3 \sum_{i=1}^M \sum_{j \in N_s(i)} \frac{|d_i(s) - d_j(s)|}{\lambda_{2rs}} \right)$$

# A Typical Example of MRF

## Gaussian Model

- $d_i(s)$  depths for multiple scales  $s = 1, 2, 3$ .
- enforce a hard constraint that depths at a higher scale are the average of the depths at the lower scale.

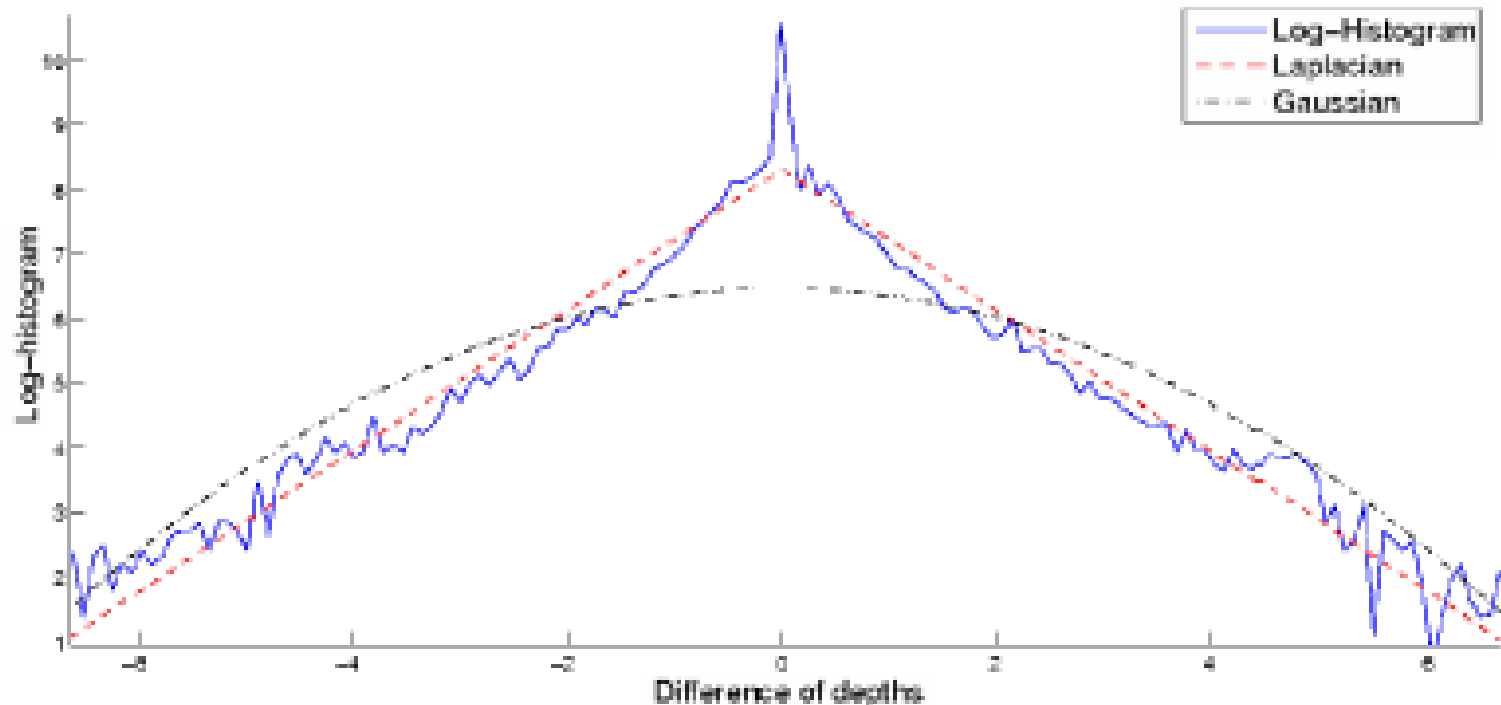
$$d_i(s+1) = (1/5) \sum_{j \in N_s(i) \cup \{i\}} d_j(s)$$

- $N_s(i)$  are the 4 neighbors of patch  $i$  at scale  $s$ .

$$P_G(d|X; \theta, \sigma) = \frac{1}{Z_G} \exp \left( - \sum_{i=1}^M \frac{(d_i(1) - x_i^T \theta_r)^2}{2\sigma_{1r}^2} - \sum_{s=1}^3 \sum_{i=1}^M \sum_{j \in N_s(i)} \frac{(d_i(s) - d_j(s))^2}{2\sigma_{2rs}^2} \right)$$

# A Typical Example of MRF

## Laplacian Model



# A Typical Example of MRF

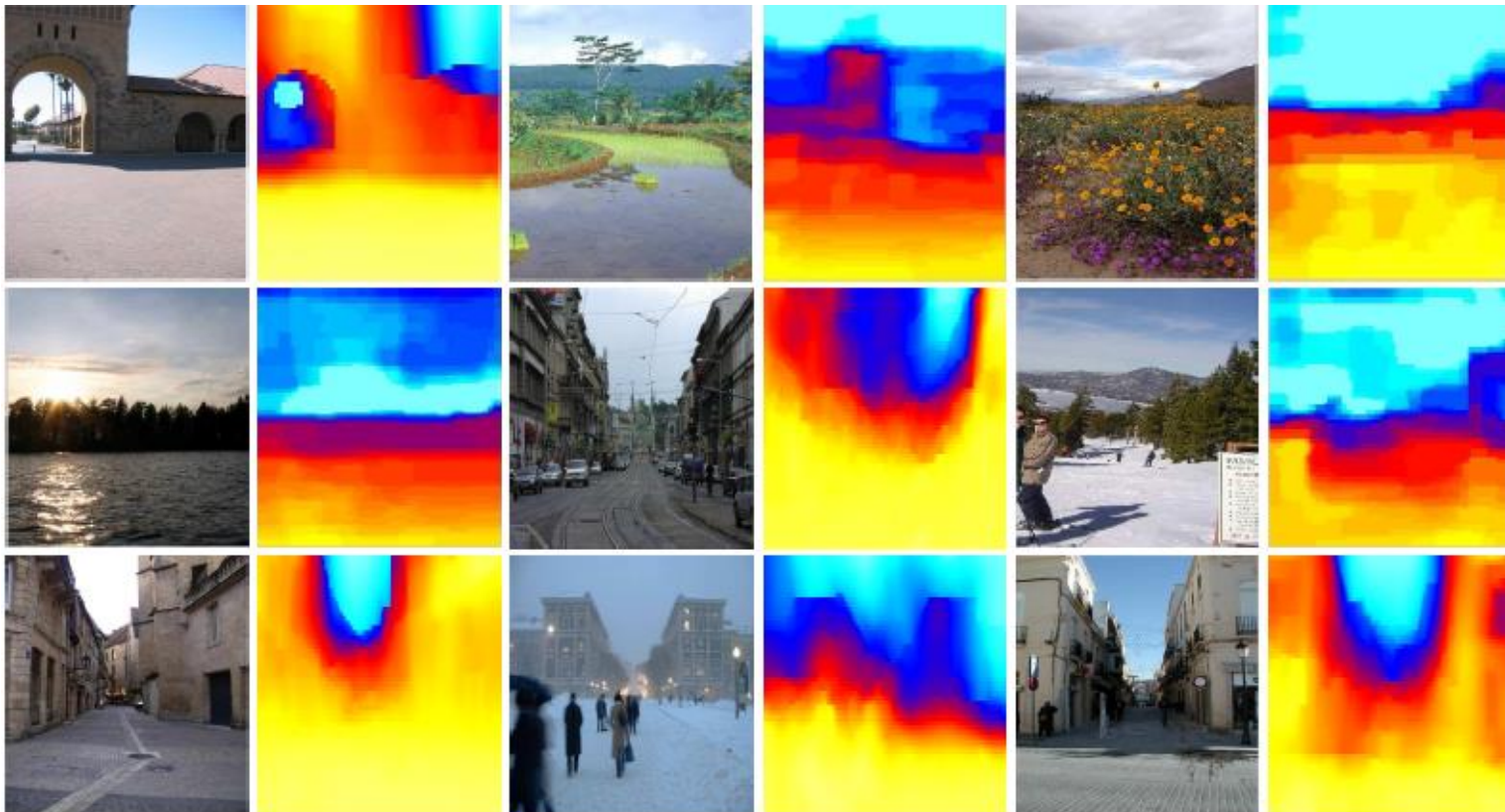
## Experiments

- Data collection
  - using a 3-d laser scanner to collect images and their corresponding depthmaps



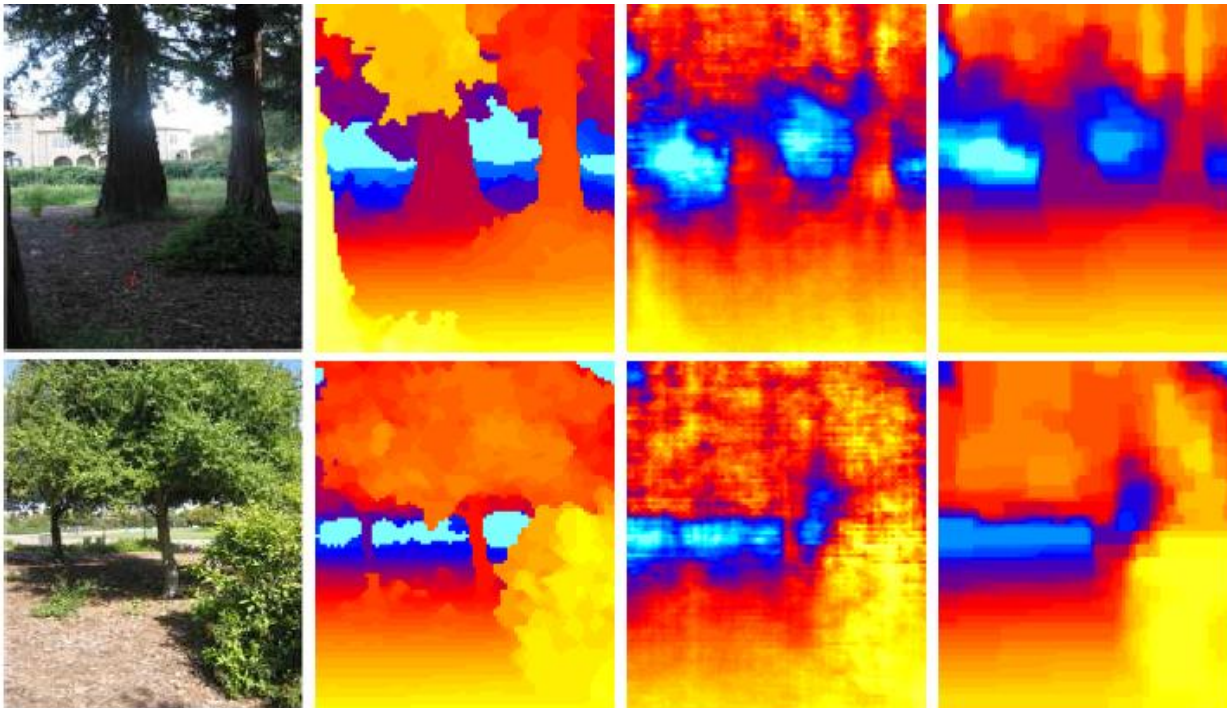
# A Typical Example of MRF

## Experiments



# A Typical Example of MRF

## Experiments

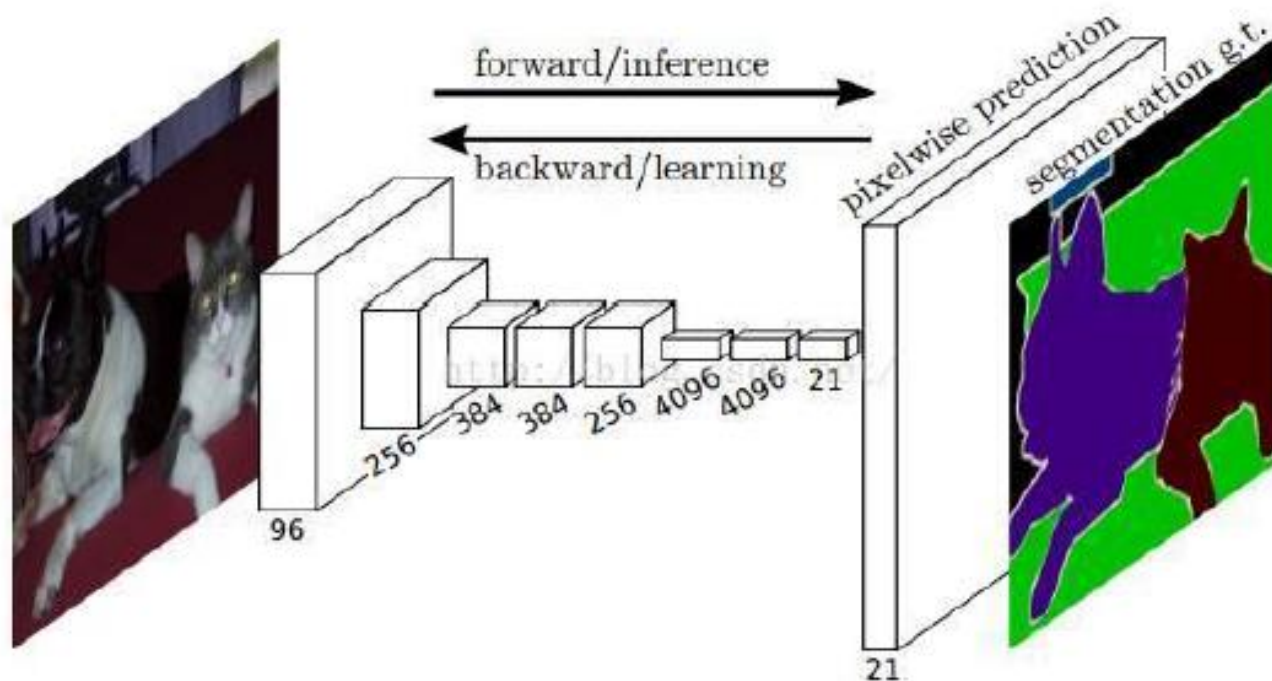


(a)original image, (b) ground truth depthmap, (c)predicted depthmap by Gaussian model, (d) predicted depthmap by Laplacian model.

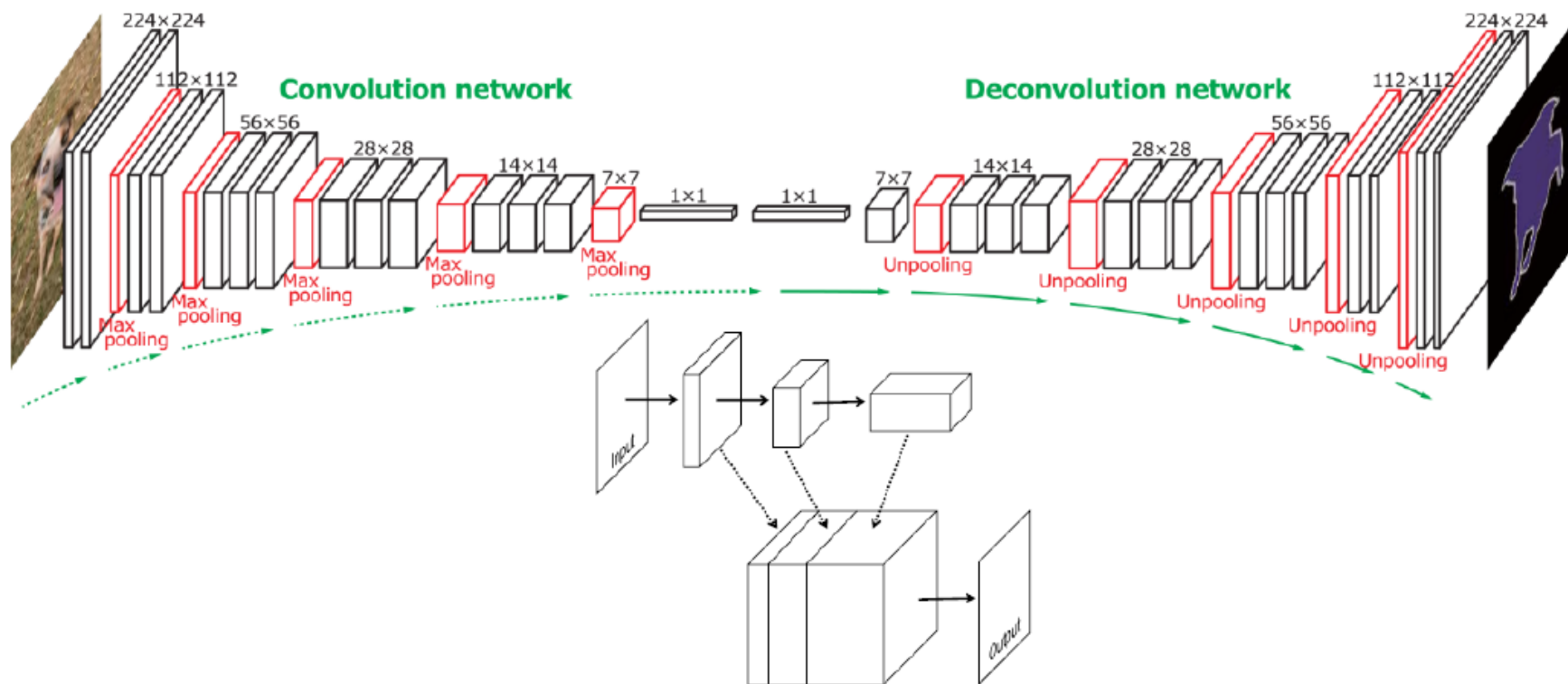


## 卷积神经网络的实例-FCN网络

- 网络基本架构
  - 全连接层->卷积层
  - 1000维向量->原图大小



# 卷积神经网络的实例-FCN网络

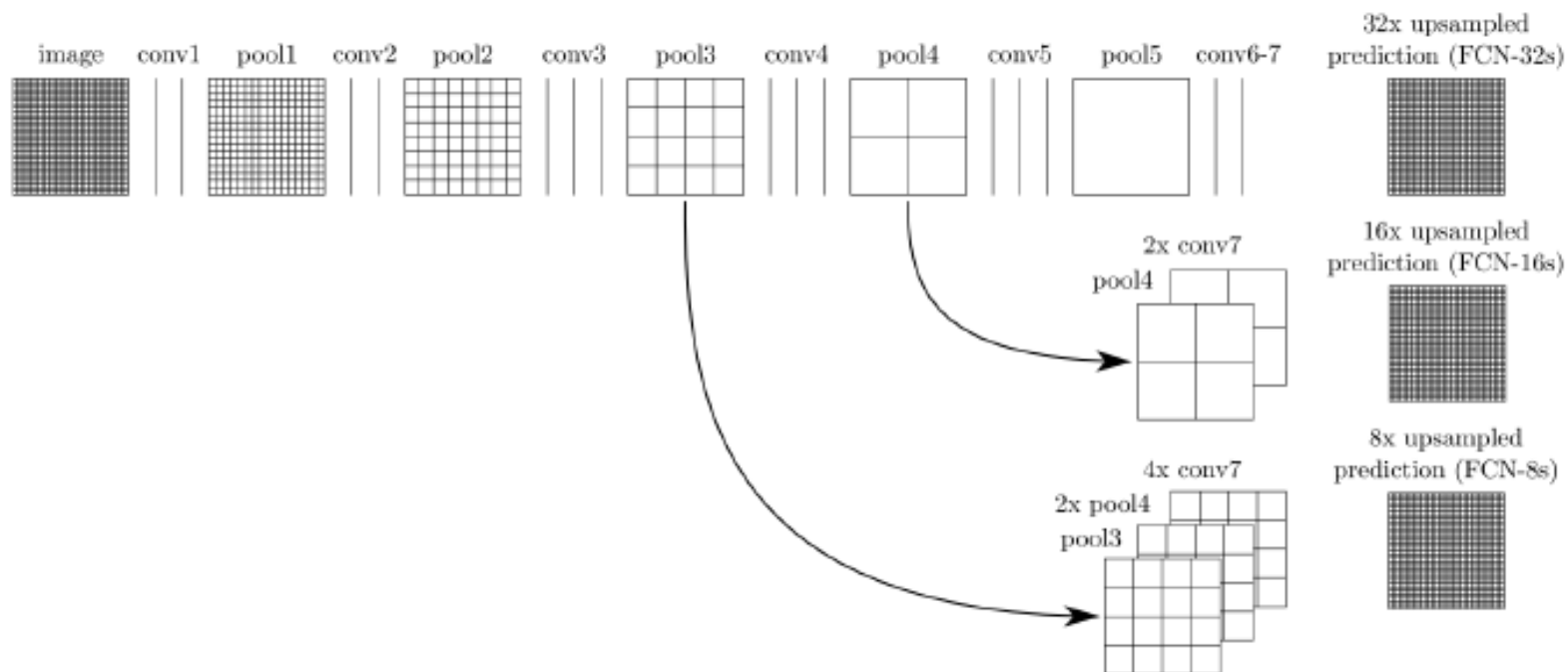




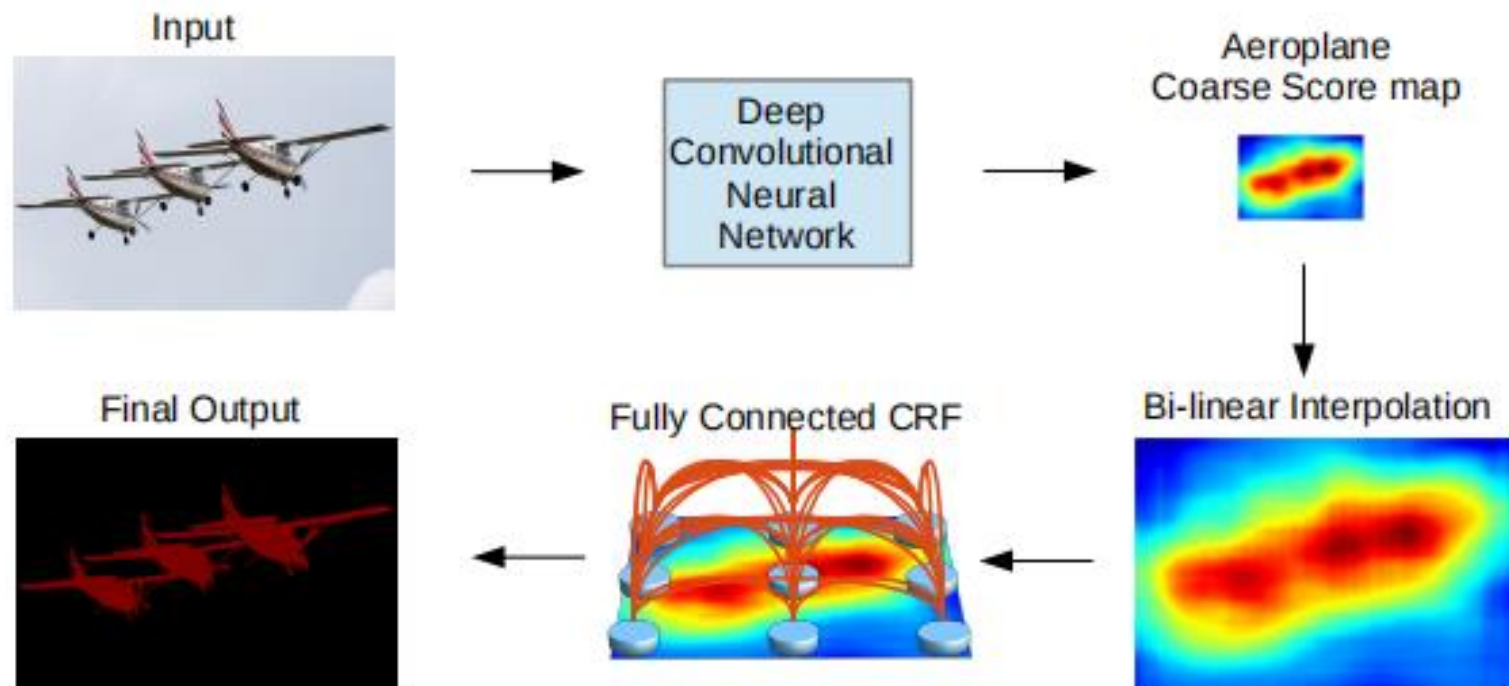
# 卷积神经网络的实例 - FCN网络

## ● Deconv

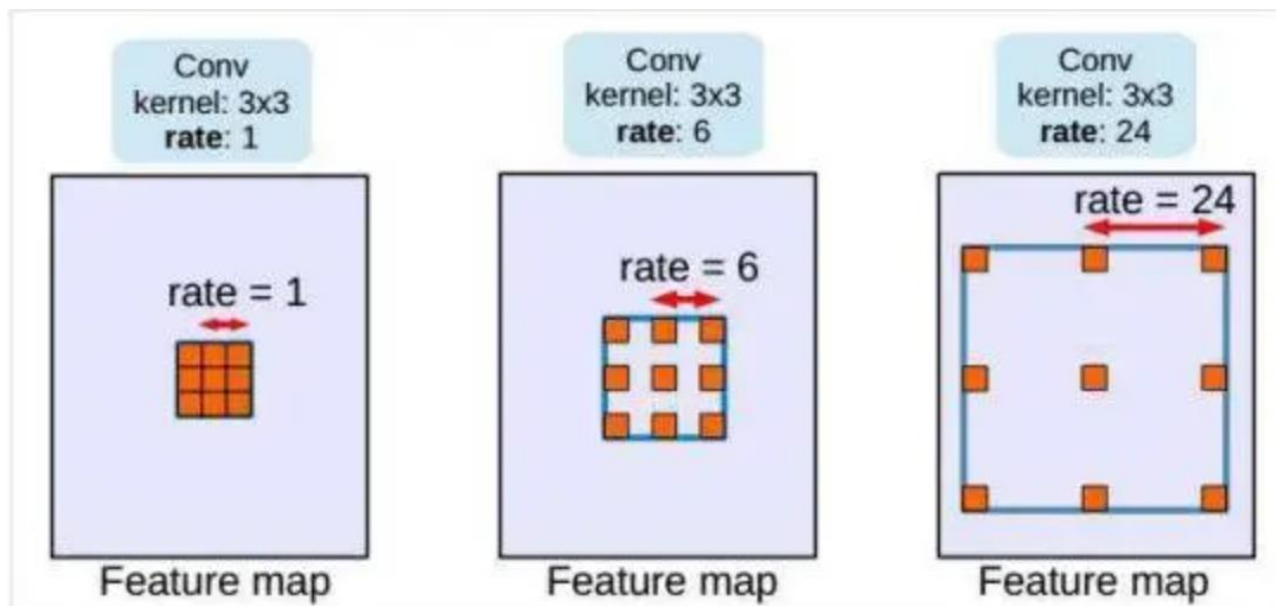
- 简单上采样结果很差
- 可以看成引入了高层信息
- 局部信息有所丢失



## 卷积神经网络的实例 - Deeplab系列



## 卷积神经网络的实例 - Deeplab系列



### 空洞卷积

作用：扩大感受野；捕获多尺度上下文信息

不足：kernel不连续导致的网格效应；小目标性能有限

## 卷积神经网络的实例 - Deeplab系列

CRF模型的能量函数:  $E(x) = \sum_i \theta_i(x_i) + \sum_{ij} \theta_{ij}(x_i, y_j)$

$$\theta_i(x_i) = -\log P(x_i)$$

刻画观测序列对标记变量的影响

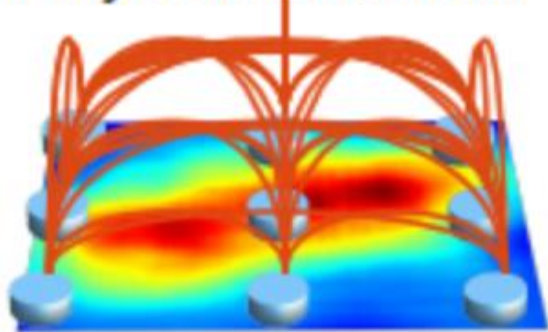
$$\theta_{ij}(x_i, y_j) = u(x_i, x_j) \sum_{m=1}^K \omega_m \cdot k^m(f_i, f_j)$$

刻画像素之间的相关性对标记变量影响,  $k$ 的权重为:

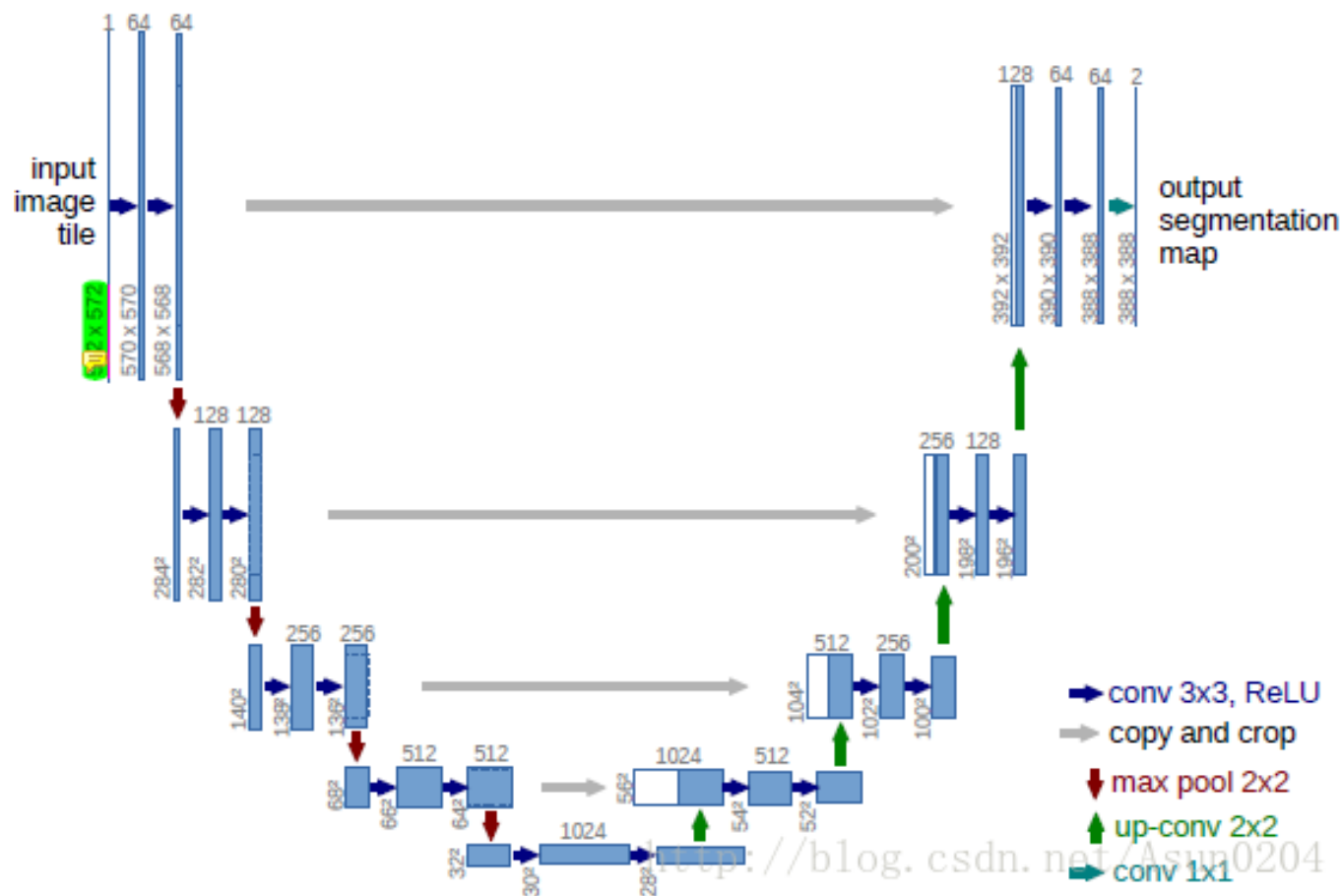
$$\omega_1 \exp\left(\frac{\|p_i - p_j\|^2}{2\sigma_\alpha^2} - \frac{\|I_i - I_j\|^2}{2\sigma_\beta^2}\right) + \omega_2 \exp\left(-\frac{\|p_i - p_j\|^2}{2\sigma_\gamma^2}\right)$$

条件随机场-CRF

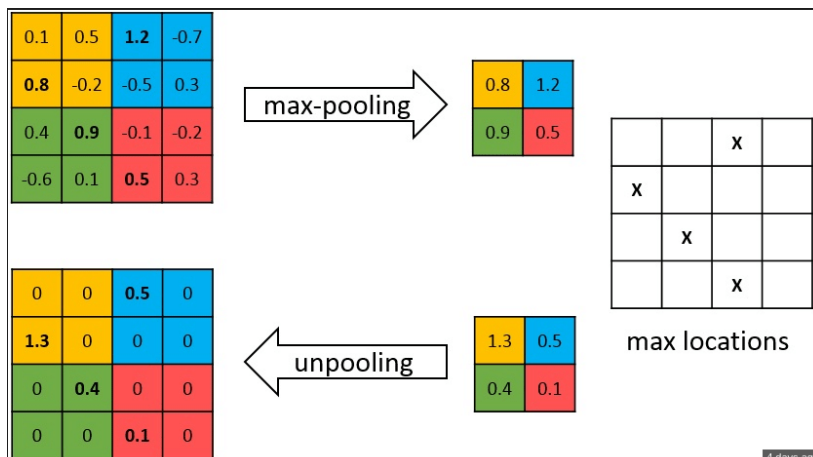
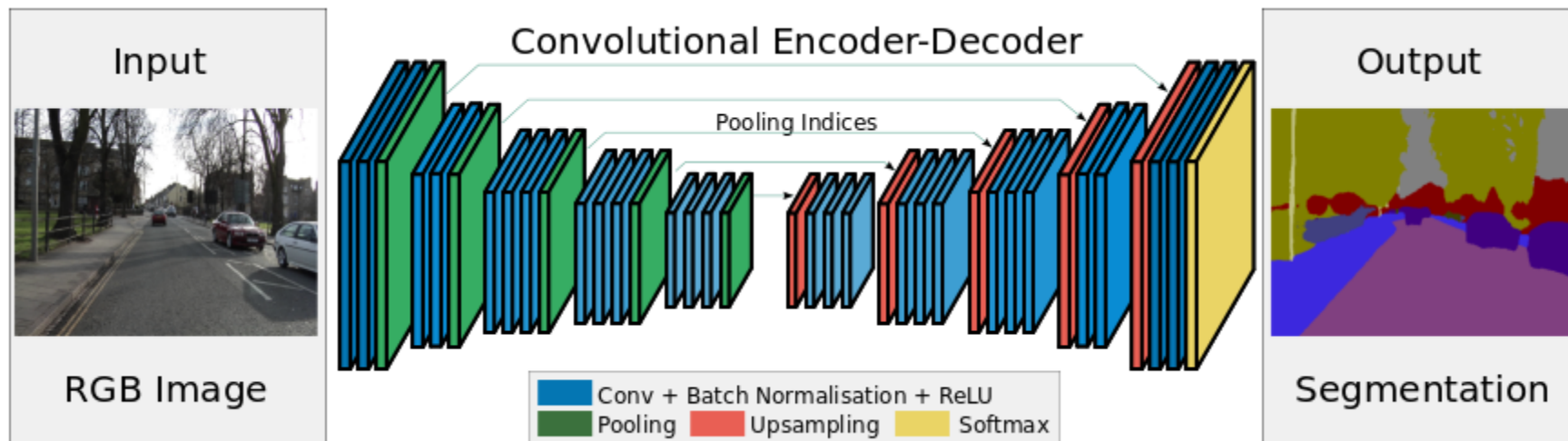
Fully Connected CRF



# 卷积神经网络的实例 - U-Net



# 卷积神经网络的实例 – Seg-Net



**SegNet** 在解码器中使用反池化对特征图进行上采样，并在分割中保持高频细节的完整性。

编码器中的每一个最大池化层的索引都被存储起来，用于之后在解码器中使用那些存储的索引来对相应的特征图进行反池化操作。