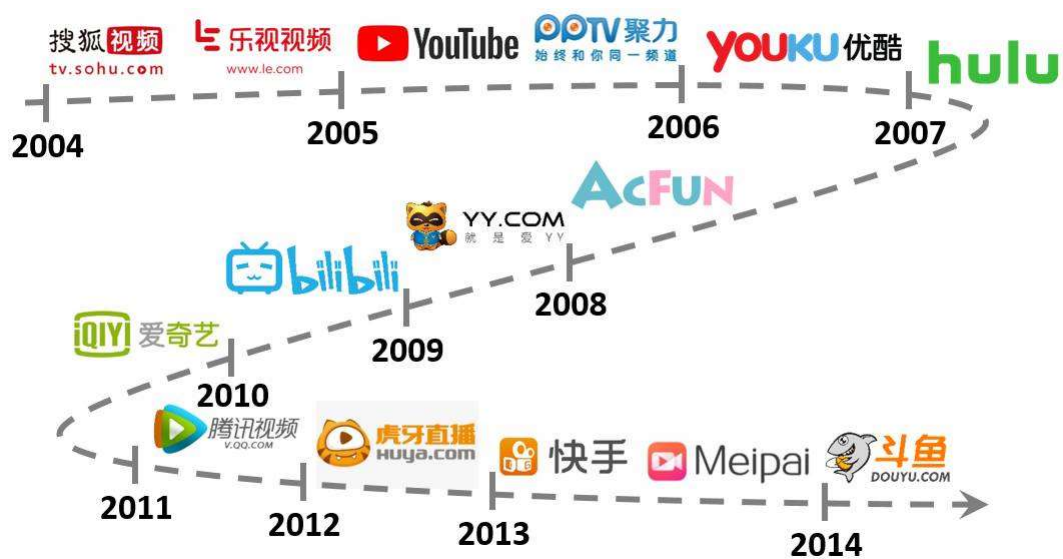




## PM5544电视测试卡

(PM5544 TV test card clock)



19世纪末期著名实验-运动感知



Think

720P, 1080P?

P的含义?

▼ 格式 M4S(mp4) ▼ 视频 00:07:57

🔥 新版推荐 高清 1080P [ 00:07:57 ]

🌈 视频 高清 720P [ 136 kbps ]

🌈 视频 1080P [ 191 kbps ]



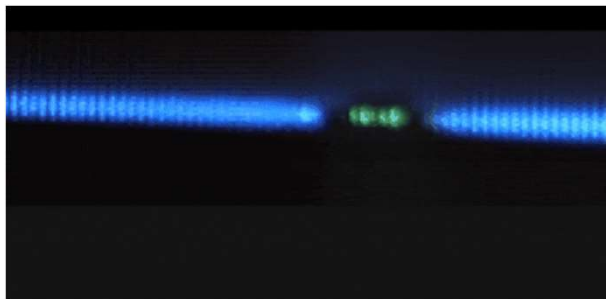
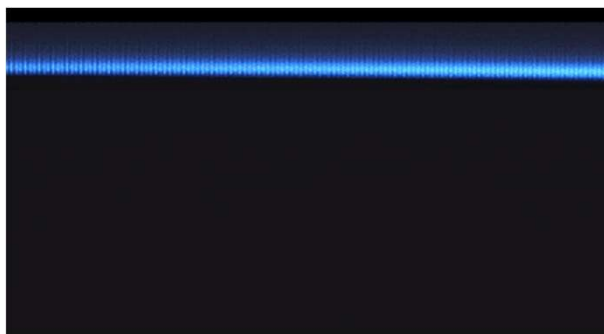
全部磁力链	4K	蓝光原盘	蓝光高清	WEB-DL	字幕下载
[无损音乐1T] [极品4K 28T] [美剧蓝光32T] [观影APP] [115网盘优惠充值]					
4K	73.84GB	True.Lies.1994.UHD.BluRay.2160p.TrueHD.Atmos.7.1.DV.HEVC.REMUX-FraMeSToR.mkv			
4K	25.01GB	True.Lies.1994.HDR.2160p.WEB.H265-SLOT			
4K	20.94GB	True.Lies.1994.2160p.WEB.H265-SLOT			
4K	15.5GB	True.Lies.1994.2160p.AMZN.WEB-DL.DDP5.1.Atmos.DV.HDR.H.265-FLUX.mkv			
4K	40.52GB	True.Lies.1994.2160p.BluRay.DoVi.x265.10bit.Atmos.TrueHD7.1-CTRLHD			
Remux	27.53GB	True.Lies.1994.1080p.BluRay.Remux.DTS-HD.5.1			
Remux	27.56GB	True.Lies.1994.1080p.BOOTLEG.BluRay.REMUX.AVC.DTS-HD.MA.5.1-FGT			
Remux	27.53GB	True.Lies.1994.1080p.BluRay.Remux.DTS-HD.5.1			
Remux	27.56GB	True.Lies.1994.1080p.BOOTLEG.BluRay.REMUX.AVC.DTS-HD.MA.5.1-FGT			
蓝光原盘	35.91GB	True.Lies.1994.1080p.BluRay.AVC.DTS-HD.MA.5.1-NOGRP			
蓝光原盘	35.91GB	True.Lies.1994.1080p.BluRay.AVC.DTS-HD.MA.5.1-NOGRP			
1080P	15.92GB	True.Lies.1994.DTheater.1080p.DualAudio.x264-CHDBits			
720P	603.82M	True.Lies.1994.720p.BrRip.x265.HEVCBay.org.mkv			

4K 超清	大会员
1080P 高码率	大会员
1080P 高清	登录即享
720P 高清	登录即享
480P 清晰	登录即享
360P 流畅	
自动(360P)	



## 视频的扫描线

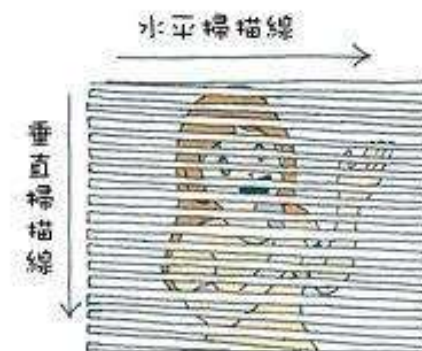
720P, 1080P?





## 扫描线

画面清晰度是以水平清晰度作为单位。  
清晰度的单位是“电视行(TVLine)”  
也称线。



### 隔行扫描

i

#### 【隔行扫描】

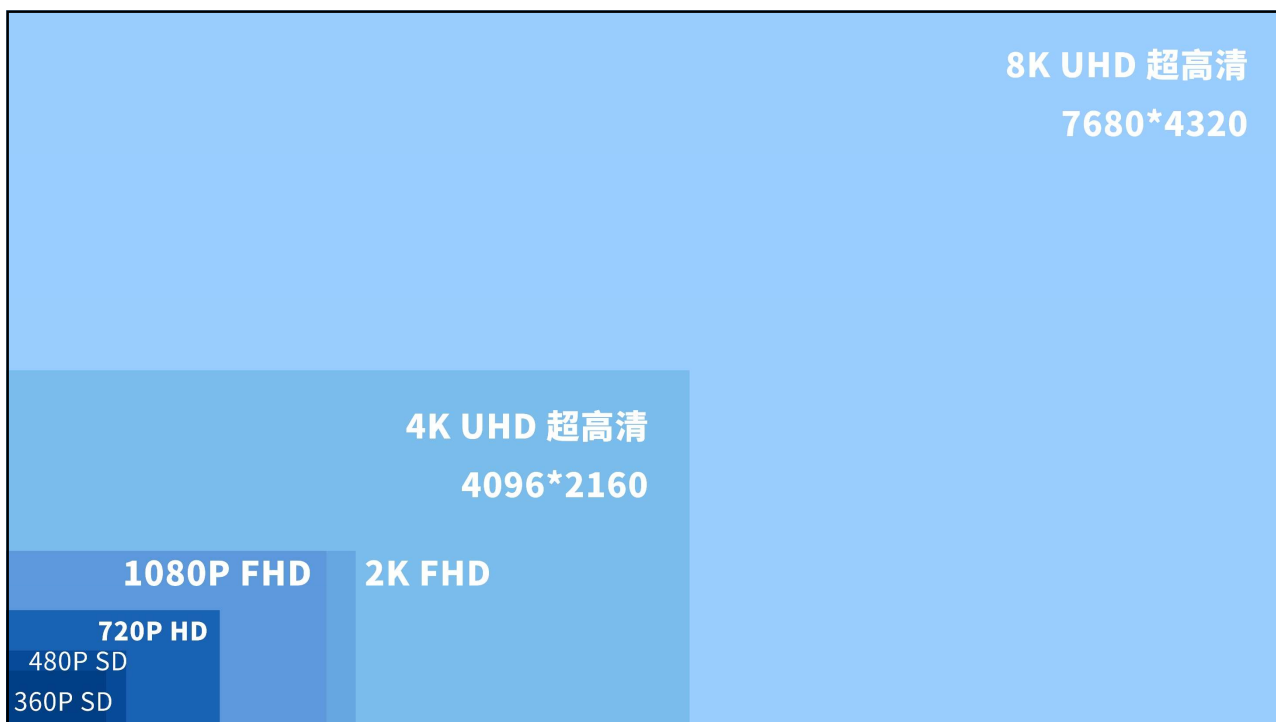
一帧图像由两部分组成:奇数场和偶数场。能够提供更流畅的运动画面,尤其在处理运动中物体的外观时表现更好。然而,隔行扫描可能导致图像上的闪烁和扭曲。

### 逐行扫描

P

#### 【逐行扫描】

图像显示效果更加清晰自然,稳定性更好,动态失真程度低,但在处理快速移动的图像时可能会出现残影。



		D1	D2	D3	D4	D5	4K	8K
	576i	480i	480p	1080i	720p	1080p		
水平扫描线	625	525	525	1125	750	1125		
垂直扫描线	576	480	480	1080	720	1080		
分辨率			640×480	1920×1080	1280×720	1920×1080	4096*2160	7680*4320
画面比例	4: 3 16: 9	4: 3 16: 9	4: 3 16: 9	16: 9	16: 9	16: 9		
行频	隔行/50Hz	隔行/50Hz	逐行/60Hz	隔行/60Hz	逐行/60Hz	逐行/60Hz		
行频		15.75KHz	31.5KHz	33.75KHz	45KHz			
	标准的PAL							





4K->8K->16K->?

**Think**

黑白电视怎么显示呢?

没有颜色?





## 彩色电视的颜色空间

视频及其应用

### ■ 彩色电视中的彩色空间变换

- 用RGB彩色空间表示的彩色信号需经过变换电路转换到一个亮度信号Y和二色差信号C1、C2的彩色空间，再进行调制，然后进行传输。
- C1、C2的含义与具体应用有关：
  - PAL中，C1、C2表示U、V；NTSC中C1、C2表示I、Q。
  - 不同的彩色空间在不同的应用中也许会比原始的RGB彩色空间具有更有用的特性。

### ■ 彩色电视中使用Y、C1、C2彩色空间的优点:

- (1) Y 和C1、C2独立，使得黑白电视和彩色电视都可以接收彩色电视信号，并且黑白电视机可以直接使用亮度信号Y；
- (2) 利用人眼对亮度较为敏感，而对色度相对不够敏感的特性，来节省信号的带宽和功率。

一个亮度Y + 二个色差C1、C2





## Y、C1、C2 颜色空间

视频及其应用

### 彩色空间之间的转换

- 用8位二进制数表示BT.601的Y'CbCr和R'G'B'，而R'G'B'颜色空间使用相同数值范围[0, 219]的分量信号，R'G'B'和Y'CbCr两个彩色空间之间的转换关系用下式表示:

$$Y = 0.299R + 0.587G + 0.114B + 16$$

$$Cr = (0.500R - 0.4187G - 0.0813B) + 128$$

$$Cb = (-0.1687R - 0.3313G + 0.500B) + 128$$

## 接口分辨率

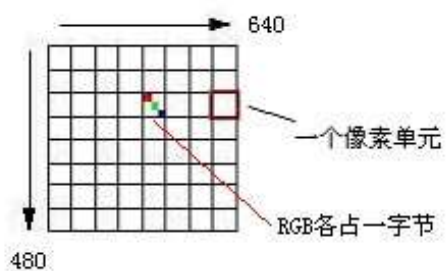
- AV接口:
  - 340×288 (576P)
- S端子:
  - 640×480 (576P)
- 色差分量:
  - 1280×720 (1080P)
- VGA:
  - 1920×1080 (1080P)



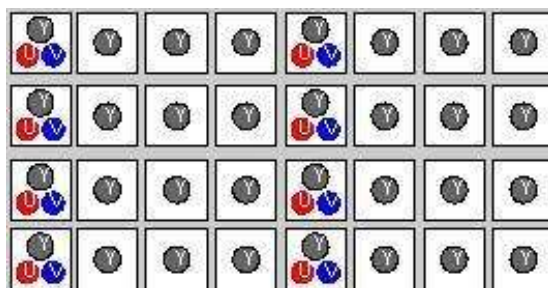


## 图像子采样

图像子采样定义和理论依据



$$640 \times 480 \times (1+1+1) = 921600 \text{ 字节}$$



$$640 \times 480 \times (1+1/(2*2)+1/(2*2)) = 460800 \text{ 字节}$$



# 色差比亮度低

色差信号使用的采样频率比对亮度信号使用的采样频率低

## 图像子采样

视频及其应用

### ■ 一、图像子采样的定义

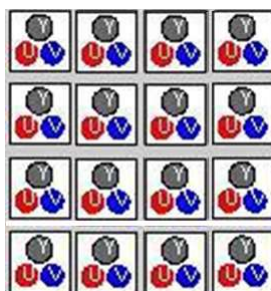
- 对**色差信号**使用的采样频率比对**亮度信号**使用的采样频率**低**的采样称为图像子采样。

### ■ 二、图像子采样的**理论依据与用途**:

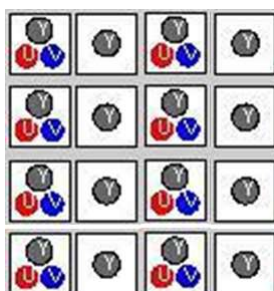
- 1、人的视觉系统的特点(图像子采样的**理论依据**):
  - 人眼对色度信号的敏感程度比对亮度信号敏感的程度低;
  - 人眼对图像细节的分辨能力有限, 取掉一些高频信号后人眼不易察觉。
- 2、图像子采样的用途:
  - 根据人的视觉系统的特点, 可以通过图像子采样实现彩色电视图像数据的压缩。



**4:4:4**



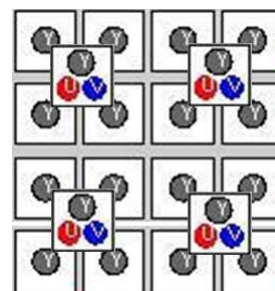
**4:2:2**



**4:1:1**



**4:2:0**



注：扫描线上每4个连续的采样点



## 视频编码

为什么巨大的原始视频可以编码成很小的视频呢？这其中的技术是什么呢？



Think

如何得到如此高的压缩比?



## 数据压缩算法

视频及其应用

- 视频图像压缩算法利用的各种冗余信息:

种类		内容	目前的主要用法
统计特性	空间冗余 时间冗余	像素间的相关性 时间方向上的相关性	变换编码, 预测编码 帧间预测, 移动补偿
图象构造冗余 知识冗余 视觉冗余 其他		图象本身的构造 收发两端对人物的共有认识 人的视觉特性 不确定因素	轮廓编码, 区域分割 基于知识的编码 非线性量化, 位分配。



## 基本原理

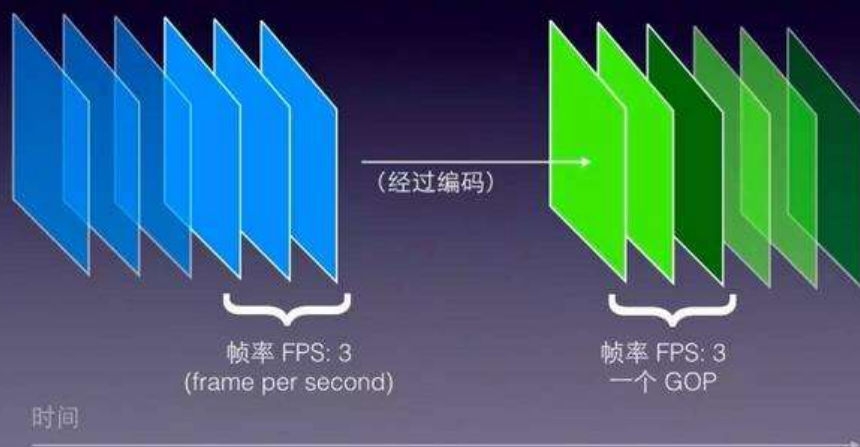
视频及其应用

### ■ 核心思想就是去除冗余信息:

- 空间冗余
  - 图像相邻像素之间有较强的相关性
- 时间冗余
  - 视频序列的相邻图像之间内容相似
- 编码冗余
  - 不同像素值出现的概率不同
- 视觉冗余
  - 人的视觉系统对某些细节不敏感
- 知识冗余
  - 规律性的结构可由先验知识和背景知识得到

**去除冗余!**

## 编码时发生了什么





## MPEG Video

### MPEG Video的数据压缩依据和方法

## MPEG Video的数据压缩依据和方法

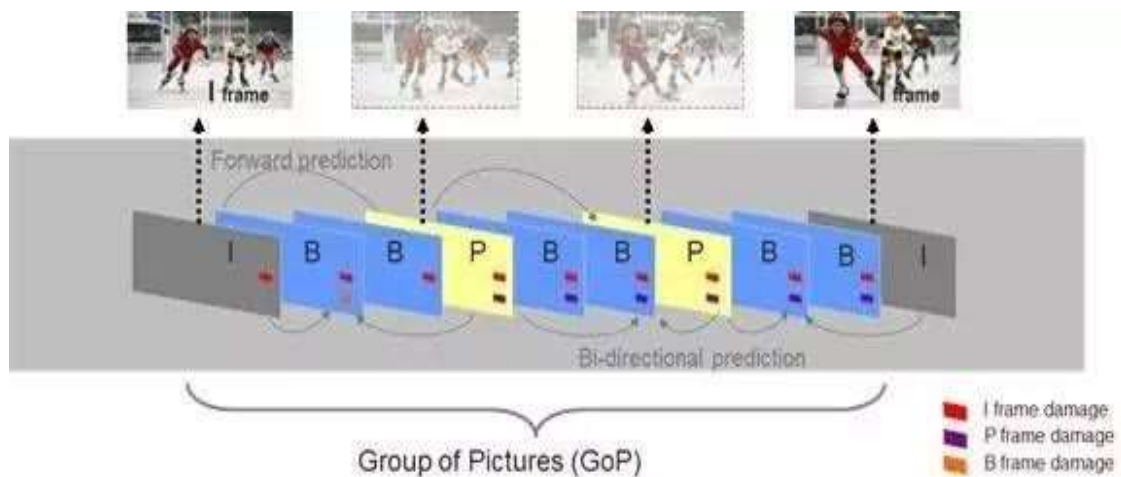
视频及其应用

### ■ 数据压缩的**依据**

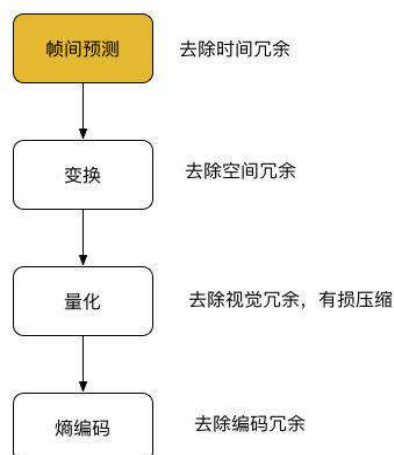
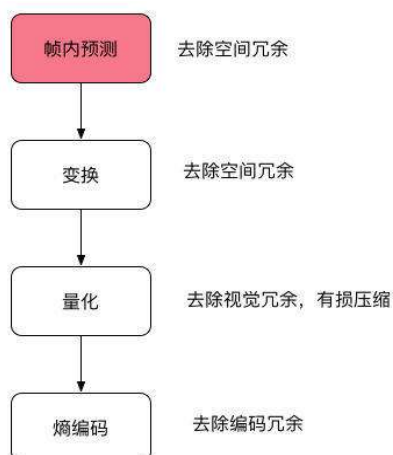
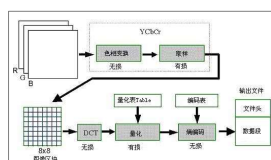
- 时间冗余、空间冗余、结构冗余、视觉冗余等多种冗余信息。
- 通过消除这些冗余就可以实现图像数据的压缩。

### ■ **基本思想和方法要点:**

- 在空间方面：对**单帧**图像数据，采用**JPEG压缩**算法来去掉帧内的冗余信息；
- 在时间方面：对连续的**多帧图**像数据，采用**移动补偿算法**来去掉帧间的冗余信息。



GOP是画面组，一个GOP就是一组连续的画面

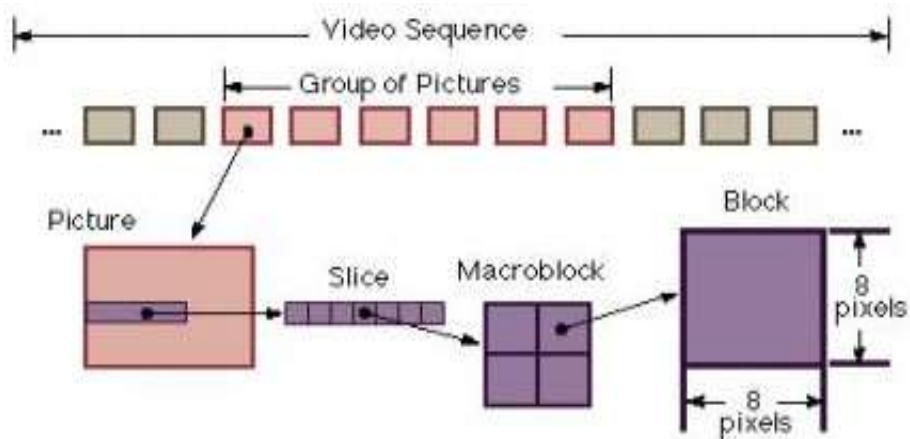






## MPEG视频的层次

视频及其应用

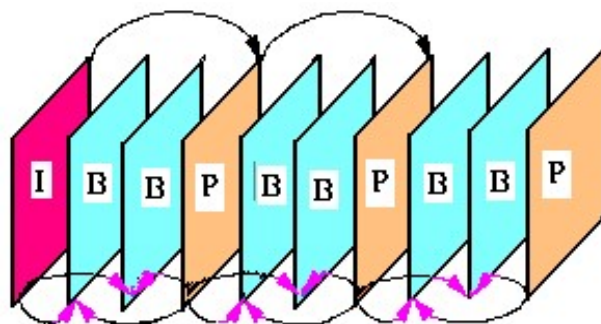


## 运动图像分类及排序

视频及其应用

### 三种图像(采用不同的方法)

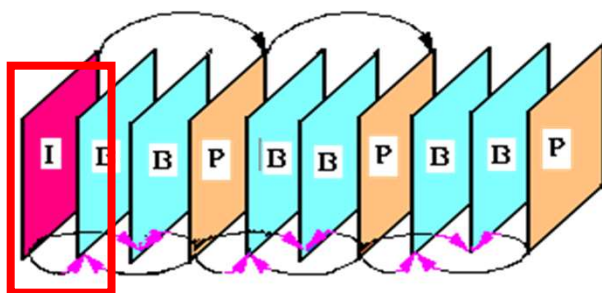
- 帧内图 (Intra Pictures **I**),
- 预测图 (Predicted Pictures **P**)
- 插补图 (Bidirectional Prediction **B**) 即双向预测图





## 1.内帧(I)

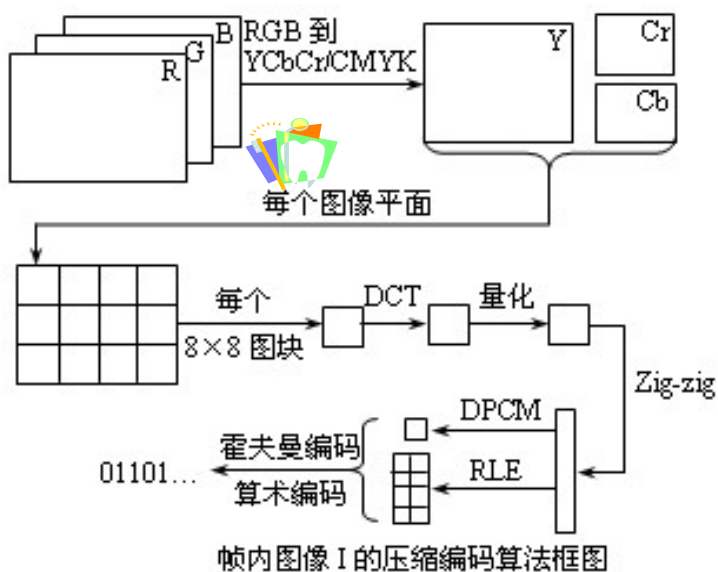
视频及其应用



- 采用JPEG算法进行中等程度的压缩压缩；
- 内帧压缩算法在编解码时不依赖其它任何帧，可作为随机访问点。

## 帧内图像I的压缩编码(类似JPEG压缩算法)

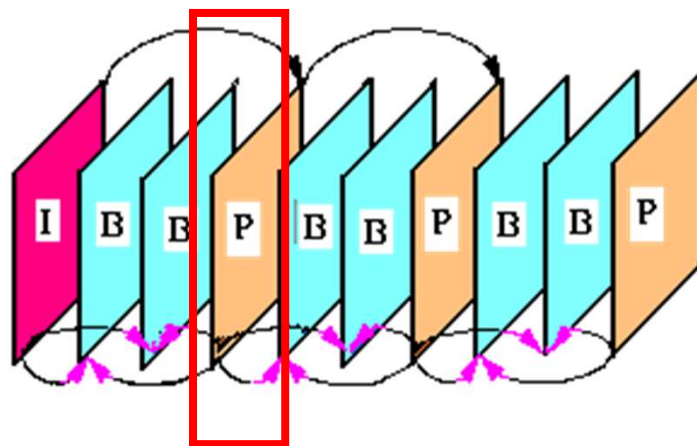
视频及其应用





## 2.预测帧(P)

视频及其应用



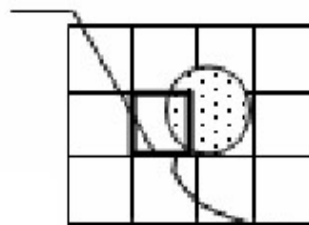
## 2.预测帧(P)

视频及其应用

■ 以前面的参照帧(I或P)为基础进行预测编码和运动补偿法压缩存储，它本身又可作为后面预测帧的参照帧。

- 编码与JPEG类似
- 以图像宏块为基本编码单元，一个宏块定义为 $I \times J$ 像素的图像块
- 以图像宏块为基本编码单元，一般 $16 \times 16$

$M_{PT}$  编码图像





## 对P图像的压缩编码

# 差值

# 运动矢量

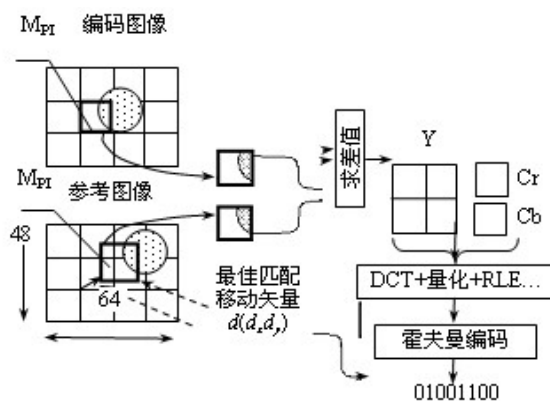
- 两种类型的参数:
  - 1)当前要编码的图像宏块与参考图像的宏块之间的**差值**
  - 2)宏块的**运动矢量**

## 解差值的方法

视频及其应用

■ 假设编码图像宏块MPI是参考图像宏块MRJ的最佳匹配块，它们的差值就是这两个宏块中相应像素值之差。

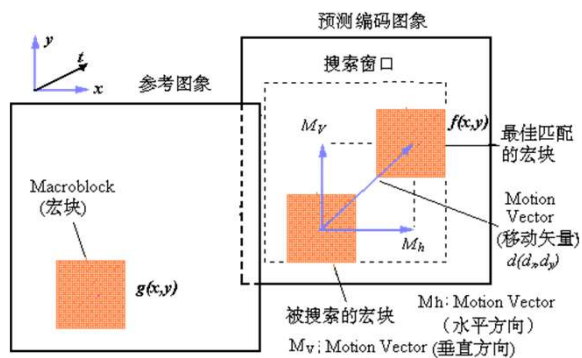
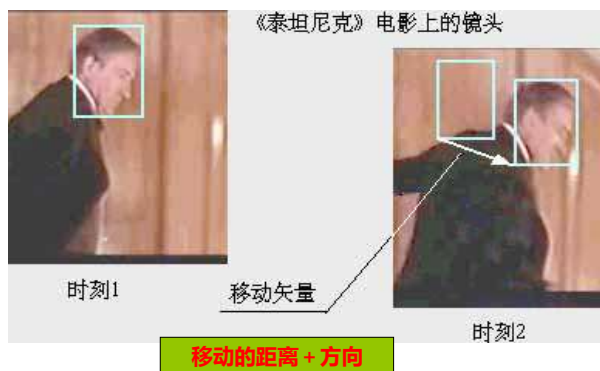
- 对所求得的差值进行彩色空间转换，并作4:1:1的子采样得到Y, Cr和Cb分量值
- 仿照JPEG压缩算法对差值进行编码
- 计算出的移动矢量也要进行霍夫曼编码



预测图像 P 的压缩编码算法框图



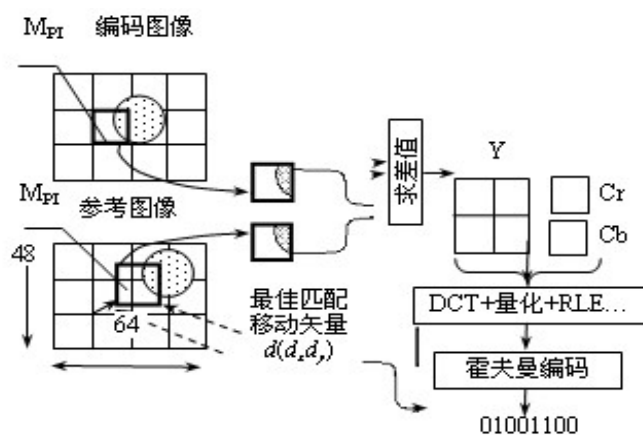
## 宏块的移动矢量



**查找与参考宏块最佳的预测图像编码宏块**



## 关键:查找与参考宏块最佳的预测图像编码宏块



预测图像 P 的压缩编码算法框图

## 寻找最佳宏块的三种简化算法

视频及其应用

### ■ 对预测帧的编码:

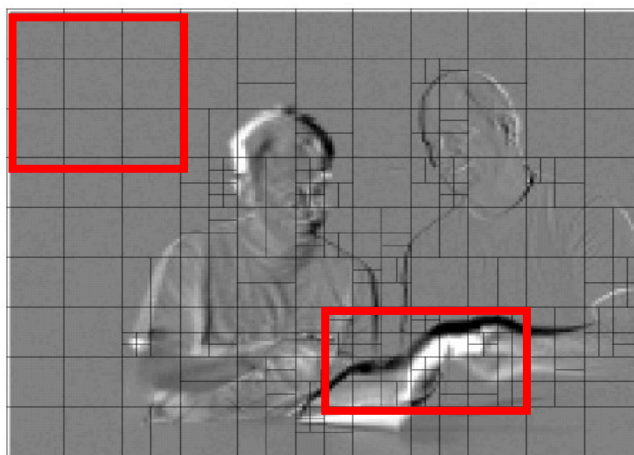
- 寻找**最佳匹配图像宏块**
- 找到**最佳移动矢量** $d(d_x, d_y)$

### ■ 主要算法:

- 1)二维对数搜索法(2D - Logarithmic Search)
- 2)三步搜索法(Three-step Search)
- 3)对偶搜索法(Conjugate Search)



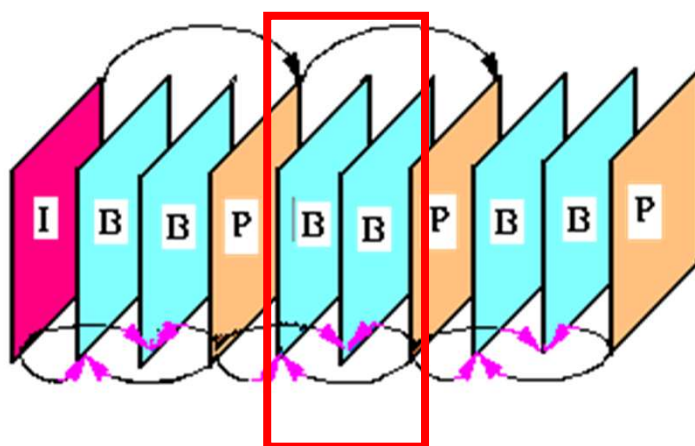
## H.264编码器：每个部分选择了最佳的分割尺寸



在帧内图像变化小的区域，即图像平坦区域(残差显示为灰色)，采用了 $16 \times 16$ 块分割。而在图像复杂区域(残差显示为黑色或者白色)，选择了更小的分割尺寸。

## 3.双向预测帧(B)

视频及其应用





### 3.双向预测帧(B)

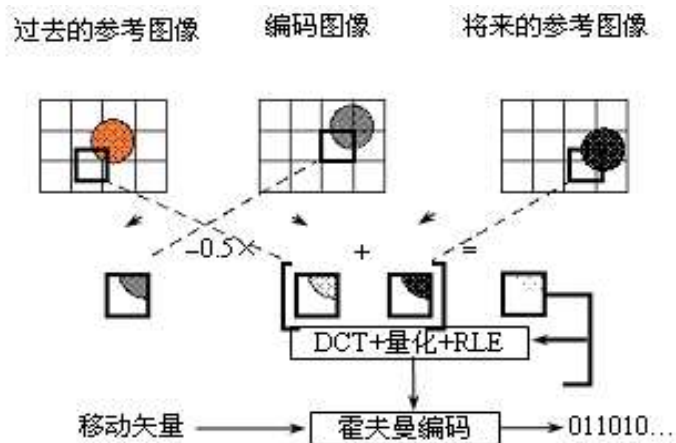
视频及其应用

#### ■ 定义

- 双向预测帧又叫内插帧
- 需要前后两个参照帧：
  - 采用双向预测编码和运动补偿法来进行图像编码
  - 本身不能作为参照帧，只是I和P之间的内插帧

#### ■ 压缩比最高。

- 内帧（I帧）是基础
- 定时提供一帧画面，称为刷新



双向预测图像 B 的压缩编码算法





## 视频及其应用

- 帧内图像的I的数据量最大，双向预测图像B的数据量最小。

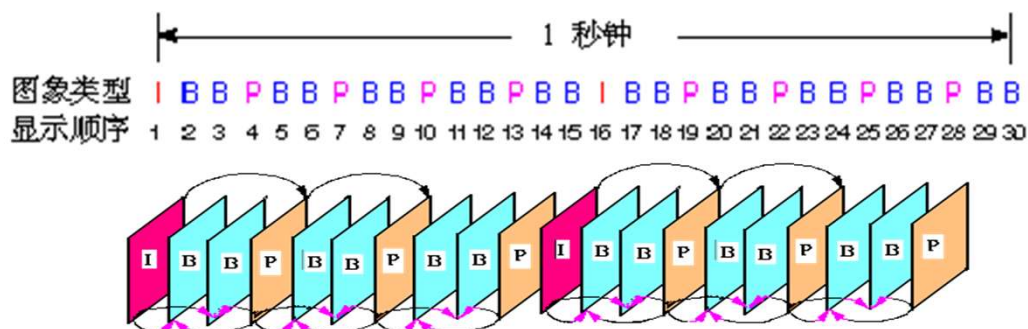
## 关键帧 vs. 参考帧

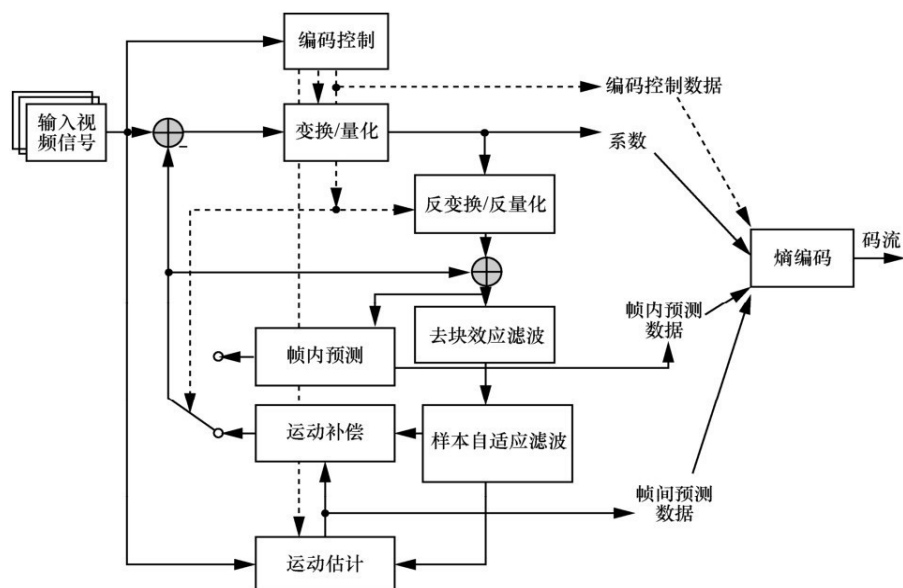
帧类型	是否可以独立解码	依赖谁解码	大吗？
I 帧	√	NA	大
P 帧	X	前	小
B 帧	X	前与后	小

图像类型	I	P	B	平均值/帧
MPEG-1 IF格式 (1.15Mb/s)	150 000	50 000	20 000	38 000
MPEG-2601格式 (4.00Mb/s)	400 000	200 000	80 000	130 000

## 视频及其应用

- 算法允许选择内帧I的频率及位置。
- 允许选择在一对内帧或预测帧P之间插入双向预测帧B的数目。
- 内帧I是基础，定时提供一帧I画面，称为刷新。





## 传统插帧方法优劣势一览

现有的较为传统的视频插帧技术。



## 帧采样

视频及其应用

帧采样示意



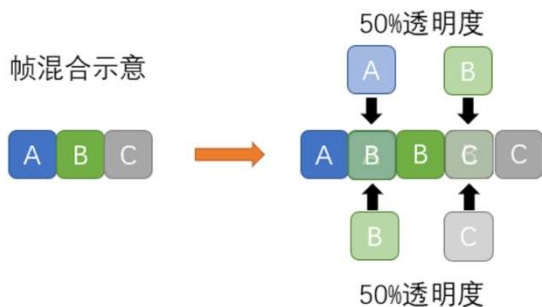
### • 帧采样

- 帧采样就是用关键帧来做补偿帧，其实质就是拉长每一个关键帧的显示时间，相当于并没有插帧。除了获得文件属性更高的帧率和在同视频质量下更大的文件体积外，不会带来任何视觉观感上的提升。
- 优点：帧采样消耗资源少，速度快。
- 缺点：可能会使得视频看起来不是很流畅。

## 帧混合

视频及其应用

帧混合示意



### • 帧混合

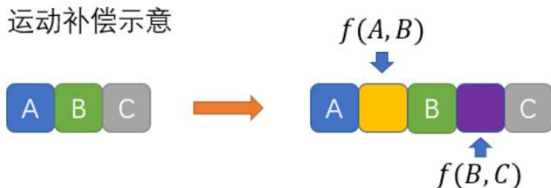
- 帧混合，顾名思义，就是提升前后关键帧的透明度，再将其混合成一个新的帧，来填补空缺。
- 优点：计算所需时长短。
- 缺点：效果不佳。由于只是简单的把原关键帧变成半透明状，运动物体轮廓在前后两帧交叠的时候，就会产生明显的模糊场景，对视频的视觉效果流畅程度提升很小。



## 运动补偿

视频及其应用

运动补偿示意



### • 运动补偿

- 运动补偿 (Motion Estimation and Motion Compensation, 简称 MEMC), 其原理是在水平和垂直两个方向上, 对两帧之间的差异寻找出现运动的块, 通过分析图像块的运动趋势, 再辅以计算得到中间帧。
- MEMC 主要被应用于电视、显示器和移动端, 提升视频帧率, 给观众更加流畅的观感。
- 优点: 减小运动抖动, 减弱画面拖尾与虚影, 提升画面清晰度。
- 缺点: 对于运动物体背景比较复杂的情况, 就会出现物体边缘运动的 bug

## 光流法

视频及其应用

光流法示意



### • 光流法

- 光流法是计算机视觉研究中的一个重要方向, 其根据上下帧来推断像素移动的轨迹, 自动生成新的空缺帧。有点类似于运动模糊计算方法。
- 优点: 画面更流畅、卡顿感弱。
- 缺点: 计算量大, 耗时长; 对光线敏感, 在光线变化较大的情况下, 容易出现画面紊乱的错误。

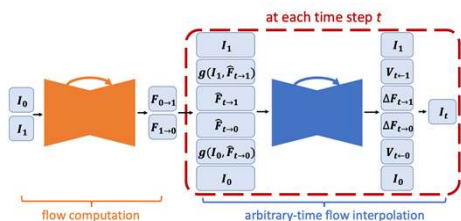


## Super SloMo: AI 插帧法

视频及其应用

### Super SloMo: High Quality Estimation of Multiple Intermediate Frames for Video Interpolation

Huaizu Jiang<sup>1</sup> Deqing Sun<sup>2</sup> Varun Jampani<sup>2</sup>  
Ming-Hsuan Yang<sup>3,2</sup> Erik Learned-Miller<sup>1</sup> Jan Kautz<sup>2</sup>  
<sup>1</sup>UMass Amherst <sup>2</sup>NVIDIA <sup>3</sup>UC Merced  
{hzjiang,eim}@cs.umass.edu, {deqings,vjampani,jkautz}@nvidia.com, mhyang@ucmerced.edu

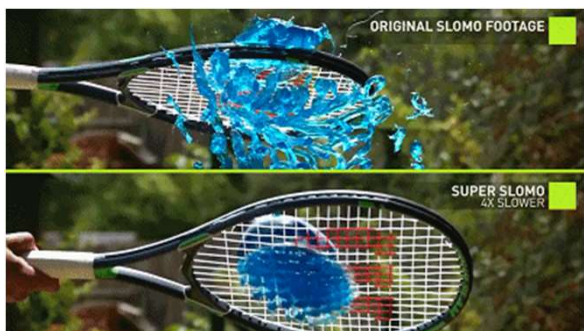


#### • Super SloMo: AI 插帧法，堪称业界经典

– Super SloMo 不同于传统方法，它利用深度神经网络来实现补帧，基本思路为：使用大量普通视频与慢动作视频进行训练，然后让神经网络学会推理，根据正常视频生成高质量的超级慢动作视频。

## Super SloMo 超级慢动作镜头补帧

视频及其应用



原始 SloMo 视频 (上)  
Super SloMo 补帧后的超级慢动作视频 (下)

#### • Super SloMo 超级慢动作镜头补帧

- 运行环境: PyTorch 0.4.1
- 语言版本: Python 3.6
- 训练可视化: TensorboardX
- 训练数据集: Adobe 240 fps
- 项目地址:
- <https://github.com/avinashpaliwal/Super-SloMo>



Think

慢视频如何制作?



## 视频帧率

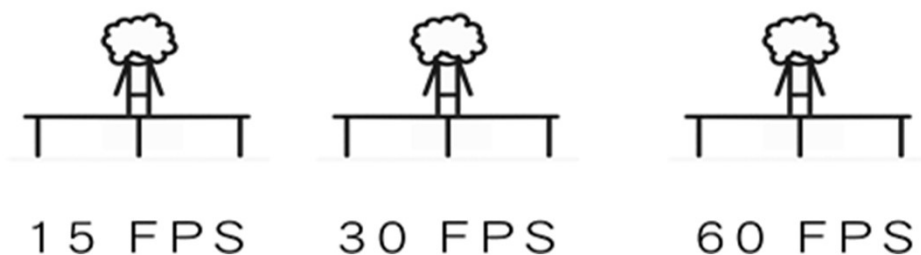
视频帧率（简称P）指的视频每秒钟播放的图片数目



# 视频帧率

(简称P)

- 数字电影或者视频是由一序列静止画面组成的，这些静止的画面称为帧。
- 一帧帧的图像组成视频，而视频帧率(简称P) 指的视频每秒钟播放的图片数目。
  - 比如25P就是指一秒钟有25张静止的图像；



我国采用的是电视标准PAL制，它规定视频每秒25帧；像电影的基本帧率为24帧每秒，帧率越大，视频画面越流畅，速度也就越快；而如果帧率低于15帧/秒，连续的运动视频就会有停顿的感觉。



## 视频帧率与快门速度

- 快门速度指每张图片曝光时间的长短，也就是每一帧快门保持打开的持续时间。
- 如果你选择拍摄60P帧率的视频时，快门速度选择30分之一秒的话，拍出来每秒最多得到30张单独的图像，另外30张是根据前面30张复制过来的。



拍摄时要注意的，你的快门速度一定要比帧率高。一般快门速度都是设定在比帧率倒数快一倍的档位上，如25P帧率使用1/50秒快门。





# 视频帧率与快门速度

播放的速率与拍摄速率的关系



## 视频处理的一些应用

简略版



## 视频清晰增强

重庆大学本科生毕业设计

H.264 错误恢复(ERC)算法研究与实现



(a)无损解码图像(b)做EC处理前的图像 (c)传统算法处理结果

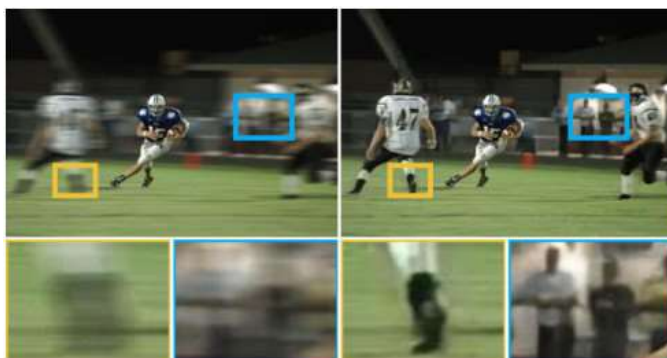
(d)RP算法处理结果



## 论文

视频及其应用

- Dynamic Scene Deblurring with Parameter Selective Sharing and Nested Skip Connections (基于选择性参数共享和嵌套跳跃连接的图像去模糊算法)



## 视频直播技术





伪实时

准实时

真实时

> 3 秒

单向观看实时  
CDN + RTMP + HLS,  
现在基本上所有的直播都  
是这类技术

1 ~ 3 秒

进行双方互动但互动有障碍

< 1秒

QQ、微信、Skype 和  
WebRTC 等都已经实现了

视频+ “交互”



## 视频分割

如何在视频里任意抠图？

## 视频物体分割

Video Object Segmentation  
简称 VOS

顾名思义就是从视频所有图像中把感兴趣的物体区域完整地分割出来。





# 视频产生新的应用入口

未来传递信息的主要载体是视频，而视频会产生新的应用入口。

## “视频+交互”成网络视频新机遇

- 交互式视频已经有哪些成熟的新玩法。
- 弹幕只是交互式视频最为简单的一种，就已经火爆到这种地步。





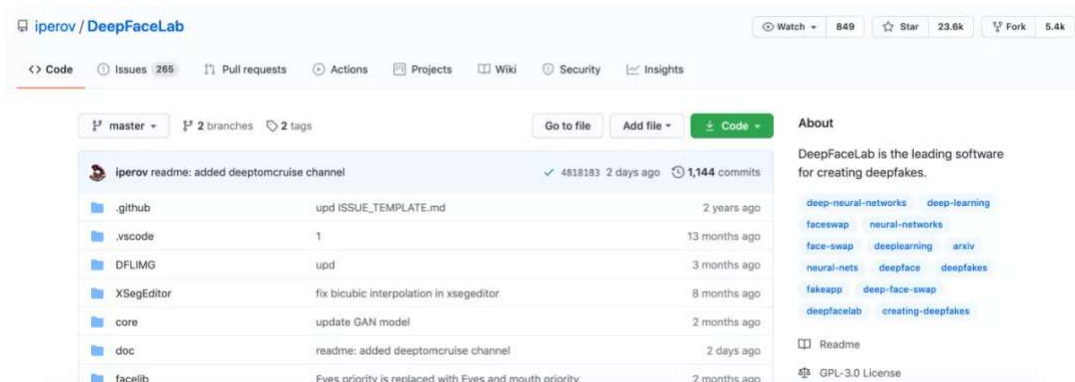
## 深度造假 (deepfaking)





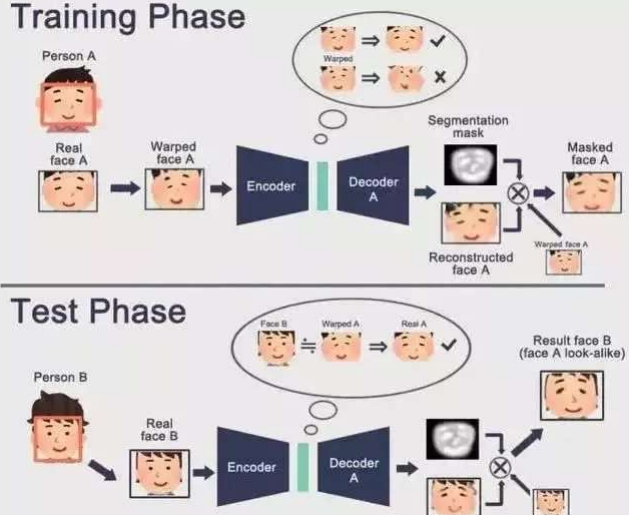
## DeepFaceLab

视频及其应用

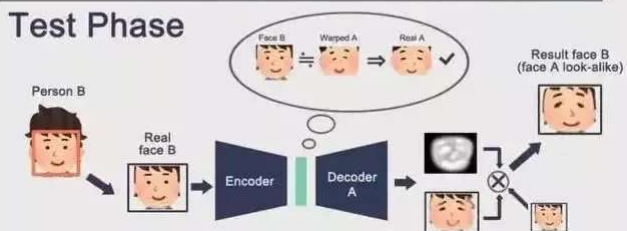


— 项目地址: <https://github.com/iperov/DeepFaceLab/>

### Training Phase



### Test Phase



直到“打假”模型区分不出真假脸的时候，就说明换脸成功了！





## Stable Diffusion能生成视频了



## 视频版Midjourney免费开放，一句话拍大片！

- 例如只需简单输入“丛林 (Jungle)”，大片镜头便可立刻呈现：
  - 而且围绕着“丛林”变换几个搭配的词语，比如“河流”、“瀑布”、“黄昏”、“白天”等，这个AI也能秒懂你的意思。





## 国外Sora还没来，国内“网课”已赚翻？

视频及其应用



- 最近，圈内最火的话题，就是OpenAI公布的文生视频工具Sora。仅靠文字就能自动生成长达60秒的高保真视频，给业界带来了一些技术震撼。
- 然而，这款产品都还没正式开放，国内不少嗅觉灵敏的人，已经开始卖课割韭菜了。





# 视频与三维数据

## SLAM

视频及其应用

- SLAM(simultaneous localization and mapping)全称即时定位与地图构建或并发建图与定位，它主要的作用就是让机器人在未知的环境中，完成定位 (Localization)，建图(Mapping)和路径规划(Navigation)。
- 由于其重要的理论与应用价值，被很多学者认为是实现真正全自主移动机器人的关键。
- SLAM技术被广泛运用于机器人、无人机、无人驾驶、AR、VR等领域。



# 追求视觉智能

追求视觉智能：对超越目标识别的探索

## 理解视觉对象关系

- ImageNet 后的目标识别路在何方：理解视觉对象关系。
- ImageNet的数据集和不同的衡量基准是对物体识别的发展有重要贡献的数据集之一。在 2010~2017 年 ImageNet 挑战赛中，目标识别错误率从 28% 降到了 2.3%



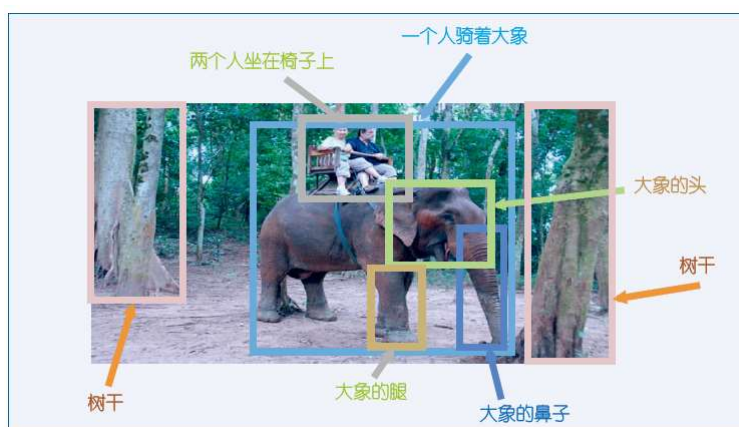
视觉对象所表现的语义



## “看图作文”

### 进一步理解视觉信息：用推理来完善描述

- 根据图像形式文字说明。



场景生成示例