

Object Detection

—Something about it

主讲人 赵海伟
幻灯片制作 赵海伟

中国海洋大学 信息科学与工程学院

September 26, 2014

content

- 1 What is Object Detection?
- 2 Why we research Object Detection?
- 3 What relates to Object Detection?
- 4 What makes up Object Detection?
 - 特征点检测 (Point Detector)
 - 背景建模 (Background Modeling)
 - 图像分割 (Segmentation)
 - 监督学习 (Supervised Classifiers)
- 5 conclusion
- 6 运动目标检测
- 7 objectness proposal

What is Object Detection?

- 是基于 {目标几何} 和 {统计特征} 的 {图像分割}。
- 将目标的 {分割} 和 {识别} 合二为一。
- {准确性} 和 {实时性} 是整个系统的一项重要能力。
- 尤其是在复杂场景中，需要对多个目标进行实时处理时，目标自动提取和识别就显得特别重要。

What is Object Detection?

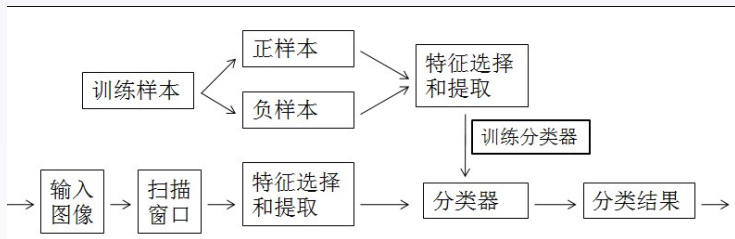


Figure: Object Detection

Why we research Object Detection?

- 计算机技术的发展和计算机视觉原理的广泛应用。
- 利用计算机图像处理技术对目标进行实时跟踪研究越来越热门。
- 对目标进行动态实时跟踪定位在智能化交通系统、智能监控系统、军事目标检测及医学导航手术中手术器械定位等方面具有广泛的应用价值。

What relates to Object Detection?

- 1.Segmentation
- 2.Object detection and recognition
- 3.Salient object detection

What makes up Object Detection?

- 特征点检测 (Point Detector)
- 背景建模 (Background Modeling)
- 图像分割 (Segmentation)
- 监督学习 (Supervised Classifiers)

特征点检测 (Point Detector)



Figure: Point Detector

特征点是在各自位置具有代表性纹理特征的点，是图像很重要的特征，对图像图形的理解和分析有很重要的作用。

特征点检测 (Point Detector)

- 点特征的优势：
 - (1) 点特征属于局部特征，对遮挡有一定鲁棒性。
 - (2) 通常图像中可以检测到成百上千的点特征，以量取胜。
 - (3) 点特征有较好的辨识度，不同物体上的点容易区分。
 - (4) 点特征提取通常速度很快。
 - (5) 对光照和摄像头视角变换具有不变性。

特征点检测 (Point Detector)

- 特征检测子(关键点)与描述子
 - FAST
 - MSER
 - StarDetector
 - StarDetector::StarDetector
 - StarDetector::operator()
 - SIFT
 - SURF
 - SURF::SURF
 - SURF::operator()
 - ORB
 - RandomizedTree
 - RandomizedTree::train
 - RandomizedTree::read
 - RandomizedTree::write
 - RandomizedTree::applyQuantization
 - RTreeNode
 - RTreeClassifier
 - RTreeClassifier::train
 - RTreeClassifier::getSignature
 - RTreeClassifier::getSparseSignature
 - RTreeClassifier::countNonZeroElements
 - RTreeClassifier::read
 - RTreeClassifier::write
 - RTreeClassifier::setQuantization

特征点检测 (Point Detector)

- 比较牛逼的几个点特征:
- 1.SIFT
- 2.SURF
- 3.Harris

背景建模 (Background Modeling)



(a) Current frames (b) Background images (c) Difference images

Figure: Background Modeling

背景建模 (Background Modeling)

- 背景建模，然后进行背景减除，剩下前景视作所求的目标，也是目标检测的一类方法。
- 背景模型的巨大变化即意味着目标移动。
- 背景建模其实就是前景检测。

背景建模 (Background Modeling)

- 比较牛逼的背景建模：
- VIBE (Visual Background Extractor) (牛逼)
- 简介：ViBe 是一种像素级视频背景建模或前景检测的算法，2009 年提出来的用于进行快速 {背景提取} 和 {运动目标检测} 的算法，具有较高的实时性和鲁棒性。

图像分割 (Segmentation)



Figure: Segmentation

图像分割就是把图像分成若干个特定的、具有独特性质的区域并提出感兴趣目标的技术和过程, 是由 {图像处理} 到 {图像分析} 的关键步骤。

图像分割 (Segmentation)

- 现有的图像分割方法主要分以下几类：
- (1) 基于阈值的分割方法
- (2) 基于区域的分割方法
- (3) 基于边缘的分割方法
- (4) 基于特定理论的分割方法

监督学习 (Supervised Classifiers)

- 监督学习 (Supervised Classifiers)

监督学习 (Supervised Classifiers)

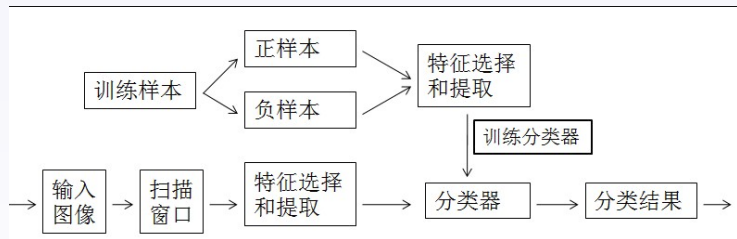


Figure: Object Detection

监督学习 (Supervised Classifiers)

- 目标检测中的监督学习方法，指的是在样本集中通过对不同视角下的目标的训练过程，学习得到不同目标视角下 {从输入到输出的映射函数}。
- 它是一个 {分类问题}，在目标检测中，学习样本由 {目标特征} 对一个相关的目标类别组成。

监督学习 (Supervised Classifiers)

- 一、特征选择
- 二、学习算法 (训练分类器)

一、特征选择

- 一、特征选择
- 特征选择是分类问题中的一个重要方面。
- 图像特征最重要的属性是 {独特性}, 能够在特征空间内方便区分目标。
- 特征可以是 {颜色}、{纹理}、{形状}、{轮廓} 等常用特征。
- 也可以是 {目标区域}、朝向、外观、{概率密度}、{直方图} 等。
- 可用于 {跟踪} 的特征有颜色、边缘、光流和纹理, 或者是其中几种的组合。

一、特征选择

- 1. 颜色特征
- 2. 纹理特征
- 3. 形状特征
- 4. 空间关系特征
- 5. 概率密度特征、直方图特征、梯度（变化率）特征等。
- 6. 边缘特征
- 7. 光流特征

一、特征选择

- 1. 颜色特征
 - 1.1 颜色直方图
 - 1.2 颜色集
 - 1.3 颜色矩
 - 1.4 颜色聚合向量
 - 1.5 颜色相关图

1. 颜色特征

- 1.1 颜色直方图
- 颜色直方图用以反映图像颜色的组成分布，即各种颜色出现的概率。
- (1) RGB
- (2) HSV
- (3) HLS

1. 颜色特征

- 缺点:
- (1) 由于颜色直方图是全局颜色统计的结果, 因此丢失了像素点间的位置特征。
- (2) 可能有几幅图像具有相同或相近的颜色直方图, 但其图像像素位置分布完全不同。
- (3) 因此, 图像与颜色直方图得多对一关系使得颜色直方图在识别 {前景物体} 上不能获得很好的效果。

1. 颜色特征

- 1.2 颜色集
- 颜色集的方法致力于实现基于颜色实现对大规模图像的检索。
- 颜色集的方法首先将颜色转化到 HSV 颜色空间。
- 将图像根据其颜色信息进行图像分割成若干 {region}, 并将颜色分为多个 {bin}, 每个 region 进行颜色空间量化建立 {颜色索引}, 进而建立二进制图像颜色索引表。
- 为加快查找速度, 还可以构造 {二分查找树} 进行特征检索。

1. 颜色特征

- 1.3 颜色矩
- 颜色矩是一种有效的颜色特征，该方法利用线性代数中矩的概念，将图像中的颜色分布 {用其矩表示}。
- 利用颜色一阶矩 { (平均值 **Average**) }、颜色二阶矩 { (方差 **Variance**) } 和颜色三阶矩 { (偏斜度 **Skewness**) } 来描述颜色分布。

1. 颜色特征

- 1.4 颜色聚合向量
- 这个是对颜色直方图的一种比较聪明的改进。
- 同样将颜色分成几个区间，不同的是在这不同的区间中，把每种颜色分成 {连接区域 (coherent)} 和 {孤立 (incoherent) 区域}，分别统计这些区域的像素数，做对比时也是分别对比。
- 这样，对直方图分类可以有改善，同样，根据这种思想也可以有很多的引申。

1. 颜色特征

- 1.5 颜色相关图
- 其含义可简述为：与颜色值为 i 的像素距离 k 的像素颜色值为 j 的 {概率}。
- 这种特征不但描述了某一种颜色的像素数量占整个图像的 {比例}。
- 还反映了不同颜色对之间的 {空间相关性}。
- 实验表明，颜色相关图比颜色直方图和颜色聚合向量具有 {更高的检索效率}，特别是查询空间关系一致的图像。

一、特征选择

- 2. 纹理特征
- 纹理是检测物体表面 {intensity} 变化的手段，是对光滑程度和规则程度的量化。
- 相对于颜色，它多了一个计算描述符的步骤，对于 {光照不敏感}。
- 一幅图像的纹理是在图像计算中经过 {量化} 的图像特征。
- 图像纹理描述图像或其中小块区域的空间颜色分布和光强分布。

2. 纹理特征

- (一) 特点:
- 全局特征。
- 仅仅利用纹理特征是无法获得高层次图像内容的。
- 需要在包含多个像素点的区域中进行统计计算。
- 不会由于局部的偏差而无法匹配成功。
- 具有 {旋转不变性}, 并且对于噪声有较强的 {抵抗} 能力。

2. 纹理特征

- 纹理特征缺点:
- (1) 当图像的 {分辨率变化} 的时候, 所计算出来的纹理可能会有较大偏差。
- (2) 由于有可能受到光照、反射情况的影响, 从 2-D 图像中反映出来的纹理不一定是 3-D 物体表面 {真实的纹理}。
- (3) 在检索具有粗细、疏密等方面较大差别的纹理图像时, 利用纹理特征是一种有效的方法。
- 但当纹理之间的粗细、疏密等易于分辨的信息之间 {相差不大} 的时候, 通常的纹理特征很难准确地反映出人的视觉感觉不同的纹理之间的差别。

2. 纹理特征

- (二) 常用的特征提取与匹配方法
 - 1. 统计方法
 - 2. 几何法
 - 3. 模型法
 - 4. 信号处理法

2. 纹理特征

- 1. 统计方法
 - (1) 灰度共生矩阵
 - 灰度共生矩阵的四个关键特征：能量、惯量、熵和相关性。
 - 灰度共生矩阵定义为像素对的 {联合概率分布}，是一个对称矩阵。
 - 它不仅反映图像灰度在相邻的方向、相邻间隔、变化幅度的综合信息，但也反映了相同的灰度级像素之间的位置分布特征，是计算纹理特征的基础。
 - (2) 从图像的自相关函数（即图像的能量谱函数）提取纹理特征（即通过对图像的能量谱函数的计算，提取纹理的粗细度及方向性等特征参数）

2. 纹理特征

- 2. 几何法
- 是建立在纹理基元（基本的纹理元素）理论上的一种纹理特征分析方法。
- 纹理基元理论认为，复杂的纹理可以由若干简单的纹理基元以一定的有规律的形式重复排列构成。
- 比较有影响的算法有两种：(1) Voronoi 棋盘格特征法 (2) 结构法。

2. 纹理特征

- 3. 模型法
- 以图像的构造模型为基础，采用模型的参数作为纹理特征。
- 典型方法：(1) 马尔可夫随机场模型法，(2) Gibbs 随机场模型法。

2. 纹理特征

- 4. 信号处理法
- 纹理特征的提取与匹配主要有：
 - (1) {灰度共生矩阵} 特征提取与匹配主要依赖于能量、惯量、熵和相关性四个参数。
 - (2) {Tamura 纹理} 特征基于人类对纹理的 {视觉感知心理学} 研究，提出 6 种属性，即：粗糙度、对比度、方向度、线像度、规整度和粗略度。
 - (3) {自回归纹理模型} 是马尔可夫随机场模型的一种应用实例。
 - (4) {小波变换}。

一、特征选择

- 3. 形状特征
- (一) 特点：可以有效地利用图像中感兴趣的目标来进行检索。
- 一些共同的问题：
 - (1) 缺乏比较完善的 {数学模型}。
 - (2) 目标有 {变形} 时检索结果往往不太可靠。
 - (3) 许多形状特征仅描述了目标局部的性质，要全面描述目标常对 {计算时间} 和 {存储量} 有较高的要求。
 - (4) 许多形状特征所反映的目标形状信息与人的直观感觉不完全一致 (特征空间的相似性与人视觉系统感受到的相似性有差别)。
 - (5) 从 2-D 图像中表现的 3-D 物体实际上只是物体在空间某一平面的投影，从 2-D 图像中反映出来的形状常不是 3-D 物体真实的形状，由于视点的变化，可能会产生各种 {失真}。

3. 形状特征

- (二) 常用的特征提取与匹配方法
- 1. 几种典型的形状特征描述方法
- 通常情况下, 形状特征有两类表示方法, 一类是 {轮廓} 特征, 另一类是 {区域} 特征。
- 图像的轮廓特征主要针对物体的 {外边界}, 而图像的区域特征则关系到 {整个形状区域}。
- (1) 边界特征法
- (2) 傅里叶形状描述符法
- (3) 几何参数法
- (4) 形状不变矩法
- (5) 其它方法
- 2. 基于小波和相对矩的形状特征提取与匹配

3. 形状特征

- 1. 边界特征法
- 该方法通过对边界特征的描述来获取图像的形状参数。
- 两个经典方法：
 - (1) Hough 变换检测平行直线方法：利用图像全局特性而将 {边缘像素连接} 起来组成区域封闭边界，其基本思想是 {点一线的对偶性}；
 - (2) 边界方向直方图方法：首先微分图像求得图像边缘，然后做出关于边缘大小和方向的直方图，通常的方法是构造 {图像灰度梯度方向矩阵}。

3. 形状特征

- 2. 傅里叶形状描述符法
- 基本思想是用物体边界的傅里叶变换作为形状描述，利用区域边界的封闭性和周期性，将二维问题转化为一维问题。
- 由边界点导出三种形状表达，分别是 (1) 曲率函数、(2) 质心距离、(3) 复坐标函数。

3. 形状特征

- 3. 几何参数法
- 形状的表达和匹配采用更为简单的区域特征描述方法，例如采用 {有关形状定量测度} (如矩、面积、周长等) 的形状参数法 (shape factor)。
- QBIC (基于内容的图像检索系统) 系统中，便是利用圆度、偏心率、主轴方向和代数不变矩等几何参数，进行基于形状特征的图像检索。
- 形状参数的提取，必须以 {图像处理} 及 {图像分割} 为前提，参数的准确性必然受到分割效果的影响，对分割效果很差的图像，形状参数甚至无法提取。

3. 形状特征

- 4. 形状不变矩法
- 利用目标所占区域的 {矩} 作为形状描述参数。

3. 形状特征

- 5. 其它方法:
- (1) 有限元法。
- (2) 旋转函数。
- (3) 小波描述符。

3. 形状特征

- 2. 基于小波和相对矩的形状特征提取与匹配:
- 先用小波变换模极大值得到 {多尺度边缘图像}。
- 计算每一尺度的 7 个不变矩, 再转化为 10 个相对矩。
- 然后将所有尺度上的相对矩作为 {图像特征向量}, 从而统一了区域和封闭、不封闭结构。

一、特征选择

- 4. 空间关系特征
- (一) 特点:
 - 所谓空间关系, 是指图像中分割出来的多个目标之间的相互的空间位置或相对方向关系, 这些关系也可分为连接/邻接关系、交叠/重叠关系和包含/包容关系等。
 - 相对空间位置信息 & 绝对空间位置信息。
 - 空间关系特征的使用可加强对图像内容的描述区分能力, 但空间关系特征常 {对图像或目标的旋转、反转、尺度变化等比较敏感}。
 - 另外, 实际应用中, 仅仅利用空间信息往往是不够的, 不能有效准确地表达场景信息。为了检索, 除使用空间关系特征外, 还 {需要其它特征来配合}。

4. 空间关系特征

- (二) 常用的提取图像空间关系特征方法:
- (1) 首先对图像进行自动分割, 划分出图像中所包含的对象或颜色区域, 然后根据这些区域提取图像特征, 并建立索引。
- (2) 简单地将图像均匀地划分为若干规则子块, 然后对每个图像子块提取特征, 并建立索引。

一、特征选择

- 5. 概率密度特征、直方图特征、梯度（变化率）特征等。

一、特征选择

- 6. 边缘特征
- 边缘通常伴随着剧烈的 {intensity} 变化。
- 对比于颜色特征，边缘特征 {对光照变化不敏感}。

监督学习

- 1. 微分算子法
 - (1) Sobel 算子
 - (2) robert 算子
 - (3) prewitt 算子
- 2.Laplacian 算子
- 3.Canny 边缘检测法 ({最流行的边缘检测算法})

一、特征选择

- 7. 光流
- 光流特征是定义区域内每个像素变化的 {稠密位移向量域}。通过计算区域内的 {光照对比度变化} 得到。通常应用在 {基于运动的分割和跟踪} 上。
- (1) 基于梯度的方法
- (2) 基于匹配的方法
- (3) 基于能量的方法
- (4) 基于相位的方法
- (5) 神经动力学方法

一、特征选择

- 如果是视频文件的话，还有基于运动的特征提取算法。
- 以上各种算法都有自己的优缺点，适用场合也不尽相同，通常需要多个特征 {组合} 在一起才能得到比较好的结果。

监督学习

- 二、学习算法 (训练分类器)
- 选择特定特征之后, 采用合适的学习 {算法} 来训练 {分类器}。

二、学习算法 (训练分类器)

- 1. Boosting
- 2. SVM(支持向量机)
- 3. 神经网络
- 4. 决策树
- 5. 贝叶斯

二、学习算法 (训练分类器)

- 1.Boosting
- Adaptive Boosting
- 通过一些低精度的分类器 {组合迭代}{调整权重} 以找到高精度分类器的一种方法。

二、学习算法 (训练分类器)

- 2.SVM(支持向量机)
- 是一种 {二分类模型}。
- 其基本模型定义为特征空间上的间隔最大的线性分类器 (即学习支持向量机的策略便是 {间隔最大化})。
- 最终可转化为一个凸二次规划问题的求解。
- {LinearSVM} (程明明 BING)

二、学习算法 (训练分类器)

- 3. 神经网络
- 是一种模仿生物神经网络的结构和功能的数学模型。
- 是一种非线性统计数据建模工具。
- 常用来对输入和输出间的关系进行建模，或用来探索数据的模式。
- 其一个比较贴切的定义是：{人工神经网络是由人工建立的以有向图为拓扑结构的动态系统，它通过对连续或断续的输入作状态响应而进行信息处理}。

二、学习算法 (训练分类器)

- 4. 决策树
- 决策树是由一个 {决策图} 和 {可能的结果} 组成，用来创建到达目标的规划。
- 建立并用来辅助决策，是一种特殊的树结构。
- 是一个 {预测模型}，他代表的对象属性和对象值之间的一种映射关系。
- 树中每个 {结点} 表示某个对象，而每个 {分叉路径} 则代表某个可能的属性值。
- 数据挖掘中决策树是一种经常要用到的技术，可以用于分析数据，也可以用来预测。

二、学习算法 (训练分类器)

- 5. 贝叶斯
- 朴素贝叶斯是贝叶斯决策理论的一部分，贝叶斯理论的核心思想是{为数据选择高概率的类别}。
- 而朴素贝叶斯之所以“朴素”，是因为这个形式化过程只做最原始，最简单的假设。

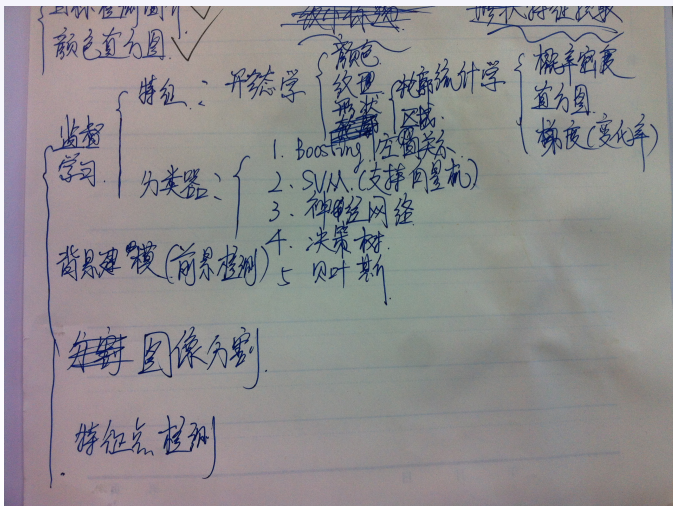
监督学习

- 部分热点:
- (1) BING (BING 是 NG 的二进制版本, 加快速度)
- (2) NG (梯度绝对值: 相邻像素颜色相减的绝对值)
- (3) DPM (HOG 的进化版)
- (4) HOG (方向梯度直方图) (轮廓特征明显)
- (5) LBP (局部二值模式) (纹理特征明显)
- LBP 局部纹理提取算子, 已经成功应用在指纹识别、字符识别、人脸识别、车牌识别等领域。
- (6) Haar-like (纹理特征明显)
Haar 特征分为三类: (1) 边缘特征 (2) 线性特征 (3) 中心特征和对角线特征。以上三类组合成特征模板。

监督学习

- 较为成功的目标检测算法:
- (1) 检测行人: HOG+SVM
- (2) 检测人脸: Haar+AdaBoost
- (3) 检测指纹: LBP+ AdaBoost

conclusion



运动目标检测

- 运动目标检测：
- 一、静态背景下的目标检测
- 二、运动背景下的目标检测

运动目标检测

- 静态背景下的目标检测：
 - 1. 背景差分法 (背景消除法)
 - 2. 多形态混合统计模型
 - 3. 帧间差分法
 - 4. 光流法
 - 5. 平均背景法
 - 6. 码本 (codebook)
 - 7. 边缘检测法
 - 8. 运动矢量检测法
 - 9. 将图像表达为离散事件状态
 - 10. 特征空间分解方法

运动目标检测

- 1. 背景差分法 (背景消除法)
 - (1) 中值法背景建模 (直方图法)
 - (2) 均值法背景建模
 - (3) 卡尔曼滤波器模型
 - (4) 单高斯分布模型
 - (5) 多高斯分布模型
 - (6) 高级背景模型

运动目标检测

- 2. 多形态混合统计模型
 - (1) 混合高斯模型 (GMM) (可归为背景差分法)
 - (2) 非参数核密度估计
 - (3) 纹理颜色特征融合
 - (4) 分层模型

运动目标检测

- 4. 光流法
- (1) 基于梯度的方法
- (2) 基于匹配的方法
- (3) 基于能量的方法
- (4) 基于相位的方法
- (5) 神经动力学方法

运动目标检测

- 7. 边缘检测法
 - 7.1 微分算子法
 - (1) Sobel 算子
 - (2) robert 算子
 - (3) prewitt 算子
 - 7.2 Laplacian 算子
 - 7.3 Canny 边缘检测法

运动目标检测

- 运动背景下的目标检测：
 - 1. 块匹配 (BMA)
 - 2. 光流估计
 - 3. 像素递归算法 (PRA)
 - 4. 自回归运动平均过程 (ARMA)

运动目标检测

- 1. 块匹配 (BMA)
 - (1) 穷尽搜索法 (Exhaustive Search) ES
 - (2) 三步法 (Three Step Search) TSS
 - (3) 新三步法 (New Three Step Search) NTSS
 - (4) 精简三步法 (Simple and Efficient Search) SES
 - (5) 四步法 (Four Step Search) FSS
 - (6) 菱形法 (Diamond Search) DS
 - (7) 自适应法 (Adaptive Rood Pattern Search) ARPS

objectness proposal

- 2010 CVPR:What is an object?journal version published in IEEE TPAMI 2012
- 2010 ECCV:Category Independent Object Proposals,journal version published in IEEE TPAMI 2014.
- 2011 CVPR:Proposal Generation for Object Detection using Cascaded Ranking SVMs
- 2011 ICCV:Segmentation as selective search for object recognition,journal version published in IJCV 2013.
- 2014 CVPR:BING: Binarized Normed Gradients for Objectness Estimation at 300fps

objectness proposal

That's all,thank you very much!