

Technology and Application Analysis of Indexing, Captioned and Keyword Searchable of ICS Lecture Video

Yingshan Shen, Yong Tang

The School of Computer, South China Normal University, Guangzhou
Email: esancn@gmail.com, ytang@scnu.edu.cn

Received: Oct. 8th, 2014; revised: Nov. 10th, 2014; accepted: Nov. 19th, 2014

Copyright © 2014 by authors and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Under the promotion and popularization of MOOCs (Massive Open Online Courses), lecture video as one of the main teaching medium, begins to catch peoples' attention. As a means of classroom and remote network instruction, curriculum lecture videos can record more versatile and multi-functional learning contents compared to a textbook. As the restrictions of the network speed and fast browsing, learners' acceptable level of lecture video is limited. In this paper, ICS lecture video was introduced with technology of video indexing, keywords search, online subtitle rendering and editing, which can improve and strengthen the users' experience. Through analysing and presenting the effective algorithm of the indexing of lecture video, one frame per sample interval was examined and the binary searching method was used. Image pretreatment and OCR tools were used to identify the option, keywords matched to search the required index of lecture video; option of lecture video is automatically recognition to text by a semi-automatic generate processing. ICS lecture video provides a more flexible way of teaching lessons to help learners as master roles in a more flexible time, in any place, to learn effectively. Also it provides a wisdom reference for the development and recording of lecture video in MOOCs.

Keywords

Index, Caption, Keywords, Lecture Video, Search

索引式可搜索的ICS课程教学视频 技术与应用分析

沈映珊, 汤庸

华南师范大学计算机学院, 广州
Email: esancn@gmail.com, ytang@scnu.edu.cn

收稿日期: 2014年10月8日; 修回日期: 2014年11月10日; 录用日期: 2014年11月19日

摘要

MOOCs (Massive Open Online Courses, 大规模网络开放课程)的大力推广与盛行下, 作为其中的主要教学方式的课程教学视频也开始备受关注。作为课堂教学和远程网络教学的一种手段, 课程教学视频能够记录下比课本更丰富、更多功能的在线学习内容。由于网络速度与快速浏览的限制, 学习者对课程教学视频的接受的程度有限。本文提出了具有视频索引、关键字搜索、在线字幕呈现与编辑的ICS (Indexed, Captioned and Searchable)教学视频, 改善与强化用户体验。通过分析并介绍课程教学视频中高效的索引算法、运用隔帧采样与二进制搜索方法分析教学视频, 利用对图片的预处理与OCR技术识别视频中的文字, 使用关键字匹配来搜索与识别所需教学视频; 并运用语音自动转换成文字, 进行半自动化的字幕校正与呈现。ICS课程教学视频的应用, 提供了更灵活的教学视频课程方式, 能帮助学生在更灵活的时间、任何地方, 以学生为主控角色进行更有效的学习。也为MOOCs中教学视频的发展与录制提供智慧参考。

关键词

索引, 字幕, 关键词, 课程视频, 搜索

1. 概述

由美国麻省理工倡导的大型开放式网上课程 MOOCs (Massive Open Online Courses)自推行以来, 虽然教师的参与程度在学习过程中降到最低, 但却因为基于互联网、免费、课程类型选择多, 学习方式灵活, 能在任何时间、任何地点、任何情境下进行学习, 吸引了大量学习爱好者[1]。学生通过观看简短的讲课视频, 然后完成作业, 由机器或其他学生批改。这样一来, 单一教师可以教授一节具有数千或数万名学生的课程。从 MOOCs 推行以来, 很多大学都开始开发并推行 MOOCs, 如 Coursera, edX, Udacity, iCourse 等, 学生多达上亿[2]。随着智能设备与互联技术的发展, MOOCs 不止能在 PC 上运行, 还可以通过智能移动设备, 如平板电脑、智能手机上运行。但大多教学视频具有这样的问题[3]:

- 1) 课程讲授水平与课堂讲授不同, 偏向平铺直叙, 学生无法获得与在课堂面对面学习时的情境;
- 2) 教学视频只能从头播放到最后, 无法快速访问感兴趣的内容, 即视频内容不可搜索。
- 3) 教学视频基本上都无法提供字幕, 大多数的教学视频都是英文, 给非英语国家, 如中国的学习者在听讲与理解过程造成了障碍。

笔者 2013 年底到美国 The University of Houston (UH)访学, 参与到该校 MOOCs 课程项目——索引式可搜索课程视频(ICS, Index and Caption course video for STEM)的开发与运行工作, 此课程视频的开发与运行正是针对当前 MOOCs 教学视频的缺点而设。

2. ICS 教学视频

在 UH 中, 教师在课堂上使用平板电脑(Tablet PC)作为教学设备来讲课。这种教学平板电脑可以在打开 PPT 的同时, 老师可以用手或鼠标在桌面上添加注释或插图进行板书, 老师不需要再转身在黑/白板上进行板书, 也不再需要通过其他板书软件在 PPT 与软件之间来回切换。与此同时, 该教学用平板电脑还具备同步录音与录像的功能, 教师完成一节课的讲授, 正好可以完成一节教学视频的录制。从这点上

来看,它避免了像一般 MOOCs 一样,需要教学在课后自己录制教学视频。从听讲课程的角度来看,学生更容易获得现场感。而 ICS 的教学视频项目的目标是旨在改善通过平板电脑所获取的教学视频流中,通过添加索引、字幕和搜索能力,增强 MOOCs 在全球运行的功能,吸引更多的非英语语种学生的参与。

ICS 教学视频的优点[4]在于:

1) 索引点:录制的教学视频讲座将具有一系列逻辑索引点,在每一个代表一个视频段的快照地方建立索引点列表;

2) 文本字幕:整个音频流,包括课堂上老师的讲课、教师与学生的互动话语,都会在一个单独的窗口上打出文本;

3) 关键字搜寻:包含该关键词的音频视频片段将被显示出来,并按关键字出现的频率排序;

如图 1 所示的是 ICS 的教学视频播放主界面,它是一个基于标准 H.264 编码的流媒体播放器,主要包括三个部分:

第一部分是主窗口,即教学视频的主窗口。学生通过视频主窗口观看教师的 PPT 和课堂演示与板书进行自主学习,可以在播放视频的过程中随时控制视频的播放与暂停,也可以通过“播放控制面板”控制视频播放的速度。在视频下方显示教师讲课的实时字幕。

第二部分是教学语音字幕索引窗口,显示这节课的所有语音字幕(教师或学生的),主要是通过教师说话的间断时间进行分段,可以精确到 HH:MM:SS。它也可以随时选择打开或关闭。通过这一窗口,学习者可以随时选择没听懂或没听清的部分,单击反复观看该部分视频。

第三部分是视频索引窗口,根据教学视频内容的变化进行时间分段。在它上方有一个关键字搜索窗口,通过在这个窗口键入相关的关键字,在视频索引窗口有含有该关键字的视频将被加亮,其下方搜索的结果和该关键字在这一段视频中出现的次数。图中所示是指视频中共有 4 段含有“file”一词的视频。学习者可以点击搜索结果,观看从这一时间点开始的包含有该搜索关键字的视频进行学习。这有利于学习者重复学习自己没能理解的内容,也可以进行加强性或重复性学习。

3. 关键技术分析

ICS 教学视频是一种将标准 MPG 格式的视频完全自动转换为 ICS 流格式的视频,将流视频、字幕、文本和搜索都统一集中在 ICS 的播放器窗口中。ICS 教学视频的转换技术框架如图 2 所示:

首先,将标准格式的视频通过自动索引算法进行视频帧的分析与应用识别,将所有相同内容的视频帧分隔为不同的段,并建立索引点指数。其次,将建立的视频索引指数保存到数据库中,同时提取所有图像的索引指数发送到搜索框。第三步,每个建立索引指数的视频段通过颜色反转技术进行图像变换,使得视频帧转变为一幅幅图像;利用搜索技术从每一幅图像中提取出文字图像,再通过 OCR(光学图像转换)技术将文字图像转换为文本,并以.txt 格式保存到数据库中。第四步,提取视频中所有的语音,通过语音转换技术,自动将语音转换为文字显示到文字字幕窗口,一方面通过视频字幕技术以手工方式在当前视频帧中添加字幕,其次将转换的文字保存到数据库中。当 ICS 视频播放器在通过搜索框搜索某段包含某关键字的视频时,搜索技术会在数据库查找到通过 ICR 转换的文本,通过文本所在的视频索引指数查找到相关的视频段,输送到视频播放窗口进行播放,同时,搜索技术也会搜索到语音文字所在的语句,发送到字幕窗口进行以高亮方式同步播放;而视频在播放过程,会直接提取加载在视频上的字幕。而当选择语音字幕上的文字来播放视频时,可以通过字幕上的时间节点在数据库上查找到相同的时间节点视频进行播放。

ICS 教学视频的转换与播放过程中,主要用到三个主要技术支持:1) 索引指数(Index);2) 字幕(Caption);3) 搜索(Search)。下面详细分析这三个主要技术。

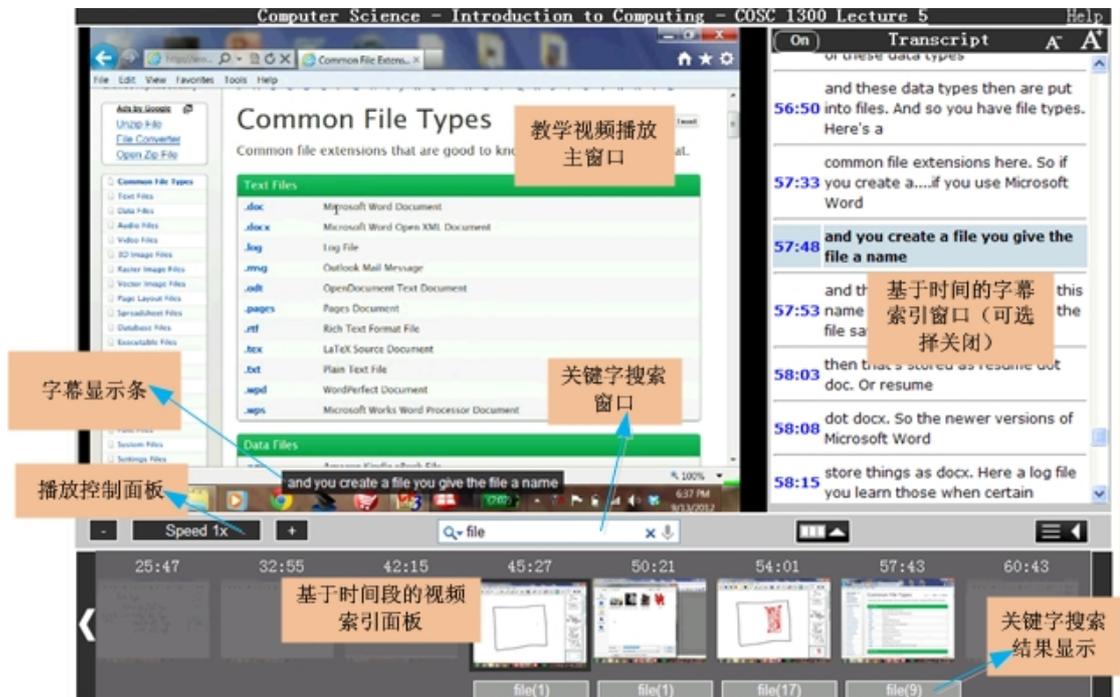
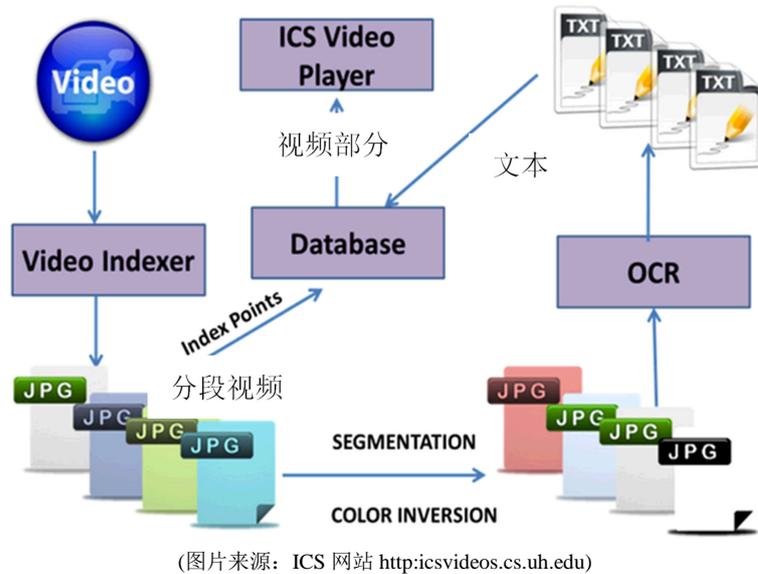


Figure 1. Interface of ICS
图 1. ICS 播放界面



(图片来源: ICS 网站 <http://icsvideos.cs.uh.edu>)

Figure 2. Main technology of ICS lecture video
图 2. ICS 教学视频主要技术

3.1. 索引指数(INDEX)

ICS 教学视频中的索引点的主要任务是将一个普通教学视频转换成包含不同主题的分类视频。所谓主题是指在一贯连续视频帧的某一帧之后场景发生了显著变化,就在这一变化帧与前面的过渡帧之间插入一个过渡点,这一过渡点可作为用户搜索某段视频的新起点。所有的过渡点的子集就是一个索引点。要确定一个过渡点,首先需要进行过渡点检测。最基本的方法是对连续帧中的视频像素信息进行比较,

如图 4-A 所示。比较两个相邻视频帧的像素 RGB 值，如果两帧之间所有像素的 RGB 值不一样，就认为两帧之间具有最小的 RGB 阈值 RGB_{mix} 。通过 RGB 阈值的比较，可以更好的适应两个相邻帧具有同一场景的状况(比如两帧都是相同的表格或图表，但表格或图表中的内容并不一样)，避免了伪过渡点的生成。只要是基于 RGB 标准的视频，当两帧间的 RGB 阈值大于 RGB_{min} ，那么两帧间的转换点或者说过渡点就可以确定。如图 3：第 2 帧与第 3 帧之间的 RGB 阈值达到最大，第 3 帧就是新的转折点/过渡点，由此可确定在第 3 帧前面的视频属于同一内容，建立索引指数。为使索引指数能区别开来，经过分析大量的教学视频转折点情况， RGB_{min} 取值 10%。

为了提高辨识视频转折点效率，项目用了两种算法：1) 顺序查询法；2) 二进制查询法。这两种方法的视频采样不再跟之前那样对每一帧采样，而是采用隔帧采样。即如果采样间隔取值为 2，就比较视频帧 1、3、5、……，如果采样间隔取值为 3，则取样视频帧 1、4、7、10……，如图 4B 或图 4C 中所示，灰色帧将不进行对比。

1) 顺序查询法

根据课堂现场教学视频录制的情况，所录制的教学视频的过渡并不频繁，也就是可能有很多视频帧展现的都是相同的内容，顺序查询搜索法既通过跳过相同内容帧的块的方法进行优化。不再是比较连续的视频帧，而是取样第一帧，然后按采样取值比较跳跃帧，如果两帧是相同的，则说明两帧之间的所有帧也是相同，即间隔帧可以被忽略，搜索标记向前移动。再取下一个跳跃帧进行比较，如图 4B 所示。如果发现当前帧与跳跃帧不同，则将搜索标记所在的那一帧作为当前帧，与跳跃帧再进行比较。如果两帧相同，再比较下一个跳跃视频。一直到出现第一帧与最后的跳跃帧不同，搜索标记处所在的帧与最后的跳跃帧也不同时，可以确定最后跳跃帧前就是视频的一个转折点，可以插入一个索引指数。

2) 二进制查询搜索法

二进制查询搜索法是将所有的采样视频分为平均两部分，将第一段的起点帧与结束帧进行比较。如果两帧相同，那么这一段中所有的帧都不包含过渡点。如果两帧不同，就将这一段再分为平均两段，再进行起点帧与结束帧之间的比较。运用这种递归方法，当某一段视频达到最小采样值时，比较帧仍是不一样，说明这是一个视频转折点，可以插入一个索引标记。搜索结束。如图 4C 所示。

3.2. 关键字搜索与 OCR 技术

关键词搜索是在所有的视频段中通过关键字段匹配来获取相应视频段的过程。这个过程如图 2 所示，当通过索引创建了包含转折点视频帧的视频段，视频段中的所有帧上的文字将被检测出来并存储于数据库中。当用户通过键入关键字段去搜索某段视频时，ICS 视频播放器就从数据库中搜索并加载相应视频段，并同时为用户在索引窗口呈现包含关键字段的一系列视频段。

在教学视频中，所有的文本图像都包含在每一帧中，而这些视频帧并不可能都是由纯文本构成，可能含有复杂的图像、或者具有不同的颜色、不同形状的背景，或者文字与图例在这一起。要将所有文本识别出来，项目采用了 OCR (光学字符识别)技术来作为识别的工具。网上有很多开源的 OCR 工具可供选择，如微软的 MODI (微软办公文档影像)、GOOCR (开源的 OCR 光学识别程序)和 Tesseract (HP 图像识别类库)，但这些工具主要是为了识别扫描文档黑白格式的文本[5]。基于识别准确性、易于集成性和可用性原则，项目确定了以 MODI 工具来作为主要的文字识别工具，并在识别之前采用简单图像增强技术对静止视频帧进行预先的图像处理。图像增强处理过程主要分为两步，即分割文本和反转颜色。分割文本是将图像中的文本检测并识别出来的过程，其过程如图 5 所示：

如图 5 所示，首先使用简单的图像像素阈值统计方法将彩色图像转化为二进制的黑白图像[6]，即图像阈值等于加权的像素总和除以他们的总量。然后通过扩张方法来保留图像上的需要留存下来的对象，



Figure 3. Main technology of ICS lecture video

图 3. 转折帧确定

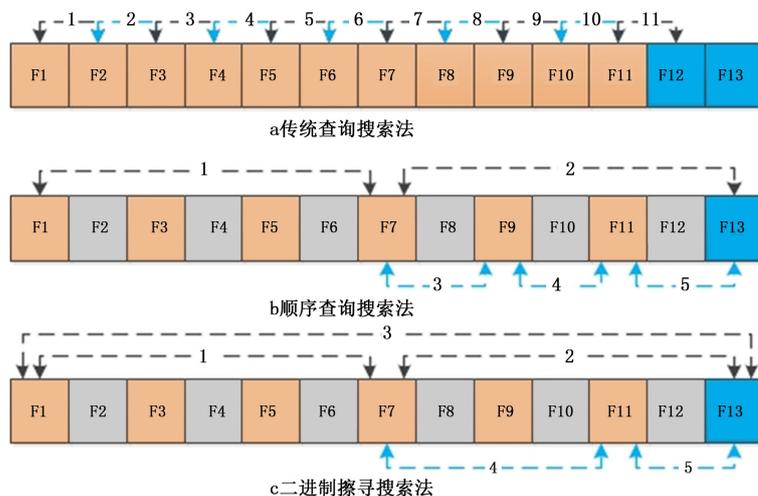


Figure 4. Compare schematic of video query searching method

图 4. 视频查询搜索法比较示意图

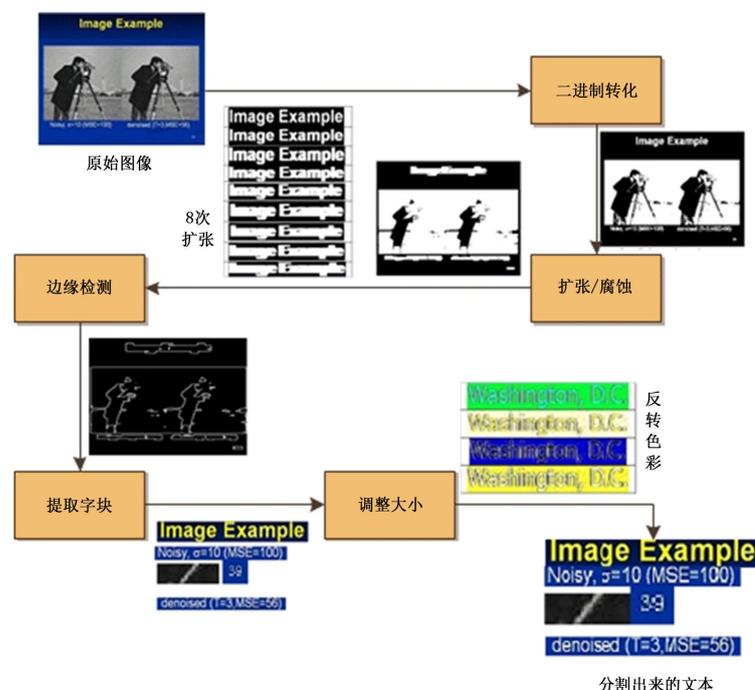


Figure 5. The processing of text segmented from images

图 5. 从图像中分割文本的过程

即字符。如果是一张图片具有黑色背景，白色前景，通过扩张，可以用白色填充字符间某些小尺寸的干扰点(漏点)，使单独的字符可以连成一片。但是当反过来用相同的方法去扩张一张白色背景、黑色前景的图片的时候，扩张方法就会将白色背景除去，这时应该用腐蚀法，用黑色去腐蚀漏点。为了避免这种情况，对每张图片，先进行平均光密度的计算，再决定是否用扩张还是用腐蚀处理。计算公式如下：

$$\text{AOD}(I) = \left(\frac{1}{N^2} \right) \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} I(i, j)$$

其中，AOD 表示表示二进制图像的灰度值，取值 $\text{AOD} = (0, 1)$ ， $\text{AOD} = 0$ 为白色， $\text{AOD} = 1$ 为黑色。当 $\text{AOD} > 0.15$ 时，指向一张黑色前景和白色背景的图片，用腐蚀方法处理；反之，当 $\text{AOD} \leq 0.15$ 时，指的是一张白色前景和黑色背景的图片，用扩张方法处理。经过试验证明，这种扩张或腐蚀的方法，需要经过 8 次迭代的过程才能得到适用的效果。

在扩张处理后，字符等小块已经可以被辨识出来，但由于这些小模块的边界存在着不完整的扩张或腐蚀，会影响分割的正确性。因此，需要对这些小块进行边缘检测，采用 Sobel 算法将不完整的边界补充完整并连接起来，保证小块边缘的统一与正确性。Sobel 算法是将所取得的小块图片中的每个像素取一个与原始图片像素点的接近的绝对梯度幅值[7]。其算法如下：

$$|G| = |G_x| + |G_y|$$

这里 G_x ， G_y 取值：

GxGy					
-1	0	+1	+1	+2	+1
-2	0	+2	0	0	0
-1	0	+1	-1	-2	-1

由 Sobel 算法可知某一像素 X 为：

P1	P2	P3
P8	x	P4
P7	P6	P5

因此： $|G| = |P1 + 2P2 + P3 - P7 - 2P6 - P5| + |P3 + 2P4 + P5 - P1 - 2P8 - P7|$ 。

接下来，对于进行了边缘检测的字符块进行块提取。由于字符块之间有可能存在其它非字符信息，因此，在提取之前首先对字符块进行过滤与筛选。过滤的原则是基于在之前的处理技术下，字符块是由斑点组成，而根据教学内容来说，图片上的字符长度一定比其宽度长，如果两个斑点之间的距离小于两个字符的长度，这些斑点将被过滤。或者说斑点连接起来的块长、宽的比例达不到字符块的常规比例，这些块也将被过滤。提取出来的文字可能太小或者太大，根据 OCR 识别字符的要求，字体的最佳大小为 10~12 个点阵[8]，需要对文本块进行调整，使之达到 OCR 识别的要求。为了增强 OCR 对文字的准确识别能力，可以对已校正大小的文本进行不同程度的色彩反转，增加文字与背景的对比度，从而可准确的检测与识别出文字。

3.3. 字幕(CAPTION)

ICS 教学视频与一般 MOOCs 课视频不同的地方之一，就是提供了字幕和对应视频索引的字幕索引窗口。这对于听力有障碍者、听力有问题的、或母语为非英语的学生来说，在听不到、听不懂或不能理解的时候，可以直接从字幕对照学习，提高了自主学习的效果[9]。

ICS 教学视频中的字幕是将教师讲课的声音，或者教师与学生在课堂上互动的语音通过工具转换为

文本并添加到字幕面板及视频上的。如果采用目前网络上流利的语音转换工具，如 IBM VOIVOICE 等工具，虽然转换准确率较高，却需要手工进行，转换慢，需要专门的录音设备，还需要事先对转换工具进行语音培训，不利于所有课堂视频的语音转换[10]。如果采用 ASR (Automatic Speech Recognition, 自动语音转换)转换工具，但转换的准确度却很低，即使经过培训，也只能达到 50%~80%的准确率，不利于学习[11]。为了解决语音自动转换、转换率高，准确度大的问题，项目组开发了一个基于 Web 的半自动的集成字幕编辑工具——ICS 字幕编辑工具(ICS Caption Editor)。所谓集成是因为字幕搜索窗口与视频索引窗口集成在同一界面中，能够自动运行语音转换成文本，从而提高了 ICS 教学视频的价值。所谓半自动化，是因为这个工具需要在语音转换成文本后，进行手动的校正。

ICS 字幕编辑工具通过内部 Google API 将要转换的具有编号的教学视频上传到 Youtube, Youtube 能够高速且准确的将语音转换成文本字幕，然后再将字幕打包传回编辑器的数据库。编辑器将所获得的字幕每 5 句分为一组，学生或其他专业人员可以通过点击播放视频校正每一组句子，其过程如图 6 所示。由于是基于 web 的，这种校正可以同步进行，提高了校正的与准确率。

4. 评价分析

ICS 教学视频项目从在休斯敦大学开展以来，已超过 10 专业的 48 门课程进行录制与播放，从早期面向全校师生都目前的面向全球，采用的是免费获取的方式。项目成员针对项目的各项指标进行了不同方面的测试与评价。这里对 ICS 视频播放仅从用户使用的角度进行评价。主要分为以下几方面：

- 视频播放窗口模式

目前大多数的 MOOCs 课程中，教学视频总是单独的一个播放界面，只有简单的控制按钮，可以控制播放速度，对于视频搜索只能从拖拉进度条来完成，很难做到精确定位。学生要想重复听讲某一个内容，只能依靠观看后的经验值来粗略估计播放位置。如图 7 所示。

而 ICS 教学视频播放，如图 1 所示，更强调其导航与索引搜索的能力。除了一般的视频播放功能外，还具有索引视频栏、关键词搜索窗口与字幕索引窗口。更方便学习者重复或针对性的快速获得所需学习的视频部分。通过搜索关键词，能将教学视频中包含有该关键字的字频部分都查找出来并高亮显示，学习者可以通过筛选很快确定自己要观看、学习的部分。对于学习内容来说，更容易获得重点学习的内容。相对一般的教学视频，可以在达到快速查找，重点加强学习的优点。

- 字幕展现形式

目前大多数 MOOCs 都有字幕显示与否的功能，只需要点击视频播放窗口上的字幕控制按钮，教师的讲课语音可以直接转换成字幕在屏幕下方显示。方便了第一语言非英语、或者听力有障碍的学习者直观的从字幕上获得学习内容，帮助到学习者的课程学习。通过让观看完 ICS 教学视频的学习者完成网上的调查问卷的数据显示，在参与调查的第一语言非英语的学习者中有 90%认为通过字幕可以很好的帮助他们理解教师讲授的内容。ICS 教学视频的字幕在这一点上与所有 MOOCs 没有什么差别。但是 ICS 教学视频播放窗口集成了按索引时间分配的所有字幕，对于教师正在讲解的一句，会以高亮方式呈现，方便了学习者知道教师正在讲哪一句，提前了解教师的下一句话是什么，有利于学习者从所有的教学字幕中查找要学习的某段视频。提供了另一种方式的教学视频搜索方法。另外，通过现场语音转换的字幕，准确性并不可能达到 100%。ICS 教学视频中内嵌的字幕编辑工具可以让学习者在确定教师的讲解内容时修改不正确的字幕，从而将 MOOCs 开发的一部分任务转移到学习者手上，这对于 MOOCs 开发从教授者集中的方法向学习者转移，让学习者能更好的掌握学习内容。特别是随着教学视频也可以让学习者参与到在线编辑的出现，真正做到主动学习，这将是云计算与大数据时代发展的趋势。

- 录制手段

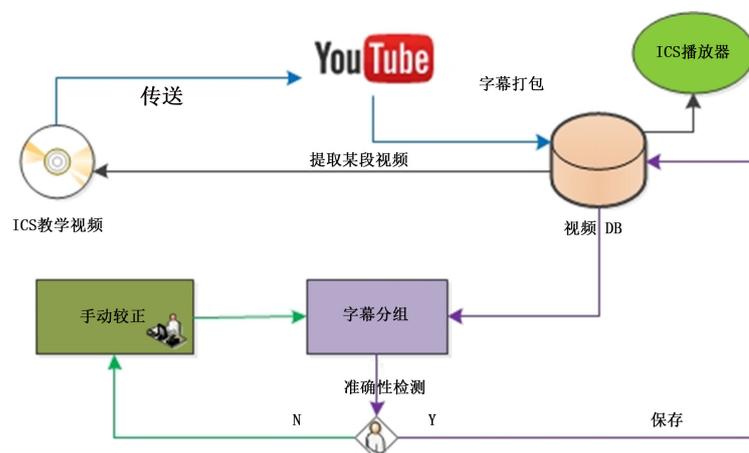


Figure 6. The working principle of ICS subtitle editor tool

图 6. ICS 字幕编辑工具工作原理



(来源 <https://class.coursera.org/insidetheinternet-005/lecture/182>)

Figure 7. Lecture video capture of the history, technology and security of Internet from Coursera

图 7. Coursera 上《互联网历史、技术与安全》教学视频截图

经过对网上现有 MOOCs 教学视频录制方法的调查发现，大多数教学视频采用外录式或内录式的方法(表 1)。但由于内录式的方法所需的设备简单、软件容易获取，对环境要求不高，教师可以在他人协助下或单独完成，而成为 MOOCs 课常用的录制方法。通常教师在准备好 PPT 和教学内容的情况下，使用平板式电脑、麦克风、录屏软件、手写笔等，在相对独立和安静的房间进行教学内容的讲解和录制，然后通过视频编辑软件对录制的教学视频进行编辑与整理，一般根据学习者关注度时间的要求整理成为 20 分钟以内的教学视频[17]。ICS 教学视频也是采用内录式，是教师在课堂环境下，向学生讲授教学内容的同时使用录屏软件进行录制的。因此除了教师的讲授声音外，还会有教师与学生进行交互的声音。这样做的好处是：教师不需另花时间去录制教学视频；学习者可以通过教学视频身临其境的感觉到教学的真实性；而教师与学生的交互过程可以让学习者了解某些教学的重点、难点。但缺点是一般课堂教学时间为 40 分钟，从观看者的角度上视频显得冗长，而且教师与学生的交互也并不是常常集中在教学重点、难点上。但是，由于有了后面视频编辑软件的支持，可以对已录制的视频进行编辑，改善教学视频的缺点。

- 视频教学内容展现

MOOCs 的最大差别是将教师的教学过程从课堂转换到网络，其教学过程不再是在课堂上，而扩展到

Table 1. The classification of recording method of lecture video

表 1. 教学视频录制方法分类

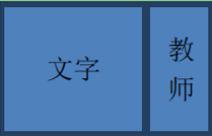
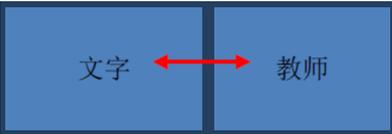
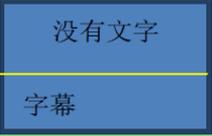
录制方式	授课视频的种类	定义	所需软硬件	优、缺点	引用资料或来源	
外录式	出镜讲解	授课者的形象出现在屏幕中来讲解知识点, 容易抓住学习者的注意力, 形成一对一的授课的感觉, 绝大多数课程都可以采用				
	手写讲课	通过板书推导、演算过程的授课方法, 具有板书讲解的全部优点; 减去拖沓内容, 提升讲解效率				
	实景授课	到实验室中边实验边讲解, 或都相应的场外讲解教学内容, 是对传统课堂的补充, 没有了教师空间的限制, 学习者能感受现场氛围的同时掌握知识。	需用到“录播室”。利用摄像机对教师的授课过程进行拍摄, 全程需要摄像人员配合教师完成拍摄。	可以更充分地展示教师的授课魅力, 但也需要较大的人力、物力及时间的投入。视频制作要求高, 还需要进行抠像、多机位拍摄等。时间投入跟产出比高。	[13] [14] [15]	
	动画演示	运用播放动画的同时形象讲解抽象知识, 生动活泼, 易于理解, 可以调动学习者的学习兴趣, 提升知识讲解的效率。				
	访谈式教学	在线课堂中师生访谈, 循序渐进地将知识寓于对话之中, 整个授课内容富有故事性, 能够吸引学习者的注意力, 也能让学习者接触更多人的真知灼见				
内录式 (屏录式)		利用计算机的录屏软件自动录制教师讲课使用的计算机屏幕, 比如 PPT, 同步录制画外音, 有时也会录制教师影像, 通常可以由教师自己完成录制。通过录屏的形式保存手	1) 桌面式工作站外接手绘板、麦克风、摄像头; 2) 带笔触输入的平板电脑外接麦克风、摄像头; 3) 直接使用带摄像头的笔记本电脑; 4) 屏幕录像软 Camtasia Studio、视频剪辑软件等	视频制作的设备要求简单, 教师自己就可以完成	[12] [16]	

教学视频中, 通过互联网可供不同地点、不同时间、不同语言、不同学习方式的人们自主学习。学习者的眼睛不再只单纯跟随教师, 而放眼到网上的教学视频。因此, 教师与教学内容如何在屏幕上分布与呈现就非常重要, 它涉及到如何在短时间使学习者能集中注意力, 听从视频中教师的讲解而学习到知识。从对 MOOCs 教师与教学内容在屏幕上分布的了解与分析, 目前教师与教学内容在屏幕上的分布情况如表 2 所示。

除开早期的教学内容讲解只有语音而没有画面内容外, 一般的教学视频都具有以教学内容文字为主的内容。从传统学习的角度来看, 教学视频画面上以教学文字为主而教师画面为辅的教学视频最易于请人接受, 因为它与传统课堂的感觉相近, 学习者可在在学习内容文字与教师之间视线切换, 更容易创造身临其境的感觉。但这种方式显示需要使用外导式的拍摄方式, 对教师与教学设备都提出较高的要求。而画面上只有教学内容文字, 或只有实例操作的画面, 虽然没有教师本人出现, 但因为有了教师的讲解, 教师可以自己使用内导式方式制作教学视频, 显然成为当前 MOOCs 教学视频大多数的使用方式。学生只能从听力方面来感受教师的情绪, 不容易做到学习情感接受, 所以一般视频不能太长。而具有教学内容文字的画面与教师画面交差出现的方式, 情感上易于被接受, 但做到教学内容的文字伴随教师活动出现, 会使学习者遗失某部分的学习内容, 而且画面不断的交差出现, 容易造成学习者的不稳定。这种目

Table 2. Screen distribution of contents of MOOCs

表 2. MOOCs 视频教学内容屏幕分布方式

教学活动内容屏幕分布	说明, 来源
	教学内容文字与教师画面并列 History of Internet, technology, security 来源: Coursera
	教学内容文字与教师讲解画面互相切换 Web Intelligence and Big Data 来源: Coursera
	只有具体的活动内容或实践活动, 没有教学内容文字 Interaction programming in python 来源: Coursera
	只有教学内容文字 Preparing for the AP Calculus AB and BC Exams Introduction to Computing 来源: Coursera
	只有音乐, 没有任何教学内容的文字和教师讲解的画面, 但下方有字幕 Game Design (MIT) 来自 MIT

前在教学视频中出现较少。ICS 教学视频由于其录制的方式与现场, 采用的是全桌面的教学内容呈现方式, 因为是课堂教学现场录制的, 桌面上内容就是呈现给学习者的内容, 由于 UH 采用平板式 Classroom Presentation 软件, 它可以导入 PPT、word、PDF 或视频、FLASH 等文档, 也可以直接切换到白板供教师板书, 所以在画面上可以呈现多元化的教学内容, 甚至是教师的推算过程。使到学习者在观看、学习 ICS 教学视频时不会感觉枯燥。

5. 总结

在当今 MOOCs 盛行并且还将被不断扩大推广的情况下[18], 基于自动索引、关键字检索和字幕编辑技术下的 ICS 教学视频的开发与运行在远程与个性化、自主学习课程下, 证明是有效且高效率的。首期的 ICS 教学视频在 UH 全校开展, 在 1000 多个不同学科学生中进行了严格的评估, 得到了学生正面的肯定, 认为开发与推广此类课程是有价值的, 有益的, 它可以用在不同的方面, 如强化学习、考试复习等, 而且使用方法简单、方便, 易于掌握[19]。随着对 ICS 教学视频不断的深入研究, 其愿景是以更灵活的教学视频课程方式, 帮助具有特别需要的学生可以在更灵活的时间、任何地方, 以学生为主控角色进行更有效的学习。

参考文献 (References)

- [1] (2014) Speech Processing in Google [EB/NL].
- [2] Walsh, T. (2011) UNLOCKING the GATES: How and why leading universities are opening up access to their courses. Princeton University Press, Princeton, 5-10.

- [3] Subhlok, J., Johnson, O., Subramaniam, V., Vilalta, R. and Yun, C. (2007) Tablet PC video based hybrid coursework in computer science: Report from a pilot project. SIGCSE'07, 7-10 March, Covington, Kentucky, USA.
- [4] Tuna, T., Subhlok, J. and Shah, S. (2011) Indexing and keyword search to ease navigation in lecture videos. *Applied Imagery Pattern Recognition Workshop*, Washington, DC, 11-13 October 2011, 1-8.
- [5] Lienhart, R. and Effelsberg, W. (2000) Automatic text segmentation and Text recognition for video indexing. *Multimedia Systems*, **8**, 69-81.
- [6] Pang, D., Halawa, S., Cheung, N.-M. and Girod, B. (2011) Mobile interactive region-of-interest video streaming with crowd-driven prefetching. *International ACM Workshop on Interactive Multimedia on Mobile and Portable Devices, ACM Multimedia (MM'11)*, Scottsdale, November 2011, 7-12.
- [7] Odhabi, H. and Nicks McCaleb, L. (2011) Video recording lectures: Student and professor perspectives. *British Journal of Educational Technology*, **42**, 327-336.
- [8] (2013) Best OCR Font Size, Computer vision [EB/NL].
<http://www.cvisiontech.com/pdf/pdf-OCR/best-font-and-size-for-OCR.html?lang=eng>
- [9] Kim, J., Guo, P.J., Seaton, D.T., Mitros, P., Gajos, K.Z. and Miller, R.C. (2014) Understanding in-video dropouts and interaction peaks in online lecture videos [J/OL]. Publication rights licensed to ACM. ACM 978-1-4503-2669-8/14/03, Atlanta.
- [10] (2014) Speech Processing [EB/OL].
- [11] Heigold, G., McDermott, E., Vanhoucke, V., Senior, A. and Bacchiani, M. (2014) Asynchronous stochastic optimization for sequence training of deep neural networks. *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Firenze, 2014.
- [12] (2013) 电信手机大全基于内录式的 MOOCs 视频制作与分析 [EB/NL].
- [13] (2014) MOOCs 制作与运营——第二章「课程制作」 [EB/NL].
<http://team.pkubyttime.com.cn/post/53.html>
- [14] Anderson, T. and Dron, J. (2011) Three generations of distance education pedagogy. *International Review of Research in Open and Distance Learning*, **12**. <http://www.irrodl.org/index.php/irrodl/article/view/890>
- [15] Daniel, J. (2012) Making sense of MOOCs: Musings in a maze of myth, paradox and possibility. *Journal of Interactive Media in Education*, 1-20.
- [16] Kolowich, S. (2013) The professors who make the MOOCs [EB/NL]. The chronicle of higher education.
<http://publicservicesalliance.org/wp-content/uploads/2013/03/The-Professors-Behind-the-MOOC-Hype-Technology-The-Chronicle-of-Higher-Education.pdf>
- [17] Cooper, S. and Sahami, M. (2013) Reflections on Stanford's Moocs. *Communications of the ACM*, **56**, 28-30.
- [18] (2014) Coursera 联合创始人达芙妮·科勒: MOOC 课程范围越广越好 [EB/NL].
<http://mooc.guokr.com/opinion/437560/>
- [19] Tuna, T., Subhlok, J. and Shah, S. (2012) Development and evaluation of indexed captioned searchable videos for STEM coursework. *Proceedings of the SIGCSE Technical Symposium on Computer Science Education—SIGCSE'12*, March 2012, Raleigh, 129-134.