

doi: 10.12068/j.issn.1005-3026.2018.02.005

基于重叠式元胞移位的音频广告统计方法

刘纪红, 解丽岩

(东北大学 信息科学与工程学院, 辽宁 沈阳 110819)

摘 要: 根据广播广告监测的需求,提出一种基于重叠式元胞移位的音频广告检测算法.由于广告类型的多样性,该算法以多种类型的广告作为研究对象,并结合元胞自动机的基本理论和音频本身所具有的特点,采用重叠式元胞移位与互相关函数相结合的检测算法,实现音频广告的检测.实验结果表明:基于重叠式元胞移位的音频广告检测算法可以快速准确地从音频广播节目中检测出广告所在时间段,统计得到广告出现的次数,具有很好的鲁棒性.

关键词: 广告检测;广告统计;互相关函数;元胞自动机;重叠式元胞移位

中图分类号: TP 39 **文献标志码:** A **文章编号:** 1005-3026(2018)02-0172-04

Audio Commercials Statistics Method Based on Overlapping Cellular Shift

LIU Ji-hong, XIE Li-yan

(School of Information Science & Engineering, Northeastern University, Shenyang 110819, China. Corresponding author: XIE Li-yan, E-mail: 810361814@qq.com)

Abstract: To meet the demand of commercials-monitoring needed by the radio stations, an algorithm of audio commercials detection is proposed based on the overlapping cellular shift (OCS). Different types of commercials are used as research objects in the algorithm because of the diversity of commercials. Based on cellular automaton theory and the characteristics of audio, audio commercials are detected by using the method of the OCS combined the cross-correlation function. The experiment results show that the OCS-based algorithm can quickly and accurately detect the time section when commercials arise, and count the frequency of the commercials broadcasting from the radio, and the algorithm proposed is robust.

Key words: commercial detection; commercial statistics; cross-correlation function; cellular automaton; overlapping cellular shift

广告是商业信息的重要载体,广播电台在播出广告时可能会遇到一些不可避免的特殊情况,造成误播、延播,甚至漏播等现象,给广告主带来巨大的经济损失;因此广告主对于如何监播广播电台是否正常播放广告的问题十分关注^[1].

姜洪臣等^[2]提出了一种基于音频语谱图像识别的广告检索方法,该方法首先把每个广告片段的音频转换成语谱图像,然后通过 boosting 训练算法,从中提取可以区分的、适合索引的特征来训练分类器.该方法虽然能够检测出广告,但是现在广告特征不明显,在训练分类器时不易提取特

征. Borras 等^[3]设计了一种自动广告识别系统,通过提取被检测广告的音频特征进行广告检测,该方法能够检测到广告,但是占用较大存储空间;如何减少存储空间的占用问题还需要进一步的研究. Koolagudi 等^[4]提出了一种实时的广告检测算法,该算法提取音频广告的基本特征,利用机器学习方法分析广告和非广告特征,但是该算法的广告识别率仅为 87.05%,查全率相对较低.

现有的广告检测算法通常只是处理文字叙述类广告,这类广告特征较明显,通过提取音频特征,再对特征进行匹配.本文将音频广告分为文字

叙述和音乐播放两类,在此情况下进行广告检测.通过 MATLAB 仿真软件进行广告提取、单声道转换、音频数组转换、降低采样频率等一系列处理,并运用元胞自动机理论,提出一种基于重叠式元胞移位的音频广告检测算法(OCSD),检测出广告所在位置,统计出广告出现的次数.

1 元胞自动机和互相关函数

1.1 元胞自动机

元胞自动机(cellular automata, CA)是在有限状态的元胞组成的离散元胞空间上,按照一定的局部规则,在离散时间上进行演化的动力系统^[5].它由元胞、元胞空间、状态集、邻居和规则五个基本部分组成,具有同质性、齐性、空间离散性、元胞状态离散有限性、并行性、确定性、高维性等特点^[6-7].元胞自动机有很多种分类方式,按照维数可以分为一维、二维、三维和高维空间元胞自动机,本文在一维空间元胞自动机的基础上进行研究.

1.2 互相关函数

互相关函数是两个信号之间相似性或共享性的量度,可以应用于互谱分析、噪声信号的检测与恢复、模式匹配,以及延迟测量等^[6-7].互相关函数的定义为

$$r_{xy}(m) = \sum_{n=-\infty}^{\infty} x(n)y(n+m). \quad (1)$$

式中: x 和 y 为两个随机信号; n 和 m 为任意两个不同的时刻.

2 音频广告检测

在一维空间元胞自动机上,所有元胞等距地排列在一条可以无限延伸的直线上,每一个元胞的状态都是有限集合,某一时刻一个元胞的状态只与前一时刻此元胞状态以及前一时刻相邻元胞状态有关^[8].音频广告检测时,将检测的音频广告样本与音频广播在相邻时间段内进行比对,这与一维元胞状态的特点有着极其相似之处,因此本文考虑用一维元胞自动机建立音频广告检测的仿真模型.但是,音频连续性的特点使得无法在元胞自动机的基础上进行优化,因此设计重叠式元胞移位的方法并运用在音频广告检测仿真模型中.

2.1 音频广播中的元胞状态

在仿真模型中,把音频广播看作由一系列元胞组成,元胞大小与音频广告长度 L 相同,元胞

状态表示音频广告与音频广播的相似度,每个元胞有 6 种状态(0, 1, 2, 3, 4, 5).音频广告检测仿真模型如图 1 所示.

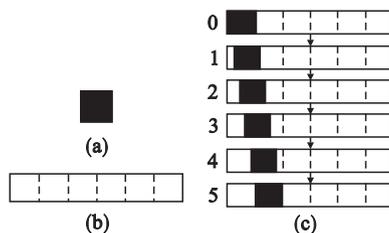


图 1 音频广告检测仿真模型

Fig. 1 Simulation model of audio commercials detection

(a) — 音频广告样本; (b) — 音频广播样本;
(c) — 元胞状态.

首先假设音频广告在音频广播起始位置,此时元胞上音频广播与音频广告样本的相似度最高,根据式(1)计算出的互相关值的最大值为 R_0 .根据式(1)计算出的每个元胞与音频广告样本互相关值的最大值若在 $(R_{i+1} + \delta_{i+1}, R_i + \delta_i]$ 范围内($i=0, 1, 2, 3, 4, 5$),则此元胞状态记为 i 状态,通过大量实验训练得到阈值 $\delta_i = 0.02R_i$.元胞状态 i 代表广告右移 $iL/5$ 后的状态,每次移位后元胞与音频广告的相似度减弱.通常 $R_i > R_{i+1}$,具体数值因音频广告与广播的不同会有所不同.

2.2 基于重叠式元胞移位的音频广告检测算法

图 1c 代表音频广告检测仿真模型中的元胞状态变化规律:音频广告自左向右重叠式移动,移动步长记为 S , $S = L/5000$.利用互相关检测算法计算每个元胞状态上音频广播与音频广告的互相关值,记录各个元胞状态的互相关最大值,标记各元胞的状态,进而得出音频广播与音频广告最匹配的状态.

OCSD 算法如下:

①采集音频广告样本 AdSample,音频广播样本 WavSample.

②判断数据是否为双声道音频信息,若是则进行变换,获得单声道数据.

③利用 $[Y, FS, NBITS] = \text{wavread}(\text{FILE})$ (FILE 为音频文件),提取出音频样本数据 Y ,再利用 $X = \text{decimate}(Y, R)$ (R 为采样要降低的倍数)降低采样频率,得到抽样后的音频数据 X ,以此提高检测速率.分别对广告样本和广播样本进行上述处理得到序列 $x(n)$ 和 $y(n)$.

④将音频广告样本作为序列 $x(n)$,采用重叠式元胞移位方式,在音频广播样本中按步长 S 进行移位,获得与 $x(n)$ 等长的序列,即元胞序列

$y_i(n)$, i 为移动的次數。

⑤计算 $x(n)$ 和 $y_i(n)$ 的互相关值, 并将最大值存入 a_i . 对 a_i 序列作出移位次数 - 互相关最大值曲线图, 利用 2.1 节中的音频广告检测仿真模型及表 1 计算得到的 R_i , 判断当前元胞序列 $y_i(n)$ 的元胞状态, 根据元胞状态的变化特征判断广告出现的粗略位置。

⑥在广告出现的粗略位置范围内, 根据互相关最大值 a_i , 求出最大互相关值 $a_{i,\max}$ 和最小互相关值 $a_{i,\min}$, 计算相对偏差:

$$\text{Diff} = (a_{i,\max} - a_{i,\min}) / a_{i,\min}.$$

⑦通过多类广告的多次试验, 设置阈值 ThreValue.

⑧比较 Diff 和阈值 ThreValue 的大小, 如果 $\text{Diff} \leq \text{ThreValue}$, 此音频广播段没有广告, 算法结束; 否则, 此音频广播段有广告, 继续下一步。

⑨去除互相关最大值曲线波动的干扰. 设置阈值 $T = (a_{i,\max} - a_{i,\min}) / 2$. 若 $a_i \leq T$, 将相关系数 a_i 归零; 若 $a_i > T$, a_i 数值不变。

⑩对经过步骤⑨处理后的互相关最大值 a_i 序列作出移位次数 - 互相关最大值曲线图. 利用连续曲线的极值法, 找到极值点, 并根据移位次数、移位步长和音频频率的关系, 计算广告所在位置, 并统计广告次数。

由 OCS D 算法可知, 序列 $x(n)$ 与 $y_i(n)$ 要作

i 次互相关运算 (i 为移动次数, N 为提取后的广告样本点数), 每次计算得到 N 个互相关值, 并对其求最大值. 所以算法的时间复杂度为 $O(i \times N)$. 这表明此算法的时间开销会受步长 S 以及音频频率的影响. 提高算法速度可以通过增大步长 S 及利用步骤③来降低 N 值, 但是如果 S 过大, N 过小, 会降低广告的定位精度, 所以应该适当选择步长及音频频率. 本文设步长 $S = L / 5\,000$, 广告时长为 T_{ad} , 定位精度为 $T_{\text{ad}} / 5\,000$.

3 仿真结果

本文所有实验数据均采集自辽宁广播电台的真实音频广播节目, 电台广播广告主要分为文字叙述类(商品宣传)和音乐播放类(以音乐为背景播放). 本文分别选择 10 段不同时长的两类广告. 广告 1 是最常听到的商品宣传广告, 以文字叙述为主. 广告 2 是在有背景音乐的情况下进行宣传的广告, 为了使检测更具有说服力, 本文采集一段在背景音乐下播放正常节目的音频广播(其中包含广告 2), 增加广告与广播之间的相似度. 将广告及广播均转换为 44.1 kHz, 16 bit 单声道的 wav 数据. 根据 OCS D 算法, 在 MATLAB 仿真平台下编程实现测试系统, 进行广告检测. R_i ($i = 0, 1, \dots, 5$) 值及广告检测结果如表 1 所示。

表 1 同类型广告的检测结果
Table 1 Detection results of different types of advertisements

广告类型	广播时长/s	广告时长/s	广告数量	识别率/%	R_0	R_1	R_2	R_3	R_4	R_5
广告 1 (文字叙述类)	300	5	5	100	1 650	1 505	1 210	1 010	750	300
	180	15	3	100	2 380	2 100	1 750	1 450	1 100	570
	265	30	2	100	2 800	2 450	2 000	1 600	1 015	620
广告 2 (音乐播放类)	300	5	5	100	15 500	12 000	8 000	6 000	2 500	1 000
	280	15	3	100	14 500	10 000	6 770	4 800	2 055	1 000
	210	30	2	100	16 770	12 500	8 055	5 050	3 500	2 060

由表 1 可以看出, 该算法能够识别不同时长、不同类型的广告样本, 并且广告识别率均为 100%. 定位精度与 OCS D 算法的步长 S 及广告时长有关, 步长 S 越小, 定位精度越高. 例如在检测 5 s 广告时, 定位精度为 0.001 s. 文献[9]中提出的基于音频匹配的广告智能监播算法, 仅对 3 s 的广告样本与 30 s 的音频样本进行音频匹配检测, 定位精度为 0.01 s, 广告识别准确率为 99.46%. 文献[9]的算法要求广告时长固定且没有对广告进行分类检测, 精确度和识别率均不如本文 OCS D 算法, 充分说明 OCS D 算法的有效性

和鲁棒性. 对广告 1 中一段时长为 15 s 的广告样本(出现 3 次)和广告 2 中一段时长为 30 s 的广告样本(出现 2 次)的检测结果如图 2、图 3 所示。

图 2a 中有 3 个峰值, 峰值大小为 2 343, 2 384, 2 380, 均在 $(2\,100 + 42, 2\,380 + 48)$ 范围内, 由 2.1 节知, 峰值对应的状态为元胞状态 0, 即状态 0 在音频广播中出现了 3 次, 表明在这 3 个峰值所对应的时间段内音频广告与音频广播相关性最大, 粗略估计广告就在这 3 个峰值所对应的时间段内; 由图 2b 可见, 经过 OCS D 算法步骤⑨处理后去除了图 2a 中的波动现象, 避免了图

2a 中微小峰值的干扰; 图 2c 中广告播放次数的文本框中显示为 3, 表明通过 OCSD 算法中步骤 ⑩ 的进一步检测, 该音频广播中此类广告确实出现了 3 次, 第一次出现在第 13.421 s, 第二次出现在第 83.471 s, 第三次出现在第 150.551 s.

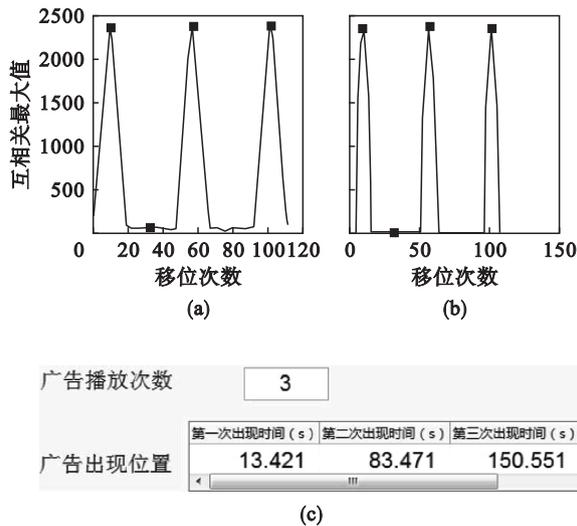


图 2 广告 1 的仿真结果

Fig. 2 Simulation results of commercial 1

(a)—OCSD 算法步骤 ⑤ 中的 a_i 序列曲线;
(b)—步骤 ⑨ 处理后 a_i 序列曲线; (c)—检测结果.

由图 3 可知, 该音频广播中广告 2 出现了 2 次, 第一次出现在第 75.001 s, 第二次出现在第 141.010 s. 另外, 广告 2 中音频广告与音频广播都是以音乐作为背景, 所以两者的相关性比较大 (见图 3a), 互相关值明显高于普通广告的互相关值 (见图 2a), 这表明即使在相关性很大的情况下, OCSD 算法也可以精确检测出广告.

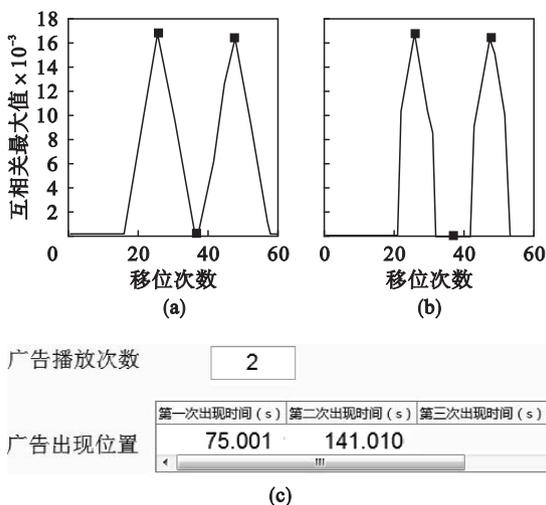


图 3 广告 2 的仿真结果

Fig. 3 Simulation results of commercial 2

(a)—OCSD 算法步骤 ⑤ 中的 a_i 序列曲线;
(b)—步骤 ⑨ 处理后 a_i 序列曲线; (c)—检测结果.

4 结 论

1) 与传统的基于音频匹配广告检测算法相比, OCSD 算法简单, 抗干扰能力强.

2) OCSD 算法不仅适用普通的以文字叙述为主的广告, 还适用以音乐为背景的广告, 适用范围广, 识别准确度极高, 通过选择合适的步长 S , 定位精度可达到 0.001 s, 可以有效地统计和监测广告播放情况, 具有很好的实用性.

参考文献:

- [1] Qin Y. Efficient audio recognition algorithm based on simple multiple kernel learning [J]. *Journal of Multimedia*, 2015, 9 (1): 59 - 66.
- [2] 姜洪臣, 任晓磊, 赵耀宏, 等. 基于音频语谱图识别的广告检索 [J]. *清华大学学报 (自然科学版)*, 2011, 51 (9): 1249 - 1252.
(Jiang Hong-chen, Ren Xiao-lei, Zhao Yao-hong, et al. Advertisement retrieval based on audio spectrogram image recognition [J]. *Journal of Tsinghua University (Science and Technology)*, 2011, 51 (9): 1249 - 1252.)
- [3] Borrás J E, Igual J, Fernández-Llata C, et al. A TV commercial retrieval system based on audio features [C]// *Proceedings of the 10th International Conference on Signal Processing and Multimedia Applications*. Reykjavik, Iceland: IEEE, 2013: 65 - 70.
- [4] Koolagudi S G, Sridhar S, Elango N, et al. Advertisement detection in commercial radio channels [C]// *Proceedings of 2015 IEEE 10th International Conference on Industrial and Information Systems*. Peradeniya, Sri Lanka, 2015: 272 - 277.
- [5] Gong J Z, Chen G L, He H, et al. Analysis of hexagonal network based on cellular automation [J]. *Applied Mechanics and Materials*, 2014, 6 (8): 1457 - 1463.
- [6] 刘纪红, 孙宇舫, 叶柠, 等. 数字信号处理原理与实践 [M]. 修订版. 北京: 清华大学出版社, 2014: 23 - 25.
(Liu Ji-hong, Sun Yu-fang, Ye Ning, et al. Principles and practice of digital signal processing [M]. Rev ed. Beijing: Tsinghua University Press, 2014: 23 - 25.)
- [7] 肖倩, 王建辉, 方晓柯, 等. 一种基于互相关函数的小波系数相关阈值去噪方法 [J]. *东北大学学报 (自然科学版)*, 2011, 32 (3): 318 - 321.
(Xiao Qian, Wang Jian-hui, Fang Xiao-ke, et al. A wavelet coefficient threshold denoising method based on cross-correlation function [J]. *Journal of Northeastern University (Natural Science)*, 2011, 32 (3): 318 - 321.)
- [8] Wen W H, Qiu Y H, Liu G Y, et al. Construction and cross-correlation analysis of the effective physiological response database [J]. *Science China: Information Sciences*, 2010, 53 (9): 1774 - 1784.
- [9] 张华熊, 朱诗威, 章晨衍, 等. 基于音频匹配的广告智能监测算法 [J]. *电子器件*, 2008, 31 (2): 669 - 671.
(Zhang Hua-xiong, Zhu Shi-wei, Zhang Chen-yan, et al. Intelligent advertisement detecting algorithm based on audio matching [J]. *Chinese Journal of Electron Devices*, 2008, 31 (2): 669 - 671.)