

网络技术的快速发展和多媒体产业的迅速发展已经越来越在我们的日常生活中产生重要的影响。我们也很容易就接触到很多数字音频资源。随着人们知识产权意识的不断增强，这些巨大数量的音频资源中就有很多涉及到版权问题，而且存在严重的管理问题。而经过媒体密码加密技术和数字水印技术在数字版权管理方面的研究和发展之后，音频指纹技术越来越得到大家的广泛关注和研究兴趣。除版权保护之外，音频指纹技术还有很多其它的巨大的实际应用前景，如为手机用户提供音乐识别服务。不论音频资源经过一定程度的什么变形操作，如采样率变化，噪声干扰，音频数据压缩等，通过音频指纹技术都能正确地找到音质好的源音频文件。

本文研究工作主要围绕音频指纹技术进行，致力于研究有效的音频指纹特征提取算法。本文的研究主要分成以下两个方面：音频指纹算法，音频指纹系统的设计开发。其中音频指纹算法是本文的研究重点。使用音频指纹算法是寻找音频文件中最“本质”，最“唯一”的特征属性作为该音频资源的唯一标识。为了使音频指纹能够满足音频版权保护和其它的应用的需求，音频指纹需要具备健壮性，可靠性，小尺度等特征。因此，本文中提出了基于 MPEG-7 音频描述子的音频指纹特征提取算法，并就音频指纹系统提出了相应的特征查找策略和方法。

音频指纹的研究目前已经有了一定的时间，不论是各种研究机构还是公司都在音频指纹方面取得了一定的成果。Philips 研究人员提出了一种有效的基于频谱多子带特征提取的 PRH 音频指纹提取算法，Microsoft 的研究人员提出了一种基于 OPCA 的全局降维的 RARE 的指纹提取算法，MPEG 组织提出了多个音频描述子用来描述音频特征。但公众能够获得的音频研究文档却不多，因为很多技术已经直接用于公司商业上，如 Gracenote 公司的 Mobile MusicID 音频指纹技术和 Auditude 公司的用于音乐识别的音频指纹技术等等。

基于上述研究发展情况，本文的主要工作和贡献如下：

提出了一种基于 MPEG-7 描述子 Audio Spectral Flatness 的音频指纹提取算法。该算法通过人耳听觉系统中的人耳模拟函数进一步修正音频频谱值，而通过引入一种简单而有效的平滑滤波器组能够有效地去除音频频谱中的噪声成分，而对 Audio Spectral Flatness 描述子增加权值的改进能进一步获得有效的音频指纹特征。而使用 Audio Signature (AS) 高层描述子能够有效地压缩计算获得的音频指纹特征，实现小的指纹尺度。对于音频指纹特征的查找，它不同于图像特征的查找。由于要查询的音频可能是一个完整音频的任意位置开始的一个片段，因此在指纹数据库中查找特征时，使用一种滑动的特征查找方法，每次滑动的距离要小于求出指纹特征所使用的滑动窗口。

开发了一个基于以上音频指纹算法音频指纹验证系统以验证算法的有效性以及一个实际的指纹应用系统。

**关键词：**音纹；音频指纹；平滑滤波器；加权 ASF；