1. 对输入的100组音频按人声进行分组