An experimental study on footsteps sound recognition as biometric under noisy conditions

Un estudio experimental sobre el reconocimiento del sonido de las pisadas como biométrico en condiciones ruidosas

Marisol Zeledón-Córdoba[1], Carolina Paniagua-Peñaranda[2], Marvin Coto-Jiménez[3]

Zeledón-Córdoba, M.; Paniagua-Peñaranda. C.; Coto-Jiménez, M. An experimental study on footsteps sound recognition as biometric under noisy conditions. *Tecnología en Marcha*. Vol. 35, special issue. November, 2022. International Work Conference on Bioinspired Intelligence. Pág. 153-161.

https://doi.org/10.18845/tm.v35i8.6467

1 Electrical Engineering Department. University of Costa Rica. Costa Rica.
E-mail: marisol.zeledon@ucr.ac.cr
https://orcid.org/0000-0001-7481-0207
2 Electrical Engineering Department. University of Costa Rica. Costa Rica.
E-mail: carolina.paniaguapenaranda@ucr.ac.cr
3 Electrical Engineering Department. University of Costa Rica. Costa Rica.
E-mail: marvin.coto@ucr.ac.cr
https://orcid.org/0000-0002-6833-9938

Keywords

Biometric; classification; footsteps; sounds.

Abstract

The experimentation of footsteps as a biometric has a short history of about two decades. The process of identification of a person is based on the study of footstep signals captured when walking over a sensing area, and the registering of sounds, pressure, vibration, or a combination of these measures. Application of this biometric can emerge in security systems, that identify persons who enter or leave a space, and in providing help to elderly and disabled persons. In this paper, we are focused in the exploration of pure audio signals of footsteps and the robustness of a person's classification under noisy conditions. We present a comparison between four well-known classifiers and three kinds of noise, applied at different signal to noise ratio. Results are reported in terms of accuracy in the detection an users, showing different levels of sensibility according to the kind and level of noise.

Palabras clave

Biométrico; clasificación; pisadas; sonidos.

Resumen

La experimentación de las pisadas como biométrico tiene una breve historia de unas dos décadas. El proceso de identificación de una persona se basa en el estudio de las señales de pisadas capturadas al caminar sobre un área de detección y el registro de sonidos, presión, vibración o una combinación de estas medidas. La aplicación de esta biometría puede surgir en los sistemas de seguridad, que identifican a las personas que entran o salen de un espacio, y en la prestación de ayuda a las personas mayores y discapacitadas. En este artículo, nos centramos en la exploración de señales de audio puras de pasos y la solidez de la clasificación de una persona en condiciones ruidosas. Presentamos una comparación entre cuatro clasificadores conocidos y tres tipos de ruido, aplicados a diferentes relaciones señal / ruido. Los resultados se informan en términos de precisión en la detección de los usuarios, mostrando diferentes niveles de sensibilidad según el tipo y nivel de ruido.

Introduction

In the area of biometrics, the most common elements for verifying a person's identity are fingerprints and face, usually applied in smartphones. Others, such as iris identification have also been successfully used in applications in airports [1]. These are examples of the physiological group of biometrics. Apart from the physiological group, the use of behavioral biometrics, such as voice recognition has also received considerable attention in recent years. Footsteps signals are an- other example of behavioral biometrics, with a shorter history of active research. Footstep signals are signals collected from people walking over an instrumented sensing area [1], with the aim of classification or analysis of the individuals. The proposal for analyzing and evaluating such signals was introduced by [2], remarking its simplicity for clinical practice.

But it was until 1997 that the first experiments with sensors in an active floor were presented [3]. Since then, several researchers have offered different approaches to the application of footsteps or gait analysis as a biometric. These works have demonstrated the real potential of the footstep biometric [4].

The human identification and analysis of footsteps can have applications in medicine, surveillance, sport shoe industry, smart homes, and multimedia [5, 6]. Given the relatively recent experiences with this measure, many recent studies have presented the building of datasets with particular measurements of sensors, features, classifiers, and conditions [5].

Additionally, there are areas of concern in their usage in terms of practicality, privacy, and security [7]. For this reason, the scientific literature on footsteps and gait analysis has grown significantly in the past years, with sensors based on vision, sound, pressure, and accelerometry [8], with the corresponding set of features.

In this work, we present the building of a dataset of footsteps sounds and report on the first experiments on the robustness of several classifiers for this biometric in the presence of noise. For this experience, we consider only the discrimination of footsteps between two individuals of similar age and gender, where the challenges are more relevant.

Related work

The assessment of signals for their application as a biometric involves developing datasets with a large number of labelled examples [9]. For lesser researched biometrics, such as footsteps, new databases are needed in order to assess the accuracy and study other practical aspects. In particular, the registering of only the sound of footsteps in a distant microphone requires the development of such datasets.

Several studies have reported the potential of using gait information to distinguish between people. Previous references have reported its use in criminal cases to identify perpetrators based on their walking behavior, and also in the identification of patterns such as the Parkinson [8].

Previous reports have achieved around 80% to 90% [6, 1] in the identification of individuals using footstep information. But a wide variety of conditions do not allow a precise comparison of the efforts or the establishment of benchmarks in this field for unexplored sensing conditions, such as the pure sound of the footsteps.

Among the many conditions that can affect the performance of the sound of footsteps as a biometric are the different types of footwear worn, like heels, sneakers, leathers or even barefooted, and the corresponding sound in different grounds of concrete, wood or other materials [10, 11].

A comparison of features and sensors during the first decade of studies in footsteps recognition, as presented in [5] has not registered the application of pure sound signals in the identification of individuals. For this reason, the exploration of features derived from sound signals requires the building of a dataset to explore it. With microphones being one of the simplest and cheapest sensors, its application in biometric of footsteps can be particularly useful.

One of the most common issues related to sound signals is the presence of noise. In a real-life application of biometrics with footsteps sounds, the presence of noise and sounds other than footsteps is continuous. In this work, we explore the building and testing of the sound of footsteps as a biometric, using a single microphone to register the distant sound of footsteps, and compare the performance of several classifiers under noisy conditions. Being the first dataset and the focus on the impact of noise, we consider the simplest case of binary classification of two users.

The rest of this paper is organized as follows: Section 2 presents the experimental setup for building the dataset and the experimentation. Section 3 presents the results and discussion, and finally Section 3 presents the conclusions and future work.

Experimental Setup

For evaluating how background noise can affect the identification of an individual using footsteps sounds, we developed recording sessions with several male and female participants. For this work, we only considered pairs of recordings of two women's footsteps, where the distinction between two people is more challenging, according to our first experiences.

The recordings were made using a single microphone with an Omnidirectional pattern, and the participants were asked to walk naturally and continually in a circle of 1.5 m around the microphone. Figure 1 shows the basic setup for the sessions. Each participant recorded about fifteen minutes of her footsteps.

Figure 1. Recording session.

Each recording was then edited in segments of five seconds, in order to cap- ture at least 3 footstep sounds in each segment. For each type of noise considered in this work, specific SNR levels were added to each file. This means that there are several versions of the sound for each audio segment: those with specific SNR level of each type of noise, and the one without any noise added (clean version). For each segment of the whole set of conditions, a set of features were extracted, corresponding to several categories of audio descriptors. For example, the energy, zero crossing rate, entropy, MFCC and chroma features. The complete description of the features corresponds to those presented in [12].

For each participant, about 3000 files were generated, considering the whole set of conditions. The set of features for each of such files were tested using four common classifiers: K-Nearest Neighbors (KNN), Naive Bayes, Random Forest, and Support Vector Machine (SVM). In each case, a ten-fold cross-validation was used to assess the accuracy of each classifier under each condition.

Results and Discussion

This section presents the results obtained from the different classifiers and the noise levels analyzed. In each of the tables, a noise level with the accuracy percentage obtained for that level is presented. The results are reported in terms of the distinction between two people. In other words, all the results correspond to binary classification as a way to analyze the first case of an experimental study that uses only audio signals for biometric identification.

Table 1. Accuracy of classifiers for different levels of Babble Noise.

Clean				
	KNN	N. Bayes	R. Forest	SVM
P_1 vrs P_2	96.6%	96.4%	98.2%	98.1%
P_1 vrs P_3	97.5%	98.0%	99.4%	100%
P_2 vrs P_3	100%	99.4%	100%	100%
SNR-10				
P_1 vrs P_2	18.7%	48.0%	34.7%	64.0%
P_1 vrs P_3	16.0%	57.3%	50.7%	66.7%
P_2 vrs P_3	14.7%	45.3%	30.7%	58.7%
SNR-5				
P_1 vrs P_2	46.7%	78.7%	81.3%	77.3%
P_1 vrs P_3	44.0%	81.3%	78.7%	80.0%
P_2 vrs P_3	25.3%	74.7%	81.3%	93.3%
SNR 0				
P_1 vrs P_2	62.7%	82.7%	86.7%	92.0%
P_1 vrs P_3	70.7%	88.0%	89.3%	92.0%
P_2 vrs P_3	61.3%	96.0%	96.0%	97.3%
SNR 5				
P_1 vrs P_2	84.0%	92.0%	90.7%	96.0%
P_1 vrs P_3	85.3%	92.0%	94.7%	98.7%
P_2 vrs P_3	94.7%	97.3%	98.7%	100%
SNR 10				
P_1 vrs P_2	88.0%	88.0%	93.3%	97.3%
P_1 vrs P_3	90.7%	93.3%	98.7%	98.7%
P_2 vrs P_3	98.7%	98.7%	98.7%	100%

For example, in Table 1, it can be observed how Babble noise affects the performance of the classifiers. It can be established that noise levels higher than SNR0 (such as SNR-5 and SNR-10) produce accuracy levels above 50%. In other words, they affect several classifiers even below the level of random identification.

Among the classifiers used, SVM stands out as one that maintains the most stable accuracy percentage at all noise levels, although it is also severely affected at the highest noise levels.

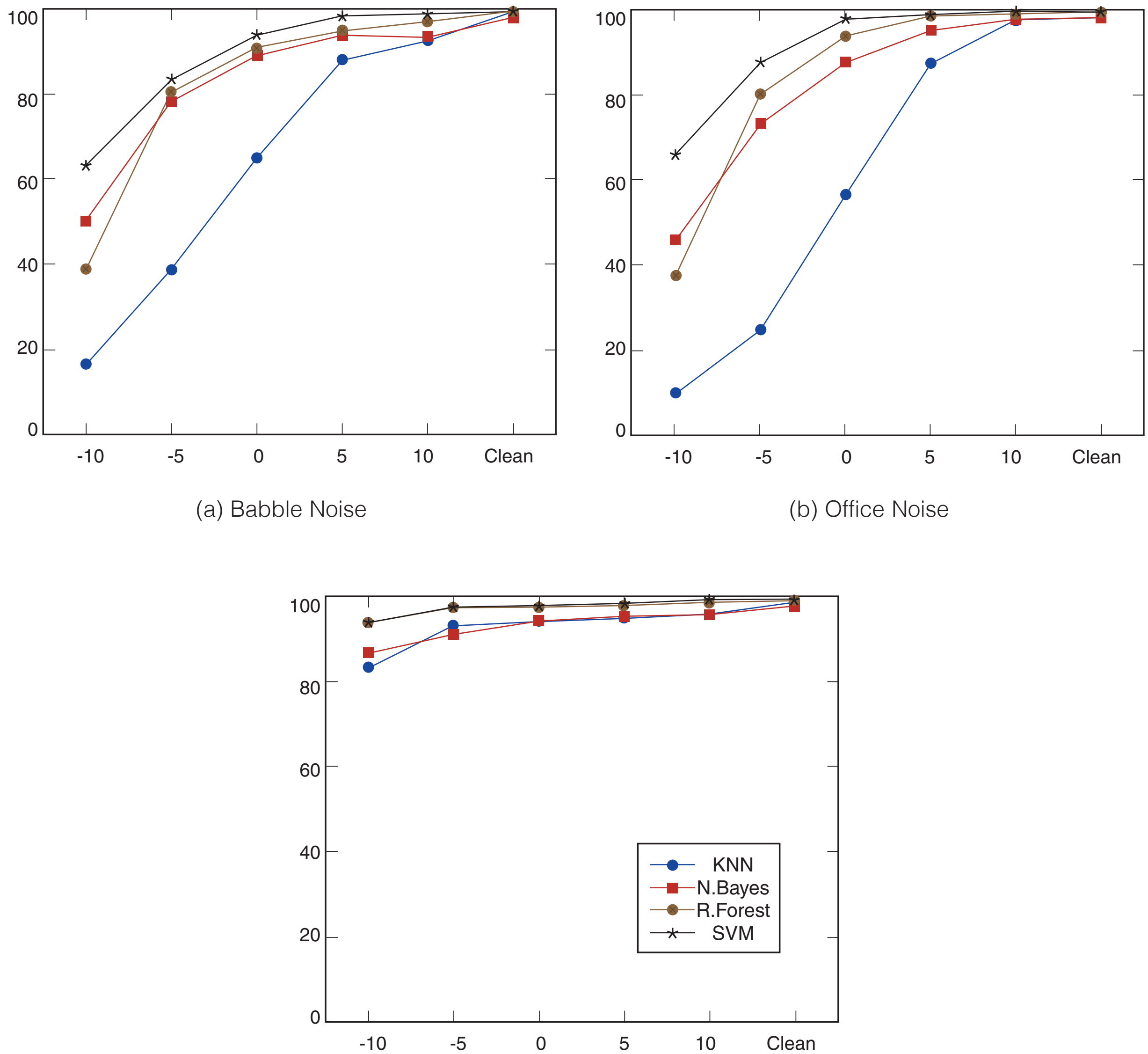
Figure 2 verifies that the KNN classifier is the one that is affected the most by Babble noise, while SVM is the one presenting the highest level of robustness. In terms of classification of Clean footstep sounds, the accuracy percentage of all classifiers is very similar.

The results obtained by adding Office noise are shown in Table 2. For this kind of noise, it can be observed that, compared with the previous case, the performance of the classifiers is affected in a differing way. For example, the accuracy of KNN decreases very quickly, while the accuracy of SVM and Naive Bayes is affected more lightly. In the case of Random Forest, it is until SNR-10 that the accuracy drops considerably.

Table 2. Accuracy of classifiers for different levels of Office Noise.

Clean				
	KNN	N. Bayes	R. Forest	SVM
P_1 vrs P_2	96.6%	96.4%	98.2%	98.1%
P_1 vrs P_3	97.5%	98.0%	99.4%	100%
P_2 vrs P_3	100%	99.4%	100%	100%
SNR-10				
P_1 vrs P_2	7.2%	50.0%	41.6%	67.6%
P_1 vrs P_3	17.0%	51.6%	47.2%	72.8%
P_2 vrs P_3	4.4%	35.2%	22.8%	56.8%
SNR-5				
P_1 vrs P_2	25.2%	80.0%	84.4%	88.8%
P_1 vrs P_3	31.6%	77.2%	83.2%	88.0%
P_2 vrs P_3	16.4%	62.0%	72.0%	85.2%
SNR 0				
P_1 vrs P_2	58.0%	87.6%	90.8%	94.8%
P_1 vrs P_3	59.2%	88.0%	94.8%	98.4%
P_2 vrs P_3	51.6%	86.8%	95.6%	99.6%
SNR 5				
P_1 vrs P_2	86.8%	92.8%	96.8%	97.2%
P_1 vrs P_3	84.0%	95.2%	98.8%	99.2%
P_2 vrs P_3	90.8%	97.2%	99.2%	99.6%
SNR 10				
P_1 vrs P_2	96.8%	98.0%	97.6%	98.8%
P_1 vrs P_3	96.0%	95.6%	99.2%	99.6%
P_2 vrs P_3	99.6%	99.2%	100%	100%

Figure 2b shows how having higher noise levels, as in the case of SNR-5 and SNR-10, the performance of the KNN classifier decreases significantly, compared to classifiers such as Naive Bayes and SVM where the performance decreases less. It can also be observed that the performance of the classifiers was affected similarly to that obtained with Babble noise, where the accuracy of the KNN, Naive Bayes, and Random Forest classifiers is below 50% for noise levels higher than SNR0.



(a) Babble Noise

(b) Office Noise

(c) White Noise

Figure 2. Accuracy of classifiers as a function of SNR (Mean of the three comparisons).

Finally, Table 3 shows the results for White noise. In these results, it can be seen that White noise affects classifiers to a lesser extent than the Babble and Office noises do. It can also be established that, in this case, White noise affects the KNN and Naive Bayes classifiers slightly more, this difference is more pronounced as the noise level increases.

None of the classifiers decrease their performance considerably; however, the ones that remain more constant are Random Forest and SVM compared to KNN and Naive Bayes.

Table 3. Accuracy of classifiers for different levels of White Noise.

Clean				
	KNN	N. Bayes	R. Forest	SVM
P_1 vrs P_2	96.6%	96.4%	98.2%	98.1%
P_1 vrs P_3	97.5%	98.0%	99.4%	100%
P_2 vrs P_3	100%	99.4%	100%	100%
SNR-10				
P_1 vrs P_2	93.6%	94.0%	98.0%	97.6%
P_1 vrs P_3	78.0%	81.6%	91.6%	92.8%
P_2 vrs P_3	79.2%	85.2%	92.4%	92.0%
SNR-5				
P_1 vrs P_2	96.8%	98.0%	99.6%	98.0%
P_1 vrs P_3	88.8%	82.0%	94.0%	96.4%
P_2 vrs P_3	94.0%	94.0%	98.8%	98.8%
SNR 0				
P_1 vrs P_2	99.2%	98.4%	98.8%	99.6%
P_1 vrs P_3	88.0%	87.6%	94.4%	95.2%
P_2 vrs P_3	95.6%	97.6%	97.6%	99.2%
SNR 5				
P_1 vrs P_2	99.2%	98.4%	99.2%	100%
P_1 vrs P_3	88.0%	88.8%	94.8%	95.6%
P_2 vrs P_3	98.0%	99.2%	100%	100%
SNR 10				
P_1 vrs P_2	98.0%	97.6%	99.2%	100%
P_1 vrs P_3	91.6%	90.4%	96.8%	98.4%
P_2 vrs P_3	98.8%	98.8%	100%	100%

Figure 2c confirms that White noise affects the KNN and Naive Bayes classifiers at high noise levels (SNR-10 and SNR-5) to a greater extent. In addition, it is observed that the performance of KNN and Naive Bayes is very similar, SVM and Random Forest also present a very similar behavior, the latter two being the ones with the best performance. As for the Clean steps, the classifiers have a similar accuracy.

Conclusions

In this work, a first study was presented on the use of the sound of steps captured with a single remote omni-directional microphone, as a means of biometric iden- tification. Since an application of this type implies its use in a real environment, where there is always noise pollution, the main interest has been to analyze how much additive noise conditions can affect the identification of a person using the sound of their steps. For this, a database of step sounds recorded with controlled environmental conditions was developed. The binary classification in pairs of participating individuals was also used.

The results show a different affectation in the classifiers, for which those that can be considered simpler, such as KNN, are greatly affected as the noise level increases, compared to others such as SVM and Naive Bayes that have a greater robustness. The type of noise also affects the affectation in a different way, where white noise presents little affectation in the identification of the person. In contrast, a natural noise such as the sounds of an office affects the classifiers much more.

As future work, the results of this work can be expanded in several directions, such as in the use of sound as a biometric identifier for a particular person with respect to a broader set of individuals and the impact of noise in this case. Noise reduction methods can also be tested for the developed database and to see if it allows for filtered sound to be considered as a means of biometric identification, in the case of favorable results. Finally, the possibility of improving the classification can be done using methods such as mixing experts and other classifiers based on deep learning.

References

[1] Rodríguez, Rubén Vera, et al. “Footstep recognition for a smart home environment.” International Journal of Smart Home 2.2 (2008): 95-110.

[2] Pedotti, Antonio. “Simple equipment used in clinical practice for evaluation of locomotion.” IEEE Transactions on Biomedical Engineering 5 (1977): 456-461.

[3] Addlesee, Michael D., et al. “The ORL active floor [sensor system].” IEEE Personal Communications 4.5 (1997): 35-41.

[4] Rodriguez, Ruben Vera, Nicholas WD Evans, and John SD Mason. “Footstep recognition. ”Encyclopedia of Biometrics; Li, SZ, Jain, AK, Eds.; Springer: Boston, MA, USA (2015): 693-700.

[5] Rodríguez, Rubén Vera, et al. “An experimental study on the feasibility of footsteps as a biometric.” 2007 15th European Signal Processing Conference. IEEE, 2007.

[6] Vera-Rodríguez, Rubén, et al. “Analysis of time domain information for footstep recognition.” International Symposium on Visual Computing. Springer, Berlin, Heidelberg, 2010.

[7] Mason, James Eric, Issa Traoré, and Isaac Woungang. “Gait Biometric Recognition.” Machine Learning Techniques for Gait Biometric Recognition. Springer, Cham, 2016. 9-35.

[8] Connor, Patrick, and Arun Ross. “Biometric recognition by gait: A survey of modalities and features.” Computer Vision and Image Understanding 167 (2018): 1-27.

[9] Vera-Rodríguez, Rubén, et al. “A large scale footstep database for biometric studies created using cross-biometrics for labelling.” 2008 10th International Conference on Control, Automation, Robotics and Vision. IEEE, 2008.

[10] Shoji, Yasuhiro, Takashi Takasuka, and Hiroshi Yasukawa. “Personal identification using footstep detection.” Proceedings of 2004 International Symposium on Intel- ligent Signal Processing and Communication Systems, 2004. ISPACS 2004. IEEE, 2004.

[11] Hori, Yuki, Takahiro Ando, and Akira Fukuda. “Personal Identification Methods Using Footsteps of One Step.” 2020 International Conference on Artificial Intelligence in Information and Communication (ICAIIC). IEEE, 2020.

[12] Giannakopoulos, Theodoros. “pyaudioanalysis: An open-source python library for audio signal analysis.” PloS one 10.12 (2015): e0144610.