COMBINING CLASSIFIERS WITH DIFFERENT FOOTSTEP FEATURE SETS AND MULTIPLE SAMPLES FOR PERSON IDENTIFICATION

Jaakko Suutala and Juha Röning

Intelligent Systems Group, Infotech Oulu
P.O. Box 4500
90014 University of Oulu, Finland
{jaakko.suutala, juha.roning}@ee.oulu.fi

ABSTRACT

Combination of classifiers is usually a good strategy to improve accuracy in pattern recognition systems. In this paper, we present a new approach to footstep-based biometric identification by combining pattern classifiers with different feature sets. Footstep profiles are obtained from a pressure-sensitive floor. Our identification system consists of two different combination stages. At the first stage, three pattern classifiers, trained with feature sets presenting different characteristics of input signal, are combined. The feature sets include the spatial domain properties of the footstep profile as well as the frequency domain presentation of the signal and its derivative. At the second stage, multiple input samples are combined, using the posterior probability outputs from the first stage, to make the final decision. The building blocks of the classification system are examined, and the methodological justifications are analyzed. The experimental results show improvements in identification accuracies compared to previously reported work.

1. INTRODUCTION

In recent years, combination of pattern classifiers has shown very promising results by improving classification accuracies in complex data sets. These combination schemes are usually based on a strategy of combining different feature presentations from the same or different source signals, different classifiers for the same feature presentation or ensembles of weak learners [1].

Footstep identification is based on a biometric identification system where the classification tasks are complex multi-class problems. Therefore, it is useful to apply combination schemes to the process to achieve the best possible classification performance. In biometric identification and verification systems, for example, different sources (e.g. face and fingerprint [2]), different feature presentations from the same source [3] or different classifiers for the same or different feature sets can be combined [4]. Furthermore, biometric identification systems usually provide a possibility to use multiple samples from the same person to improve reliability [5, 6] and even allow their fusion with multi-source data [7].

In this paper, we combine both different classifiers for a single sample and multiple samples from the same source to achieve a reliable walker identification system based on footstep profiles from a pressure-sensitive floor. At the first stage of the multiple classifier system, three classifiers with unique feature presentations are applied. The different feature sets are calculated from the geometric and amplitude spectrum properties of the signal as well as the amplitude spectrum of its derivative. As an output, each classifier produces posterior probabilities to each known class. These partially independent probabilities are then combined with a simple product rule as an input to the second stage.

At the second stage of the identification system, the knowledge of multiple consecutive footstep profiles from the floor is used. This stage fuses the posteriors of the combined feature spaces of single footsteps by utilizing a sum rule and, finally, by choosing the maximum probability of the known classes. The use of the sum rule is derived from the assumption that samples recorded from the same person are highly correlated in nature. The methodology described in this paper improves the reliability of the identification system considerably.

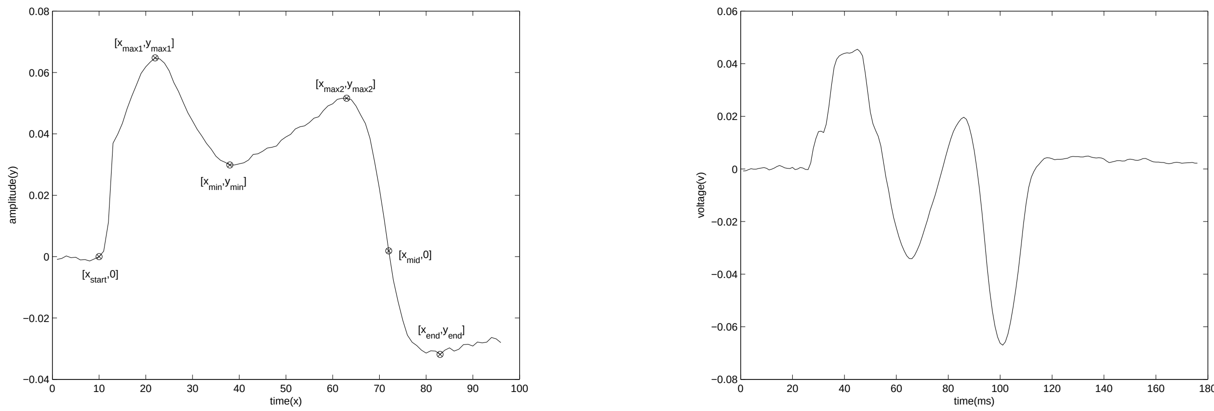
The rest of the paper is organized as follows. In section 2 the environment and the source signal are briefly introduced. Section 3 presents the building blocks of the multiple classifier system and section 4 the system as a whole. The results are reported in section 5, and the paper is concluded in section 6.

2. FOOTSTEP PATTERNS

The source signal of our system is achieved using a pressure-sensitive floor. It is based on ElectroMechanical film (EMFi) material, which provides a voltage signal when an external force makes an impact on its surface. Our floor system consists of 64 long, 30 cm wide sensor stripes, which make up a 30x34 matrix, where the cell size is 30x30 cm. The stripes were installed under the normal flooring material, providing an area of 100 square meters for the measurements. The details of the sensor material are presented in [8] and our system in, for instance, [9].

In our person identifier, single footstep profiles segmented from raw data are used as input patterns. In Figure 1(a), an example of the segmented footstep pattern of a walking person is shown. The voltage signal consists of two clearly observable local peaks resulting from the heel strike and the toe push-off. According to this, the profile is often named as *camel-back curve* and used as a basis for feature extraction. In addition, another presentation is constructed, containing a derivative of the input signal (see Fig.1(b)). This type of footstep signal has been achieved from, for example, a piezo force sensor measuring *Ground Reaction Force* [10]. In our experiments, the signal is numerically derived from the original profile using convolution with a differential mask. The aim of this feature presentation is to provide different, more dynamic, characteristics

This work was supported by TEKES and Academy of Finland.



(a) Footstep signal

(b) The derivate of footstep signal

Fig. 1. An example of the footstep profiles of (a) a signal and (b) the derivative of the signal.

of the footstep pattern to a classification scheme.

3. METHODS FOR A CLASSIFIER SYSTEM

3.1. Feature Extraction

The multiple classifier system is based on training classifiers with different feature sets. As mentioned in the previous section, two different presentations of input signals are used: the direct pressure signal and its derivative. Three different feature sets are then extracted from the presentations. The first set is based on the geometric (spatial) properties of the input signal. These features include the main coordinate points and relationals between them, as reported in our previous work [11] (see also Fig.1(a)). The second set contains a frequency domain presentation, calculated from *camel-back curve*, and the third set was constructed from the frequency domain of the derivative signal. Here, the amplitude spectrum of frequency domain presentations were used as feature sets, calculated by Fast Fourier Transformation (FFT).

These two amplitude spectrum presentations have high dimensionality and many correlated features. The statistical unsupervised feature extraction method called Principal Component Analysis (PCA) [12] was applied to decrease dimensionality and to find the most discriminative data projections. The task is to map a dataset of the vectors x_n for $n = 1, \ldots, N$ in $\mathbf{V} = \mathbb{R}^d$ to the vectors z_n in $\mathbf{U} = \mathbb{R}^k$, so that $k < d$, and to preserve as much data variance as possible. In other words, the original data set is rotated to the direction of maximum variance, where correlated high-dimensional data can be presented in a low-dimensional uncorrelated feature space with a small number of principal components (i.e., eigenvectors). PCA can be calculated by, for example, using the eigendecomposition of the input sample covariance matrix.

3.2. Classifier Design

In these experiments, we used two different pattern classifiers, which allows all the three feature sets to be presented with the same kind of classification method at once. The first method is Learning Vector Quantization (LVQ) [13], which was successfully applied in footstep identification in the authors' previous works [9, 14]. LVQ is a simple distance-based classifier, where a finite set of labeled prototype vectors are trained in a given feature space to approximate class distributions. An unknown sample is classified to the closest prototype vector (1-nearest neighbor rule [1-NN]) using some kind of distance metric (e.g., Euclidean distance). If class distributions are presented with a large number of prototype vectors, the k-nearest neighbor rule (k-NN) can also be applied. In our earlier studies, a small number of prototype vectors are used to present class distributions, and the final classification is made by the 1-NN rule [9].

When the k-NN rule is used in LVQ classification, the approximation of conditional posterior probabilities can be directly estimated from the occurrences of different classes in a k-NN set. However, when using a small number of nearest neighbors or 1-NN, the estimation criterion mentioned above is not suitable due to the limited range of outcomes [15]. Therefore, the posterior probability estimation can be based on distance calculations [15] as follows,

$$P(\omega_i|x) = 1 - \frac{(c-1)d_i(x)}{\sum_{j=1}^{c} d_j(x)}, \tag{1}$$

where $P(\omega_i|x)$ is the posterior probability of class ω_i when given a sample x, and d_i is the distance to the closest prototype vector in the class i, which is scaled with the sum of distances between the input sample and the closest prototype vectors in every class. The number of known classes is c. The probabilities are normalized so that the sum of all elements is one.

The second classifier applied to our experiments is a traditional feed-forward Multi-Layer Perceptron (MLP) neural network [12], which is trained using a backpropagation algorithm. An MLP classifier can directly estimate the conditional posterior probabilities, when the softmax activation function is used in the output layer [12]. The architecture of MLP consisted of one hidden layer with sigmoid activation functions.

3.3. Combination Strategies

The different feature presentations introduce different areas of expertise into the classification process. It has been shown in the literature [16] that when using uncorrelated and independent feature sets, a product combination rule is a good choice. Let R be the number of independent classifiers and ω ($\omega = \omega_1 \ldots \omega_n$) the known n classes. When every classifier produces conditional output probabilities $P(\omega_k|x_i)$, $k = 1 \ldots n$, according to the feature vector x_i, the product combination rule to assign an input sample to the class ω_c is presented as follows,

$$\omega_c = \underset{k=1}{\overset{n}{\arg\max}} \Big[\prod_{i=1}^{R} P(\omega_k|x_i) \Big], \tag{2}$$

where the final decision is made according to the maximum of combined values.

On the other hand, it has been found out that the summing/ averaging strategy works well when correlated outputs are used [16, 5]. In our case, multiple consecutive footstep profiles (S) recorded from the same person are assumed to correlate very closely, so that the sum combination rule can be applied to it correspondingly,

$$\omega_c = \underset{k=1}{\overset{n}{\arg\max}} \Big[\sum_{i=1}^{S} P(\omega_k|x_i) \Big]. \tag{3}$$

4. MULTIPLE CLASSIFIER SYSTEM

The multiple classifier system presented in this paper consists of two different combination stages. At the first stage, classification is applied to a single sample (i.e., footstep profile) using three different feature sets. The output of classification at the first stage is a combination of three classifiers' posterior probability outputs using a product rule (eq. 2), as different feature presentations are assumed to be independent.

Correspondingly, at the second stage, the product output of multiple correlated footsteps from a single person are combined using the sum rule (eq. 3). When these two combination rules are used one after another, the decision of the multiple classifier system can be presented as follows,

$$\omega_c = \underset{k=1}{\overset{n}{\arg\max}} \Big\{ \sum_{j=1}^{S} \Big[\prod_{i=1}^{R} P(\omega_k|x_{ij}) \Big] \Big\}, \tag{4}$$

where $P(\omega_k|x_{ij})$ is output probability of k:th class according to the i:th feature presentation of j successive input samples x_{ij}. R is the number of different classifiers at the first stage, S is the number of consecutive input samples used in the combination, and n is the number of known classes. The final decision is made similarly to the equations 2 and 3.

The example of the multiple classifier system architecture for using three consecutive footsteps is shown in Figure 2. However, the number of input samples is not restricted to three, if more of them can be used.

5. RESULTS

In these experiments, data from 11 different walkers are used, consisting of 40 segmented footsteps/person, as reported in [9]. Three different feature sets were extracted from each footstep profile. The first feature set of geometric properties contained 23 spatial features based on the extreme points of the profile (see Fig.1). Those were the 23 best features reported in [11]. The frequency domain presentations were calculated using 64-point FFT, and dimensionality was reduced with PCA. Finally, the 15 largest principal components were chosen, capturing the most of the variance.

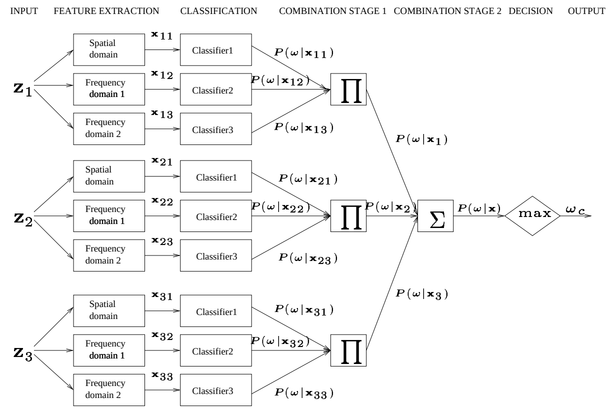


Fig. 2. An schematic example of an identification system using three input samples. At the first stage of combination, the product rule is used to combine the posterior output probabilities of classifiers, and then three consecutive samples are combined using a sum rule at the second stage.

The data set was divided into the training and test sets, so that 2/3 of the data were in the training set and the rest in the test set. The LVQ classifiers contained 18 prototype vectors / class and were trained with the program package LVQ_PAK [17]. The MLPs were implemented in Matlab using the Netlab [18]. The neural net trained with spatial features consisted of 20 hidden neurons, and both of the nets for the frequency domain features were utilized with 15 hidden neurons. The MLPs were trained using the scaled conjugated gradient optimization method. Both methods were tested using ten randomly chosen data sets, and the results are presented as average success rates and standard deviations.

The total recognition rates of single footsteps are presented in Table 1. The product rule can increase accuracies in both methods compared to different single features presentations. The combination results are more accurate than the single-feature presentations alone in both cases.

Finally, the results of the whole identification system are shown in Table 2. Both methods show very reliable recognition rates when the number of consecutive input samples is increased. For example, when using three consecutive footsteps, the system shows 91.2% and 92.4% success rates, which are very reliable compared to our earlier work [14], where a 90.0% success rate was achieved when rejecting 20.0% of input samples.

Feature Set	LVQ (%)	MLP (%)
SP	67.7 (4.9)	72.6 (3.4)
FR1	48.5 (3.7)	55.8 (4.8)
FR2	55.6 (6.2)	61.6 (4.6)
product	74.8 (8.8)	79.2 (7.5)

Table 1. The recognition accuracies of different single footstep feature presentations and the combination of the presentations. The first three rows present the total recognition rates of the spatial domain presentation (SP), the frequency domain presentation of the input signal (FR1) and the frequency domain presentation of the input signal derivative (FR2). The last row shows the combined recognition rates by the product rule.

No. samples	1	2	3	4	5	6	7	8	9
LVQ (%)	74.8 (8.8)	86.1 (5.8)	91.2 (6.1)	93.6 (3.9)	94.6 (4.7)	95.0 (4.5)	95.5 (4.3)	97.3 (4.4)	97.3 (4.3)
MLP (%)	79.2 (7.5)	89.0 (4.4)	92.4 (4.6)	92.4 (6.3)	95.0 (4.5)	95.0 (5.0)	95.9 (5.0)	96.8 (6.1)	98.2 (3.8)

Table 2. The classification accuracies of the multiple classifier system. The different feature presentations are combined by product rule at the first stage. The table presents the total recognition rates using multiple consecutive footsteps and a sum rule to combine them.

6. CONCLUSIONS

In this paper, it was shown that the combination of classifiers and multiple samples can improve the performance of footstep profile-based person identification. The combination scheme is based on a two-stage multiple classifier system, which includes a combination of different feature presentations for single footsteps at the first stage and then combines the knowledge of multiple input samples at the second stage. The combination strategies are derived from conditional posterior probability outputs of classifiers at both stages, using simple product and sum rules. The results show improvements of person identification compared to the authors' previous work.

Naturally, the next goal is to develop adaptive machine learning methodologies, which can automatically and incrementally evolve to find new classes (i.e. persons) and to adapt to changes in the behavior of persons already known to the system. Therefore, the analysis of how different feature presentations are affected by changes of the occupants' behavior (e.g., using different shoes, changes in walking style in general) will be studied. Also, more analysis will be needed to verify the best single classifiers and the combination strategies for an adaptive system.

7. REFERENCES

[1] R.P.W. Duin and D.M.J. Tax, "Experiments with classifier combining rules," in *Proc. of First International Workshop on Multiple Classifier Systems (MCS 2000)*, J. Kittler and F. Roli, Eds., Cagliari, Italy, June 2000, vol. 1857 of *Lecture Notes in Computer Science*, pp. 16–29, Springer-Verlag, Berlin.

[2] A. Ross and A. K. Jain, "Information fusion in biometrics," *Pattern Recognition Letters*, vol. 24, no. 13, pp. 2115–2125, September 2003.

[3] A. K. Jain, S. Prabhakar, and S. Chen, "Combining multiple matchers for a high security fingerprint verification system," *Pattern Recognition Letters*, vol. 30, no. 11-13, pp. 1371–1379, 1999.

[4] J. Kittler, M. Hatef, R. P. W. Duin, and J. Matas, "On combining classifiers," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 3, pp. 226–239, March 1998.

[5] J. Kittler, J. Matas, K. Jonsson, and M.U. Ramos Sanchez, "Combining evidence in personal identity verification systems," *Pattern Recognition Letters*, vol. 18, pp. 845–852, 1997.

[6] M.C. Cheung, M.W. Mak, and S.Y. Kung, "Multi-sample data-dependent fusion of sorted score sequences for biometric authentication," in *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP04)*, Montreal, Canada, May 2004.

[7] N. Poh, S. Bengio, and J. Korczak, "A multi-sample multi-source model for biometric authentication," in *Proc. of IEEE 12th Workshop on Neural Networks for Signal Processing (NNSP)*, 2002, pp. 375–384.

[8] M. Paajanen, J. Lekkala, and K. Kirjavainen, "Electromechanical film (emfi) - a new multipurpose electret material," *Sensors and actuators A*, vol. 84, pp. 95–102, 2000.

[9] S. Pirttikangas, J. Suutala, J. Riekki, and J. Röning, "Learning vector quantization in footstep identification," in *Proceedings of 3rd IASTED International Conference on Artificial Intelligence and Applications (AIA 2003)*, M.H. Hamza, Ed., Benalmadena, Spain, September 8-10 2003, IASTED, pp. 413–417, ACTA Press.

[10] P. Cattin, *Biometric Authentication System Using Human Gait*, Ph.D. thesis, ETH-Zürich, Institute of Robotics, Switzerland, 2002.

[11] J. Suutala and J. Röning, "Towards the adaptive identification of walkers: Automated feature selection of footsteps using distinction-sensitive LVQ," in *Proceedings of International Workshop on Processing Sensory Information for Proactive Systems (PSIPS 2004)*, Oulu, Finland, June 14-15 2004, pp. 61–67.

[12] C. M. Bishop, *Neural Networks for Pattern Recognition*, Oxford University Press Inc., New York, 1995.

[13] T. Kohonen, *Self-organizing Maps (3rd edition)*, Springer-Verlag, Berlin, Heidelberg, New York, 2001.

[14] J. Suutala, S. Pirttikangas, J. Riekki, and J. Röning, "Reject-optional LVQ-based two-level classifier to improve reliability in footstep identification," in *Proceedings of 2nd International Conference on Pervasive Computing (PERVASIVE 2004)*, A. Ferscha and F. Mattern, Eds., Linz/Vienna, Austria, April 18-23 2004, vol. 3001 of *Lecture Notes in Computer Science*, pp. 182–187, Springer-Verlag.

[15] R.P.W. Duin and D.M.J. Tax, "Classifier conditional posterior probabilities," in *Advances in Pattern Recognition, Proc. of Joint IAPR Int. Workshops SSPR'98 and SPR'98*, A. Amin, D. Dori, P. Pudil, and H. Freeman, Eds., Sydney, Australia, August 1998, vol. 1451 of *Lecture Notes in Computer Science*, pp. 611–619, Springer-Verlag.

[16] D.M.J. Tax, M. van Breukelen, R.P.W. Duin, and J. Kittler, "Combining classifiers by averaging or by multiplying?," *Pattern Recognition*, vol. 33, pp. 1475–1485, 2000.

[17] T. Kohonen, J. Hynninen, J. Kangas, J. Laaksonen, and K. Torkkola, "LVQ_PAK: The learning vector quantization program package," Tech. Rep., Helsinki University of Technology, Finland, 1996, Report A30.

[18] I. Nabney, *Netlab: Algorithms for Pattern Recognition*, Springer-Verlag, 2004, Software available: http://www.ncrg.aston.uk/netlab/.