Activity Related Biometrics based on motion trajectories

Anastasios Drosou[1,2], Konstantinos Moustakas[2], Dimos Ioannidis[2], Dimitrios Tzovaras[2]

[1]Imperial College London, SW7 2AZ London, UK {a.drosou09@imperial.ac.uk}

[2]CE.R.T.H. - I.T.I., P.O. Box 361 57001 Thermi-Thessaloniki, Greece

Abstract: The current paper contributes to the concept of activity-related biometric authentication in ambient Intelligence environments. The motivation behind the proposed approach derives from activity-related biometrics and is mainly focusing on everyday activities. The activity sequence is captured by a stereoscopic camera and the resulting 2.5D data are processed to extract valuable unobtrusive activity-related features. The novel contribution of the current work lies in the warping of the extracted movements trajectories, so as to compensate for different environmental settings. Authentication is performed utilizing both HMM and GMMs. The authentication results performed on a database with 32 subjects show that the current work outperforms existing approaches especially in the case of non-interaction restricting scenarios.

1 INTRODUCTION

Biometrics have recently gained significant attention from researchers, while they have been rapidly developed for various commercial applications ranging from surveillance and access control against potential impostors to smart interfaces. These systems require reliable personal recognition schemes to either confirm or determine the identity of an individual requesting their services. A number of approaches have been described in the past to satisfy the different requirements of each application such as unobtrusiveness, reliability, permanence, etc. Biometric methods are categorized to physiological and behavioral [JRP04], depending on the type of used features.

Behavioral biometrics, are related to specific actions and the way that each person executes them. They can potentially allow the non-stop (on-the-move) authentication or even identification in an unobtrusive and transparent manner to the subject and become part of an ambient intelligence environment. Behavioral biometrics are the newest technology in biometrics and they have yet to be researched in detail. They are supposed to be less reliable than physiological biometrics, however they are less obtrusive and simpler to implement [JRP04].

Recent work and efforts on human recognition have shown that the human behavior (e.g. extraction of facial dynamics features[HPL07]) and motion (e.g. human body shape dynamics during gait [ITD$^+$07]), when considering activity-related signals, provide the potential of continuous authentication for discriminating people .

Moreover, prehension biometrics belong to the general category of behavioral biometrics and can also been thought as a specialization of activity related biometrics [KCC02]. Activity related signals have exhibited the potential to accurately discriminate between subjects, while they remain stable for the same subject over time [KCC02]. Moreover, they are targeting to the user convenience (unobtrusiveness) as well as to an optimal performance in various realistic environments (invariance).

In this concept, an interesting biometric characteristic can be the user's response to specific stimuli within the framework of an ambient intelligence (AmI) environment. The present paper extends the applicability of activity-related biometric traits [DMIT10] and investigates their feasibility in user authentication applications. Specifically, it deals with the major problem of small variances in the interaction setting, which are introduced by the arbitrary positioning of the environmental objects, in respect to the user, at each trial. Thus, a generic approach for coping with this issue is attempted through the utilization of a warping algorithm, whereby the the behavioural information of the movement is not affected at all.

The overall workflow of the system follows: The user is expected to act with no constraints in an ambient intelligence environment. Meanwhile, some events, such as the ringing of the telephone, the need for typing of a password to a panel or even an instant message for online chatting, trigger specific reaction from the user. The users movements are recorded by one stereo camera and the raw captured images are processed, in order to track the users head and hands.

Then, the feature extraction step follows, where the user-specific trajectories are processed. The proposed biometric system supports two modes: a)The enrollment mode, whereby a user is registered through the training of a Hidden Markov Model or a Gaussian Mixture Model. b)The authentication mode, where the HMMs or the GMMs evaluate the claimed ID request by the user, as valid or void.

The proposed algorithm has been tested and evaluated in a large proprietary database and considerable improvements in recognition performance are seen in comparison to the state-of-the-art methods. The effectiveness of the proposed system is demonstrated by two experiments, where the potential of a person authentication using the biometric signature of just one activity as well as the combination of two separate activities activities will be examined. Additionally, the comparison results between the two proposed statistical methods (HMM - GMM) are presented.

2 ACTIVITY RELATED FEATURE EXTRACTION

Lacquaniti et. al. has proved in [LS82] that the motion pattern for a given movement is considered consistent from trial to trial and independent of the movement speed assumption. Thus, we can claim that the trajectories of each body part for a given activity can be seen as a biometric pattern.

Just like the approach suggested in [DMIT10], the estimated positions of the head and the palms on each frame are used to describe the movement performed by the user (Figure 1a). Before the actual feature extraction, a series of normalization operations are applied to the trajectories. Given the depth information provided by the the disparity image, it is easy to acquire the 3D data from the 2D image. The following trimming that is performed on these signals is a 3-step procedure [DMIT10]. Additionally, the homogenization of the extracted trajectories is further improved by resizing them to a fixed, a priori set vector length.

The acquired set of smooth trajectories (head & two palms) is shown in Figure 1b. These trajectories represent with high accuracy the movement of the corresponding body parts in XY axis, while the Z-axis (depth) is represented by the diameter of the circles.

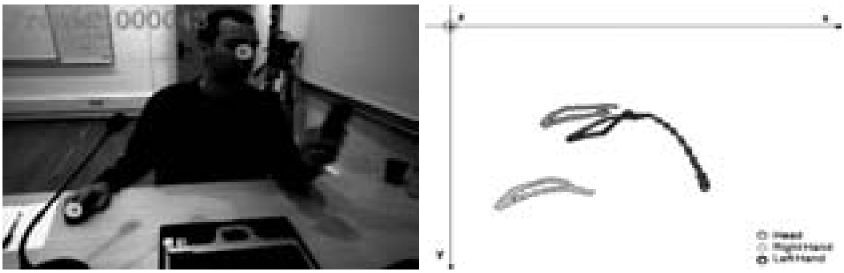


Figure 1: Set of trajectories describing a combined activity: Interaction with an office Panel during a phone conversation.

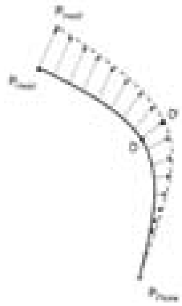
Spatial Warping A major issue is the invariance of the extracted trajectories even under small variations in the interaction setting between separate trials (different positions of the interaction objects) Normally, an increased False Rejection Rate (FRR) would apppear, due to variances in the interaction setting (object's positions, etc.) and not because the user not a genuine client for the system. In order to provide enhanced invariability to the extracted trajectories, the concept of *spatial warping* is introduced, inspired from the Dynamic Time Warping (DTW) method [SC90].

Without loss of generality regarding the environmental object, the case of a user answering successive phone calls will be studied. In the sequel, the relative distance between the user and the phone is not expected to remain fixed, either due to a shift of the user's body or due to small displacements of the phone. The same method can be applied to any environmental object with which the user is expected to interact (i.e. mouse, keyboard, pencil, book, etc.).

In a regular short phone conversation, there are two "extreme" positions of the hand that holds the telephone. Specifically, these can be seen in Figure 2 at point P_{Phone}, when the user has just grasped the phone just before he picks it up and at point P_{Head}, when the phone has touched the user's ear. The distance between these two "extreme" spots may vary even between the same user from trial to trial, since it depends on the slight variations of the environmental setting. Nevertheless, since we are mainly interested in the motion pattern of the trajectories and not its size, we apply the warping method on the hand trajectories.

Specifically, the exact location of these two points in the 3D space is automatically stored in the database for each user during the enrollment procedure. At the authentication step is warped according to the environmental characteristics of the enrollment moment.

In other words, the head-to-phone distance d is used for the deformation of an incoming set of trajectories, according to the claimed ID. Specifically, the blue line in Figure 2 indicates



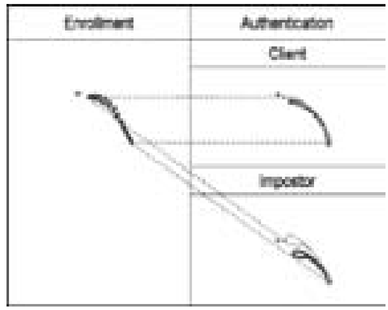


Figure 2: Normalization method for adjusting small variances in the head-to-phone distance.

the actual extracted trajectory in the authentication stage. $P_{Head'}$ and P_{Phone} are the stored locations of the user's head and the phone respectively obtained in the enrollment phase. The suggested method indicates that P_{Head} and $P_{Head'}$ as well as $P_{Phone'}$ and P_{Phone} are aligned, while all other points P_D of the XYZ signature in between are linearly transformed to the new point $P_{D'}$ as following:

$$P_{D'} = s(d)P(D) \tag{1}$$

where $s(d) = (P_{Head} - P_{Head'})d$ and $d = \sqrt{P_{Head}^2 - P_{Phone}^2}$.

In Figure 2b it can be noticed that the application of this method will the deformation (stretching/compression) of the trajectories to a common length. The advances of the suggested method towards enhanced invariancy can clearly be noticed in the ROC-curve diagrams, presented in Section 4.

3 Clustering & Classification

The training and verification step of our method has been tested under two statistical models. Namely, the Hidden Markov models and the Gaussian Mixture models have been mobilized and investigated, so as to perform authentication in the context of the proposed framework.

At the enrollment stage, both Hidden Markov Models (HMMs) and Gaussian Mixture Models (GMMs) are capable of clustering several sets $x_k(l_{head}, l_{rHand}, l_{lHand})$ of trajectories into a user specific signature. The Baum-Welch algorithm is utilized in the HMM case, while a weighted linear combination of $M = 5$ unimodal Gaussian densities and the Expectation-Maximization (EM) form the GMM. At the classification stage, on the other hand, the extracted set of trajectories is compared to the corresponding claimed signature, utilizing of the maximum likelihood criterion (HMM case) or Equation 2 (GMM case) and is finally classified as an impostor/client trajectory.

$$L(x|GMM_{n,k}) = \sum_{i=1}^{M} \log \sum_{j=1}^{L} \alpha_j \phi(j|\mu_j \Sigma_j) \tag{2}$$

whereby $M = 5$ denotes the five clusters of each GMM, α_j the weight factor of the j-th cluster, μ_j and Σ_j are the mean value and the variation of the distribution in the j-th cluster.

A user is only then authenticated, if the signature likelihood, returned by corresponding classifier, is bigger than an experimentally selected threshold.

4 EXPERIMENTAL RESULTS

The proposed methods were evaluated on the proprietary ACTIBIO-dataset [DMIT10], which consists of 29 regular subjects, performing a series of everyday office activities (i.e. a phone conversation, typing, talking to a microphone panel, drinking water, etc.) with no special protocol in 8 repetitions in total, equally split in two sessions.

Considerable improvements in the potential of recognition performance has been seen in comparison, after the application of the warping algorithm. Additionally, important outcomes have been extracted in terms of the applicability of the two utilized classifiers (Section 3) to the proposed system.

The proposed framework has been evaluated in the context of three verification scenarios. Specifically, the potential of the verification of a user has been tested, based on his a)activity-related signature during a *phone conversation*, b) activity-related signature during the *interaction with an office panel* and c) the *fusion at the score level* of the latter two activities. Specifically for scenario *c* the results from the activity *Phone Conversation* contributed with a factor of 0.2, while the the weight factor for the activity *Interaction with an Office Panel* is 0.8.

The evaluation of the proposed approach in an authentication scenario utilizes ROC-Curves and the corresponding equal error rates (EER) scores as shown in Table 1, whereby a noticeable improvement compared to original authentication capabilities of the system proposed in [DMIT10] can be seen.

Table 1: Authentication Performance - EERs.

EER	non-Warped			Warped	
	Phone	Panel	Fused	Phone	Fused
HMM	15%	9.8%	7.4%	12%	6.5%
GMM	25.2%	16.2%	15.3%	21.1%	14.9%

Given an HMM classifier, the original system could achieve an overall lowest EER of 7.4% in case there was a score level fusion in the results of two activities, namely a phone conversation and the user's interaction with the office panel. Each partial activity alone exhibited a much lower EER of about 15% and 9.8% respectively.

The evaluation of the warping method (Section 2) takes place in respect with the *phone conversation* activity, whereby the corresponding trajectory is warped for each user, according to the stored features of the claimed ID. There is a noticeable improvement with our approach, since there is a decreasing in the EER, of about 3%, compared to the simple approach. Given that the EER rate scores of the second activity remains untouched, the overall recognition performance of the system can achieve even lower EER scores, after tje score level fusion from these two separate activities.

On the other hand, the evaluation of the system, given a GMM classifier, can be characterized as rather worse in both cases (warped - nonWarped). Specifically, the overall EER lies at 15.3%, while each partial activities exhibit even lower EERs (phone conversation:25.2%, office panel:16.2%).

The reason for this deterioration in the authentication capabilities of the system, when a GMM classifier is used can be found in the fact that GMMs fail to calculate the transition probability within the given signal. A GMM can be viewed as a single-state HMM with a Gaussian mixture density. This proves to be useful in some cases, in terms of complexity, processing load and control over the statistical model, however, in this case the overall recognition performance deteriorates. Still, the proposed warping method has improved the results even with the GMM classifier, compared to the simple approach.

5 CONCLUSION & FUTURE WORK

In this paper we presented an extension to an unobtrusive authentication method that is related to activity-related biometrics and includes the dynamic characteristics derived when performing everyday activities, as a response to specific stimuli. The trajectories extracted from each user are warped towards more invariant activity related features, which are less dependent on the environmental setting but still retain the behavioural information. Moreover, the comparison between hidden Markov models and Gaussian mixture models towards user recognition exhibited the superiority of the first ones. This can be explained by the GMMs failure to take into account the transition probability between different states within the same model. Obviously, proposed method can achieve very high rates of authentication performance and therefore comprises a very interesting approach for further research in activity-related biometrics.

References

[DMIT10] A. Drosou, K. Moustakas, D. Ioannidis, and D. Tzovaras. On the potential ofactivity-related recognition. In *The International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISAPP 2010)*, 2010.

[HPL07] Abdenour Hadid, Matti Pietikäinen, and Stan Z. Li2. *Learning Personal Specific Facial Dynamics for Face Recognition from Videos*, pages 1–15. Springer Berlin / Heidelberg, 2007.

[ITD+07] D. Ioannidis, D. Tzovaras, I. G. Damousis, S. Argyropoulos, and K. Moustakas. Gait Recognition Using Compact Feature Extraction Transforms and Depth Information. *IEEE Trans.Inf. Forensics Security.*, 2(3):623–630, 2007.

[JRP04] A. K. Jain, A. Ross, and S. Prabhakar. An Introduction to Biometric Recognition. *IEEE Trans. Circuits Syst. Video Technol.*, 14(1):4–20, 2004.

[KCC02] A. Kale, N. Cuntoor, and R. Chellappa. A framework for activity-specific human identification. In *IEEE Proc. International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, volume 4, pages 3660–3663, 2002.

[LS82] F Lacquaniti and J F Soechting. Coordination of arm and wrist motion during a reaching task. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 2(4):399–408, April 1982.

[SC90] H. Sakoe and S. Chiba. Dynamic programming algorithm optimization for spoken word recognition. *Readings in speech recognition*, 1990.