Smartphone Fingerprint by Mixing Device Sensors Features

I. Amerini[♯], P .Bestagini[♭], L. Bondi[♭], R. Caldelli[♯ ⋄], M. Casini[♯], S. Tubaro[♭]

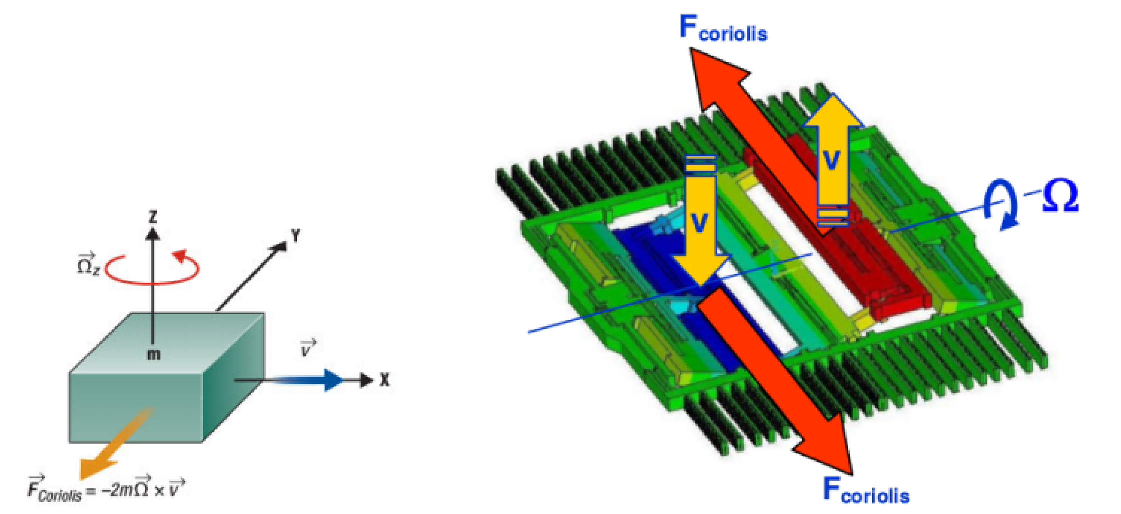
[♯] *Media Integration and Communication Center (MICC), University of Florence*
[♭] *Dipartimento di Elettronica, Informazione e Bioingegneria (DEIB), Politecnico di Milano*
[⋄] *National Interuniversity Consortium for Telecommunications (CNIT)*

Introduction

In the next coming years, many of our basic activities such as reading an e-mail, checking our bank account, buying on-line, etc., will be performed by using a smartphone in a mobile environment. It is quite obvious that the degree of security granted by a classic username-password access is not sufficient and that a stronger level of safeness is required.

We present a possible solution which envisages the use of the user's own smartphone as a mean to grant a safer and easy mobile access. A novel methodology is introduced to obtain a robust smartphone fingerprint by opportunely combining different intrinsic characteristics of each sensor. Modern mobile phones, in fact, have several kinds of sensors such as accelerometer, gyroscope and camera; such sensors can be used to uniquely identify each phone by measuring the specific anomalies left onto the signals they acquire.

Gyroscope



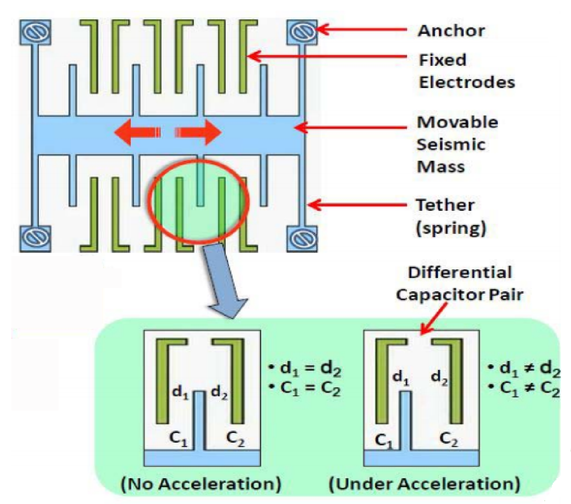
We read the output signals $\omega_x(k)$, $\omega_y(k)$, $\omega_z(k)$.

Features

We extract directly from readings a set of 21×3 features f_g in the temporal and frequency domain using the MIR Toolbox[a].

[a] https://www.jyu.fi/hum/laitokset/musiikki/en/research/coe/materials/mirtoolbox

Accelerometer



We read signals $a_x(k)$, $a_y(k)$ and $a_z(k)$ and the associated time-stamp $t_a(k)$.

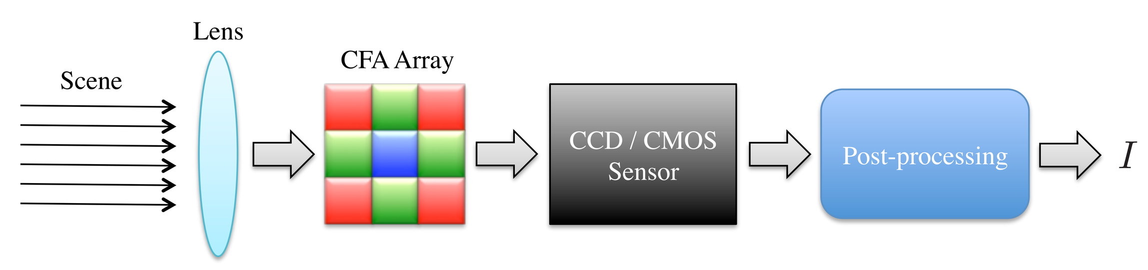
Features

We extract a set of 17×2 features[a] f_a in the temporal and frequency domain from

$$T(k) = t_a(k+1) - t_a(k),$$

$$S(k) = \sqrt{a_x^2(k) + a_y^2(k) + a_z^2(k)},$$

Camera



An image can be modelled as

$$I = I^{(0)} + I^{(0)}K + N,$$

where $I^{(0)}$ is a noiseless representation of the scene, N is an additive noise term, and K is the multiplicative PRNU.

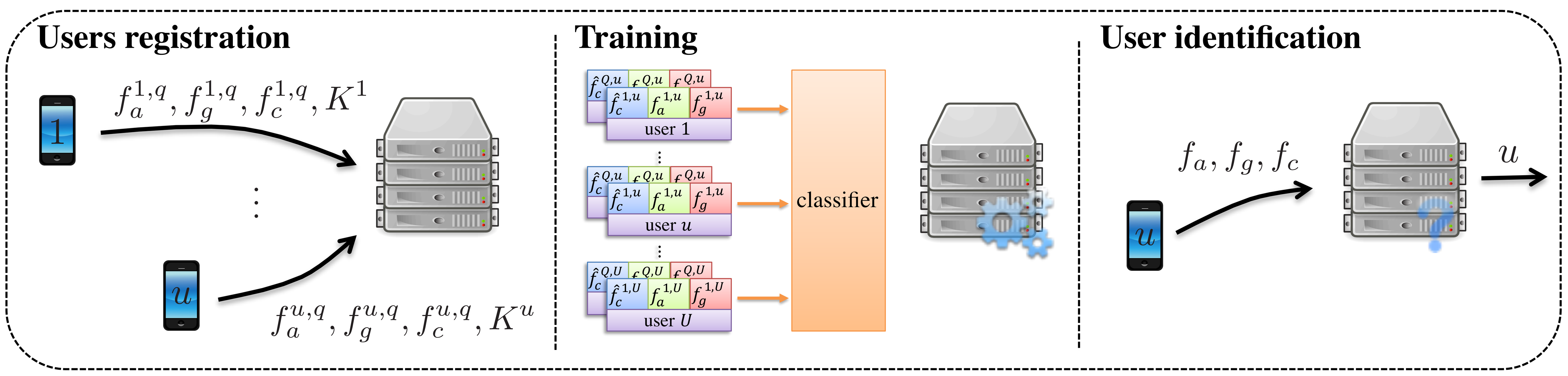
Features

Each camera is described by the vector

$$f_c = \mathrm{sign}(W_{512\times512}),$$

where $W_{512\times512}$ is the 512×512 central portion of the noise term W extracted from I.

Proposed system



PRNU estimation

$$K = \frac{\sum_p W_p I_p}{\sum_p I_p^2}$$

PRNU correlations

$$\hat{f}_c^{u,q} = \begin{bmatrix} \rho(f_c^{u,q}, K^1) \\ \rho(f_c^{u,q}, K^2) \\ \vdots \\ \rho(f_c^{u,q}, K^U) \end{bmatrix}$$

Results

Devices

Device	Amount
LG Nexus 5	5
Motorola Moto G 2015	1
Samsung Galaxy S3	2
Samsung Galaxy S4	1
Samsung Galaxy S2plus	1

Configurations

	Training set		Test set	
	Position	*Vibration*	*Position*	*Vibration*
CONF1	Table	ON	Table	ON
CONF2	Table	OFF	Table	OFF
CONF3	Hand-held	ON	Hand-held	ON
CONF4	Table	ON	Hand-held	ON
CONF5	Hand-held	ON	Table	ON