Classificazione di Immagini Digitali Attraverso Pattern di Rumore del Sensore della Fotocamera

Michael Orrù

Anno Accademico 2021/2022

Progetto Elaborazione di Immagini e Visione Artificiale

ABSTRACT

L'esperimento si propone di determinare se é possibile individuare un dispositivo, tra un gruppo, a partire da un'immagine scattata da esso. Si presuppone che siano disponibili altre immagini scattate dalle fotocamere in esame o che siano disponibili le fotocamere stesse. Il metodo di identificazione si basa sul calcolo del pattern di rumore della fotocamera, che é una caratteristica stocastica unica dei sensori ottici, presente in singole regioni dell'immagine. Il modello decisionale si basa sulla correlazione tra il pattern di rumore calcolato per ogni fotocamera e l'immagine cui si vuole scoprire il dispositivo che l'ha scattata. L'esperimento viene svolto su immagini scattate da dispositivi reali, inoltre, viene verificato in che modo, l'elabrazione delle immagini utilizzate, incide sulla capacitá di individuare il dispositivo. L'esperimento si basa sulla definizione di pattern di rumore fornita dall'articolo *Detecting Digital Image Forgeries Using Sensor Patten Noise* di Jan Lukas, Jessica Friedrich e Miroslav Goljan.

Sommario

1 Introduzione

La necessitá di individuare dispositivi attraverso immagini scattate digitalmente é sempre stata presente in tutti quegli ambiti dove é cruciale sincerarsi della natura di una immagine, ad esempio in tribunale, dove le immagini vengono presentate come prove.

1.1 Processo di acquisizione di un'immagine

Il nucleo di ogni fotocamera é il sensore ottico. Il sensore é diviso in piccoli elementi indirizzabili e minimi, i pixel, che raccolgono fotoni e li convertono in tensione che viene successivamente campionata in un segnale digitale attraverso un convertitore A/D. Prima che la luce della scena fotografata raggiunga il sensore, tuttavia, passa attraverso gli obbiettivi della fotocamera, un filtro antialising e attraverso un array di filtri colorati (CFA). L'array di filtri colorati blocca una determinata porzione dello spettro, consentedo a ciascun pixel di rilevare solo un colore specifico.

Se il sensore utilizza un CFA, l'output del sensore viene ulteriormente interpolato utilizzando algoritmi di interpolazione del colore per ottenere tutti e tre i colori di base in ciascun pixel. Il segnale ottenuto viene ulteriormente elaborato per correggere la distribuzione del colore e il bilanciamento del bianco. Infine, l'immagine viene salvata sulla memoria del dispositivo.

1.2 Il pattern di rumore

Le fonti di imperfezione che influenzano il processo di acquisizione dell'immagine, dovute alla fotocamera, sono principalmente due, una delle quali é l'obiettivo di studio di questo lavoro. Quando un'immagine, scattata in presenza di illuminazione assolutamente uniforme, presenta piccole variazioni di intensitá tra i singoli pixel, la colpa é da attribuire, in parte, a componenti casuali, come il rumore di lettura o di scatto, anche conosciuto come rumore fotonico, e in parte, al pattern di rumore, una componente deterministica che rimane approssimativamente uguale in immagini scattate dalla stessa fotocamera. Il pattern di rumore é un difetto fisico del sensore ottico ed é quindi presente in ogni immagine scattata dal sensore.

Le due componenti principali del pattern di rumore sono il fixed pattern noise (FPN) e il photo-response non-uniformity noise (PRNU). Il fixed pattern noise si riferisce alle differenze da pixel a pixel quando il sensore non é esposto alla luce, é anche conosciuto come corrente oscura. L'FPN é un rumore additivo e dipende dall'esposizione e dalla temperature.

Il photo-response non-uniformity noise é la parte dominante del pattern di rumore. Quando una quantitá fissa e uniforme di luce colpisce le celle del sensore (pixel) in una fotocamera digitale, ciascuna cella della fotocamera dovrebbe emettere esattamente la stessa tensione, ma in realtá questo non succede. Quando le celle del sensore vengono colpite da una luce uniforme, emettono tensioni leggermente diverse, questo fenomeno é causato dalle differenti grandezze dei pixel e dalle imperfezioni del materiale che ne forma il substrato. Questa differenza in risposta a una sorgente luminosa uniforme é denominata photo-response non-uniformity noise. Poiché il PRNU é causato dalle proprietá fisiche del sensore, é impossibile eliminarlo e lo si puó trovare in ogni immagine scattata dal sensore.

1.2.1 Photo-Response non-Uniformity Noise

Per comprendere meglio la natura del PRNU viene fornito il modello matematico che descrive il processo di acquisizione di un'immagine. L'output digitalizzato, non elaborato, del sensore $y = y_{ij}$ puó essere espresso nella seguente forma:

$$y_{ij} = f_{ij}(x_{ij} + \eta_{ij}) + c_{ij} + I_{ij}$$

Dove:

- $x = x_{ij}, i = 1, ..., m, j = 1, ..., n$ é il segnale grezzo acquisito dal sensore e $m \times n$ é la risoluzione del sensore;
- $\eta = \eta_{ij}$ é il rumore casuale;
- $I = I_{ij}$ é il rumore casuale additivo, come il rumore di lettura, di scatto, ecc.;
- $c = c_{ij}$ é la corrente oscura;
- I fattori f_{ij} sono prossimi a 1 e catturano il rumore PRNU, che é un rumore moltiplicativo;

Il rumore PRNU non é presente nelle aree completamente saturate di un'immagine, dove tutti i pixel del sensore sono stati riempiti a piena capacitá producendo un segnale costante. Nelle aree molto scure il PRNU é ampiamente soppresso.

2 Individuazione del patten di rumore e implementazione

Al modello decisionale che si occupa di classificare le immagini vengono forniti tutti i pattern di rumore dei sensori sotto esame e l'immagine che si vuole classificare. Pertanto l'esperimento

inizia con il calcolo del PRNU per ogni sensore. Viene quindi descritto il processo di valutazione.

Per una data fotocamera C, otteniamo un'approssimazione P_c al pattern di rumore di riferimento della fotocamera calcolando la media di piú immagini $p^{(k)}, k = 1, ..., N_p$. Questo processo puó essere accellerato sopprimendo il contenuto della scena dall'immagine prima della media. L'approssimazione al pattern di rumore sará ottenuta dalla formula:

$$mean(\sum_{k=0}^{n} n^{(k)}) = mean(\sum_{k=0}^{n} p^{(k)} - F(p^{(k)}))$$

dove F é un filtro di denoising. Si consiglia di utilizzare un numero di immagini superiore alle 50 unitá per ottenere una buona approssimazione. Nell'esperimento ne vengono utilizzate 80 per ogni sensore.

Viene fornita l'implementazione matlab della soluzione:

```
function[prnu] = estimate_PRNU(imgs)

    size_img = size(imgs{1});
    prnu = im2double(zeros(size_img(1), size_img(2)));

    for i = 1:length(imgs)
        img = im2double(im2gray(imgs{i}));
        denoised = wiener2(img, [5 5]); % denoising
        prnu = prnu + (img - denoised); % eliminazione contesto
    end

    prnu = prnu ./ length(imgs); % media

end
```

dove il denoising é effettuato attraverso un filtro di Wiener.

3 Natura del modello decisione

Per decidere se un'immagine p é compatibile con il pattern di rumore della telecamera C, calcoliamo la correlazione tra il rumore residuo $n = p - F(p)$ e il pattern di riferimento della telecamera P_c

$$\rho(n, P_c) = \frac{(n-\overline{n})\cdot(P_c-\overline{P_c})}{\|n-\overline{n}\|\|P_c-\overline{P_c}\|}$$

dove n e P_c sono vettori. La barra sopra i simboli denota il valor medio, "$\cdot$" é il prodotto scalare e $\|\cdot\|$ é la norma L2.
Si otterrá una correlazione discretamente alta per il sensore che ha effettivamente scattato l'immagine e si otterrá una correlazione vicina allo 0 per tutti gli altri sensori.

La soluzione implementata é la seguente:

```
function [label] = classify_from_PRNU(PCs, labels, img)

    img_g = im2double(im2gray(img));
    n = img_g - wiener2(img_g, [5 5]); % denoising / eliminazione contesto
    size_n = size(n);
    n = reshape(n, [1, size_n(1) * size_n(2)]);
    n_bar = mean(n, 'all');

    corrs = [];
    for i = 1:length(PCs)
        Pc = reshape(PCs{i}, [1, size_n(1) * size_n(2)]);
        Pc_bar = mean(Pc, 'all');
        correlation = dot(n - n_bar, Pc - Pc_bar) / ...
            (norm(n - n_bar, 2) * norm(Pc - Pc_bar, 2));
        corrs = [corrs, correlation];
    end

    [~, I] = max(corrs); % decisione
    label = labels(I);

end
```

4 Risultati

Per un totale di tre sensori, il database utilizzato per l'esperimento fornisce cento immagini naturali in formato jpg e trecento immagini elaborate da tre diversi algoritmi:

- Algoritmo FBH: utilizzato da Facebook per creare post ad alta risoluzione.
- Algoritmo FBL: utilizzato da Facebook per creare post a bassa risoluzione.
- Algoritmo WA: utilizzato da WhatsApp per l'invio di immagini.

Per ogni set di immagini, l'80% é stato utilizzato per calcolare l'approssimazione del pattern di rumore, il restante 20% é stato utilizzato come campione per la classificazione. La scelta dei campioni che compongono il training set e il test set é casuale.

Viene, prima di tutto, mostratun esempio di risultati ottenuti per il set di immagini naturali.

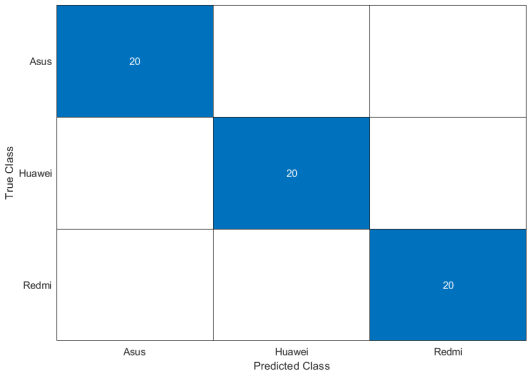


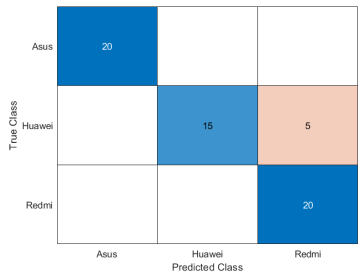
Figura 1: classificazione immagini naturali

```
"true label = Huawei"
correlation_asus = 0.0086
correlation_redmi = 0.0136
correlation_huawei = 0.0414

"true label = Asus"
correlation_asus = 0.1385
correlation_redmi = 0.0128
correlation_huawei = 0.0114

"true label = Redmi"
correlation_asus = 0.0101
correlation_redmi = 0.0405
correlation_huawei = 0.0156

"true label = Redmi"
correlation_asus = 0.0035
correlation_redmi = 0.0700
correlation_huawei = 0.0057
```

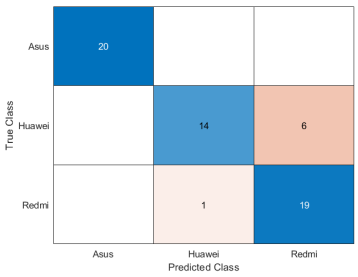
Figura 2: esempi di processo decisionale

Attraverso l'approssimazione ottenuta si é in grado di classificare ogni immagine correttamente, tuttavia, dagli esempi riportati é possibile osservare che il margine decisionale é abbastanza stretto. Sarebbe possibile allargare questo margine utilizzando piú immagini per calcolare l'approssimazione del pattern di rumore.

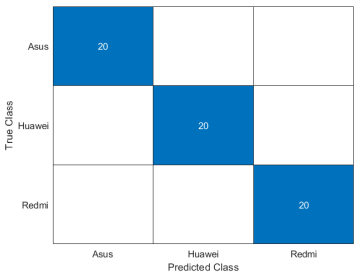
Vengono ora mostrati in ordine degli esempi di classificazione per i set di immagini elaborate con gli algoritmi: FBH, FBL e WA:




```
"true label = Redmi"
correlation_asus = 0.0383
correlation_redmi = 0.0574
correlation_huawei = 0.0430

"true label = Asus"
correlation_asus = 0.0641
correlation_redmi = 0.0385
correlation_huawei = 0.0365

"true label = Redmi"
correlation_asus = 0.0320
correlation_redmi = 0.0523
correlation_huawei = 0.0373

"true label = Asus"
correlation_asus = 0.0627
correlation_redmi = 0.0222
correlation_huawei = 0.0206
```




```
"true label = Redmi"
correlation_asus = 0.0170
correlation_redmi = 0.0236
correlation_huawei = 0.0186

"true label = Huawei"
correlation_asus = 0.0211
correlation_redmi = 0.0260
correlation_huawei = 0.0225

"true label = Huawei"
correlation_asus = 0.0462
correlation_redmi = 0.0503
correlation_huawei = 0.0466

"true label = Redmi"
correlation_asus = 0.0282
correlation_redmi = 0.0353
correlation_huawei = 0.0284
```




```
"true label = Huawei"
correlation_asus = 0.0309
correlation_redmi = 0.0353
correlation_huawei = 0.0393

"true label = Redmi"
correlation_asus = 0.0303
correlation_redmi = 0.0449
correlation_huawei = 0.0348

"true label = Asus"
correlation_asus = 0.0770
correlation_redmi = 0.0345
correlation_huawei = 0.0323

"true label = Redmi"
correlation_asus = 0.0184
correlation_redmi = 0.0435
correlation_huawei = 0.0230
```

Si possono osservare classificazioni in linea con quella del set di immagini naturali. Questo significa che l'approssimazione del pattern di rumore é ancora sufficientemente buona per permetterci di classificare correttamente (quasi) tutte le immagini. L'elaborazione ha rimosso solo in piccola parte l'informazione portata dal PRNU.

5 Conlusioni

Nell'esperimento eravamo in possesso di immagini etichettate, ovvero, eravamo a conoscenza di quale sensore avesse scattato quale immagine. Questo ci ha permesso di costruire delle confusion matrix e di comprendere quanto buona fosse la classificazione. Tuttavia, immaginiamo un caso reale, in cui il sensore delle immagini da classificare é sconosciuto. Un utente otterebbe solamente una lista di correlazioni tra immagine e PRNU dei sensori in esame (simile a quelle riportate precedentemente). Con risultati simili a quelli riportati per il primo set, l'incertezza é relativamente bassa, per risultati simili a tutti gli altri casi, il gap tra coppie di correlazione é troppo stretto per affermare, con sicurezza, che un'immagine sia stata effettivamente scattata da un dato sensore.
Possiamo concludere che il metodo illustrato é effettivamente in grado di selezionare il sensore che ha scattato una data immagine, tuttavia, lo fá con un grado di incertezza inversamente proporzionale al numero di immagini utilizzate per stimare il PRNU del sensore. In sviluppi futuri sarebbe interessante capire quante immagini siano necessarie per calcolare un'approssimazione del patter di rumore che dia un margine decisionale abbastanza largo da affermare, senza alcun dubbio, che la classificazione é corretta.

Riferimenti bibliografici

[JJM] Lukas Jan, Friedrich Jessica, and Goljan Miroslav. Detecting digital image forgeries using sensor pattern noise. *Department of Electrical and Computer Engineering.*

[SFI+17] Dasara Shullani, Marco Fontani, Massimo Iuliani, Omar Al Shaya, and Alessandro Piva. Vision: a video and image dataset for source identification. *EURASIP Journal on Information Security*, 2017(1):1–16, 2017.