ScuDo

Scuola di Dottorato ~ Doctoral School

WHAT YOU ARE, TAKES YOU FAR

XXXII Cycle

Camera Identification and Clustering from Images and Videos

Sahib Khan

Supervisor: Prof. Tiziano Bianchi

Research context and motivation

- Digital images have brought tremendous changes in human life
  - Documenting news
  - Sharing life events on social media
  - Providing evidence in the court of law
- At the same time, forensic analysis of images faces various problems
  - Source identification
  - Tampering detection
  - Grouping according to a common source
- Unique camera fingerprints can help solving the above-mentioned tasks
  - Each acquisition device leave unique intrinsic trances
  - Photo Response Non Uniformity (PRNU) is the more relevant among them
  - PRNU is unique, stable and multiplicative in nature

Fig 1. Camera fingerprint

Research Problems/Challenges and Objective

- **Problem:** clustering images according to the source device. The clustering done without any prior information about
  - The souce camera
  - Number of source cameras
  - Number of images captured with an individual camera
- **Challenges:**
  - High I/O and computational cost
  - Large memory requirements
  - Number of clusters (NC) >> size of clusters (SC)
- **Objective:** simple algorithm with reduced complexity
- **Side Project:** Along with camera identification and clustering, some work is done in the field of data hiding and an ACO based data hiding in edges technique is implemented at the start of the PhD studies.

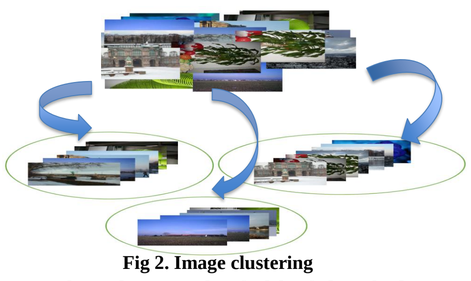


Fig 2. Image clustering

Basic Concepts

- **Fingerprint Estimation:** Estimate and Standardize caemra fingerprints
  - $M = \{F_i | F_i = \varphi(X_i - D(X_i)), i = 1, 2, \dots n\}$
  - $D(.)$ is the denoising function
  - $\varphi(.)$ normalize to zero mean and unit norm
- **NCC Computation:** The NCC between F_i and RF_k each of size d is
  - $\rho(i) = \frac{1}{d}\sum_{j=1}^{d} RF_k[j]F_i[j]$
- **Threshold Compute:** The threshold T using the desirable PFA is
  - $T = \sqrt{\frac{2}{d}}erfc^{-1}(2 \times PFA)$
- If ρ between F_i and RF_k is greater than or equal to T then they belong to the same camera

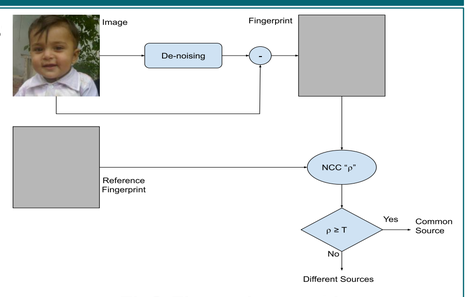


Fig 3. Fingerprints matching

Methodology

RCIC Algorithm

Initialize
- Set of unclustered fingerprints M
- k=1

Repeat
- Randomly select a fingerprint as reference RFk and assign it to cluster Ck
- For each unclustered fingerprint Fi
  - Compute NCC ρ between RFk and Fi
  - If ρ ≥ T, the fingerprint Fi is assigned to a cluster Ck, otherwise Fi is left unclustered
- k=k+1

Until all fingerprints are assigned to a cluster Ck

FICFO Algorithm

- **Ranking index computation**
  - $RI = G^{\frac{1}{\alpha}} \times (1-S)^{\frac{1}{\beta}} \times (1-T)^{\frac{1}{\gamma}}$
  - G is the average gray level, S is the saturation level, T is the texture. While, α, β and γ are the factors defining the contribution of G, S and T, respectively
- **Fingerprints ordering:** The fingerprint are arranged in the descending order of RI.
  - $M_o = SortDescend(M, RI)$
- **Clustering:**
  - Select the best fingerprint i.e. first one, as reference RFk.
  - Use the baseline clustering algorithm i.e. RCIC, to cluster the fingerprints.

Until all fingerprints are assigned to a cluster Ck

Attraction

- For each Ck and average reference fingerprint ARFk is computed by averaging all the fingerprints in Ck and normalizing the result to zero mean and unit norm
- k=1

Repeat
- Randomly select one non-merged ARFi as reference RFk
- For each non-merged ARFi
  - Compute NCC ρ between ARFk and Fi
  - If ρ ≥ T, merge ARFk and RFk clusters, other wise left ARFi non-merged
- k=k+1

Until all ARFk corresponding Ck are either treated as RFk or merged with some other cluster

CFIC Algorithm

- **Fingerprints Reduction:**
- $F_d = Decimate(F, \dot{d})$
- $F_{rp} = \Psi * F_d$
- $F_r = \begin{cases} +1 & if\ F_{rp}(i) > \delta\rho \\ 0 & if\ -\delta\rho \le F_{rp}(i) \le \delta\rho \\ -1 & if\ F_{rp}(i) < -\delta\rho \end{cases}$
- Sorting of fingerprints (F and F_r) in descending order of RI
- **Clustering:** Uses F_r and RCIC as a baseline to get Ck
- **Fine clustering:** Merged the full fingerprints of each cluster Ck and standardize the merged fingerprints to zero mean and unit norm.
- Refine the clusters using the process explained in attraction.

Novel contributions

- Image clustering on the basis of camera fingerprints.
- Complexity reduction.
- NC>> SC problem addressed.

Experimental results

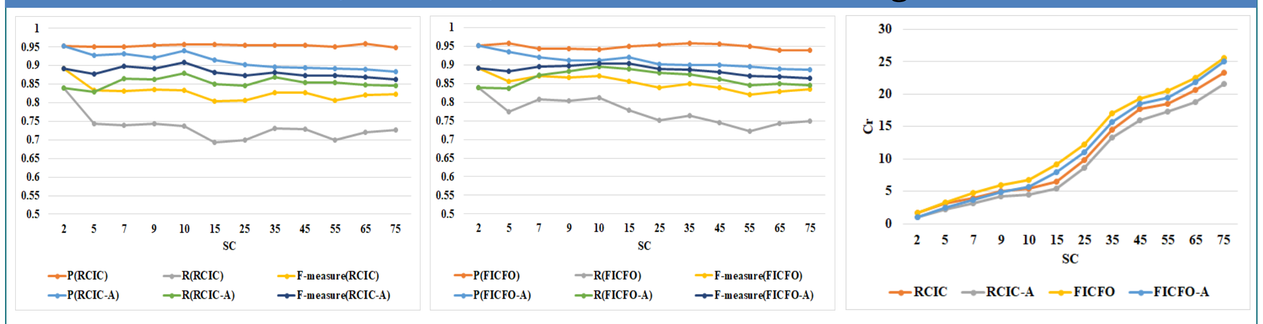
- Dresden, 53 cameras
- D2: 25 cameras, with 20, 30, 40, 50 and 60 Images
- D4: 50 cameras, with 10, 15, 20, 25 and 30 images
- All images are center cropped to 1023 x1023

- $P = \frac{\sum_k (max_j |c_k \cap \omega_j|)}{\sum_k |c_k|}$
- $F = 2 \times \frac{(P \times R)}{(P+R)}$

$R = \frac{\sum_j (max_k |c_k \cap \omega_j|)}{\sum_j |\omega_j|}$

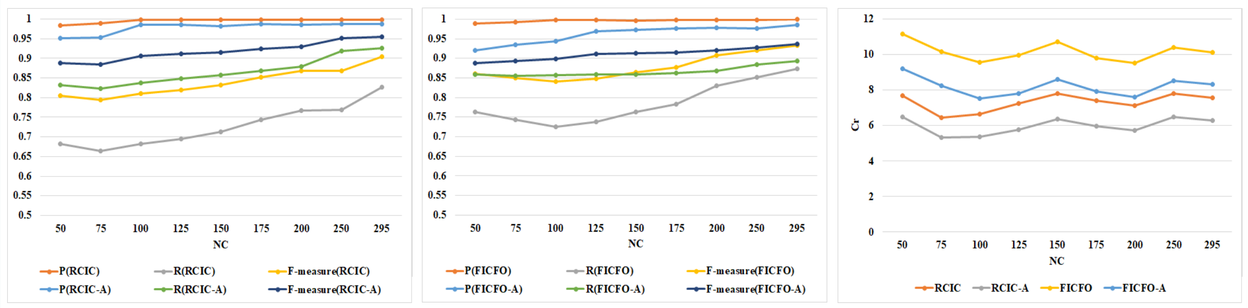
$Cr = \frac{n(n-1)}{2 \times tc}$

RCIC, RCIC-A, FICFO and FICFO-A Algorithm



(a) RCIC and RCIC-A (b) FICFO and FICFO-A (c) Cr

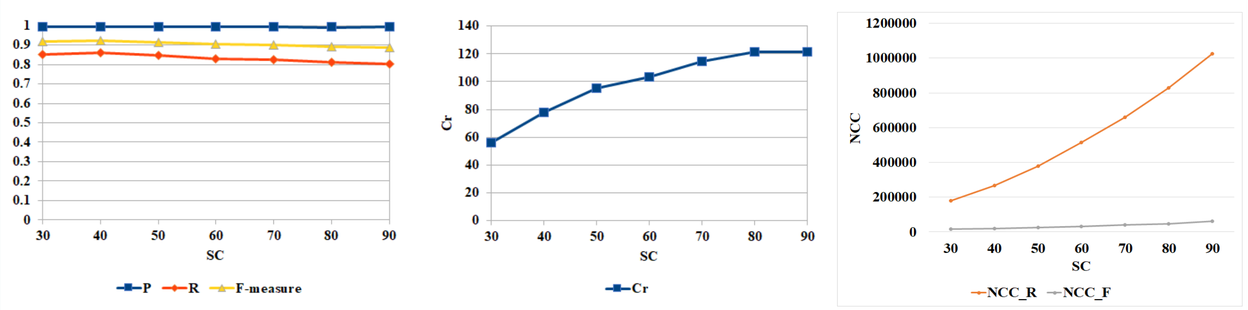
Fig 4. P, R, F-measure and Cr of RCIC, RCIC-A, FICFO and FICFO-A algorithms using natural images by varying SC and fixed NC=53



(a) RCIC and RCIC-A (b) FICFO and FICFO-A (c) Cr

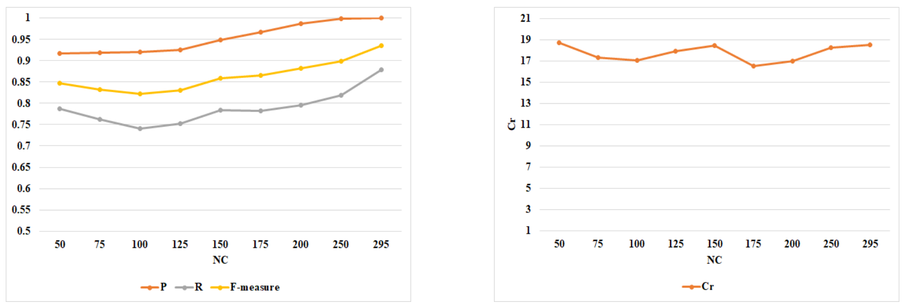
Fig 5. P, R, F-measure and Cr of RCIC, RCIC-A, FICFO and FICFO-A algorithms using varying NC and fixed SC=20

CFIC Algorithm



(a) P, R, F-measure (b) Cr (c) No. of NCC

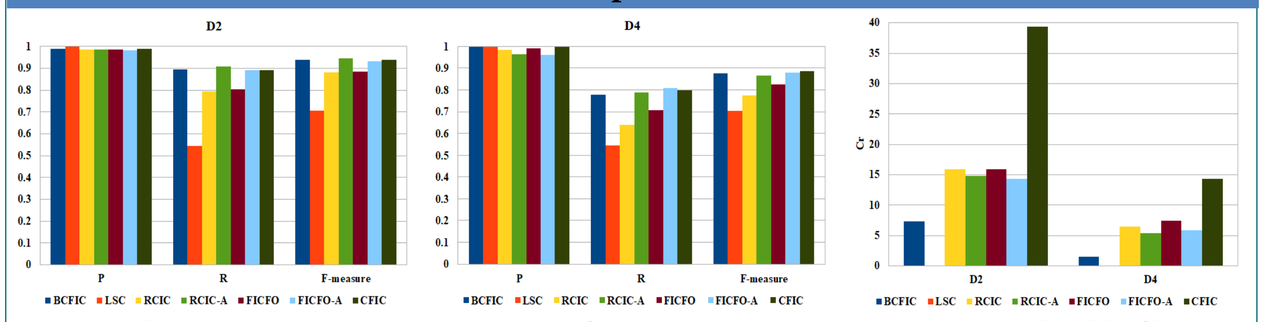
Fig 6. P, R, F-measure, Cr and No. of NCC of CFIC algorithm using natural images by varying SC and fixed NC=53



(a) P, R, F-measure (b) Cr

Fig 7. P, R, F-measure and Cr of CFIC algorithm using NC and fixed SC=20

Comparison



(a) D2 (b) D4 (e) Cr on D2 and D4

Fig 8. Comparison of image clustering algorithms

The BCFIC and LSC are state-of-the-art image clustering algorithms

Submitted and published works

- Khan, S. and Bianchi, T. *Image Clustering on the Basis of Compressed Fingerprints*. (In drafting).
- Khan, S. and Bianchi, T. (May, 2019). *Reduced Complexity Image Clustering Based on Camera Fingerprints*. International Conference on Acoustics, Speech and Signal Processing (ICASSP) , pp. 2682-2688. IEEE.
- Khan, S. and Bianchi, T. (July, 2019). *Fast Image Clustering Based on Camera Fingerprints Ordering*. International Conference on Multimedia and Expo (ICME), pp. 766-771. IEEE.
- Khan, S. and Bianchi, T., "*Ant Colony Optimization (ACO) based Data Hiding in Image Complex Region*", International Journal of Electrical & Computer Engineering, vol. 8, no. 1, 2018, pp. 379-389.

List of attended classes

- 02RHPRV – (Nov 2016-Feb 2017, 6 credits)
- 01RQXRV – (Dec 2016- Feb 2017, 8 credits)
- 01RQORQ – (Dec 2016, 3 credits)
- 01PJMRV – (Jan-Feb 2017, 4 credits)
- 01QSAIU – (Jan-Mar 2017, 4 credits)
- 01QRXIU – (Feb-Mar 2017, 4 credits)
- 01QRQRV – (Feb-March 2017, 4 credits)
- 01QTEIU – (Feb 2017, 4 credits)
- 01QFFRV – (Feb 2017, 4 credits)
- 01QRRRV – (May-June 2017, 4 credits)
- 01SHCRV – (Nov 2017-Feb 2018, 6 credits)
- 01NDLRV – (Sep 2018, 3 credits)

POLITECNICO DI TORINO

PhD program in

Electrical, Electronics and Communications Engineering