ROBUST IMAGE IDENTIFICATION WITH SECURE FEATURES FOR JPEG IMAGES

Kenta Iida, Hitoshi Kiya

Tokyo Metropolitan University, Japan

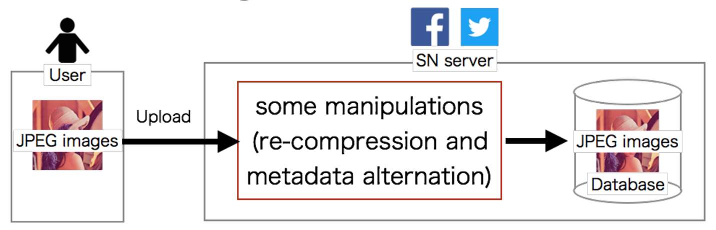
Paper ID : 2499

Summary

- Our proposed scheme can identify robustly images under various coding conditions.
- Quantization matrices and positions in which DCT coefficients have zero values are used as features.
- The features do not provide no visible information.
- The property of DCT coefficients and the features allow us to provide no false negative matches.
- Simulation results demonstrate the effectiveness of the proposed scheme.

Background

- What is "image identification"?
  ⇒Identification of images which are generated from **the same original image** under various coding conditions.
- Why is image identification required?
  ⇒Social networks(SNs) providers often **re-compress** uploaded JPEG images with **the different coding parameter** from that of uploaded one.

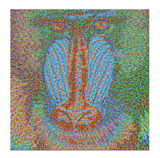


Identification schemes

Compression-method -independent schemes
- have robustness against JPEG compression
- include image hashing-based schemes(ex. [2]).

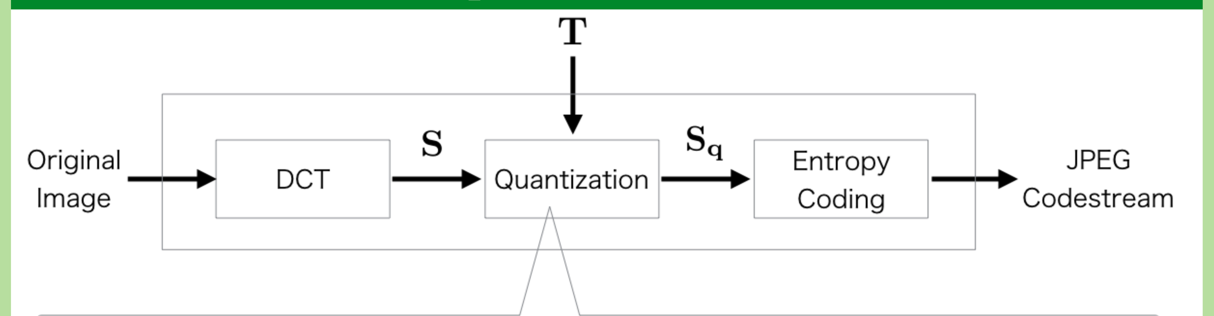
Compression-method -depend schemes
- have **strong robustness** against JPEG compression.
- do not provide **any false negative matches**.
- include **proposed scheme** and signs of coefficients -based scheme(ex. [1]).

- What are target applications?
  ⇒Relating images uploaded to SNs with downloaded ones and tamper detection.
- What limitations do conventional schemes using signs of DCT coefficients have?
  - The features have to be protected because they have **visible information.**
  - **Coding conditions** are limited.

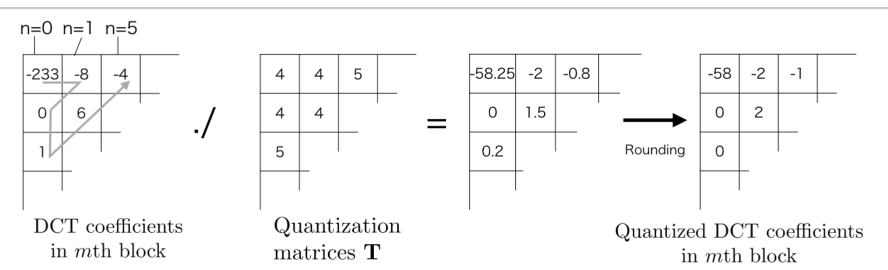


Images reconstructed from extracted features

JPEG Compression



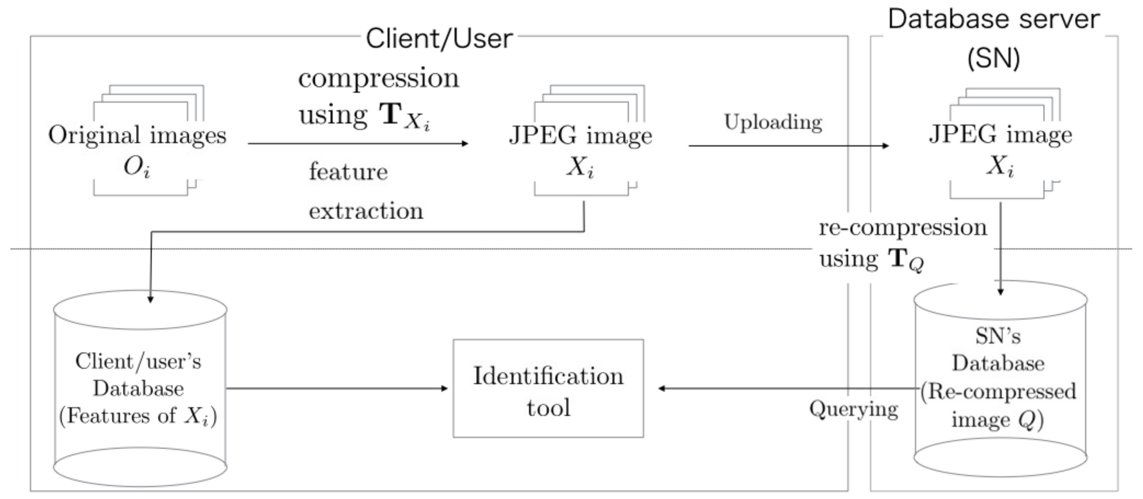
$$S_q(m,n) = \mathrm{round}\left(\frac{S(m,n)}{T(n)}\right), \text{ where } T(n) = \begin{cases} \mathrm{round}\left(\frac{T_0(n) * \lfloor \frac{5000}{QF} \rfloor}{100}\right), & 1 \le QF < 50, \\ \mathrm{round}\left(\frac{T_0(n)*(200-2*QF)}{100}\right), & 50 \le QF \le 100, \end{cases}$$



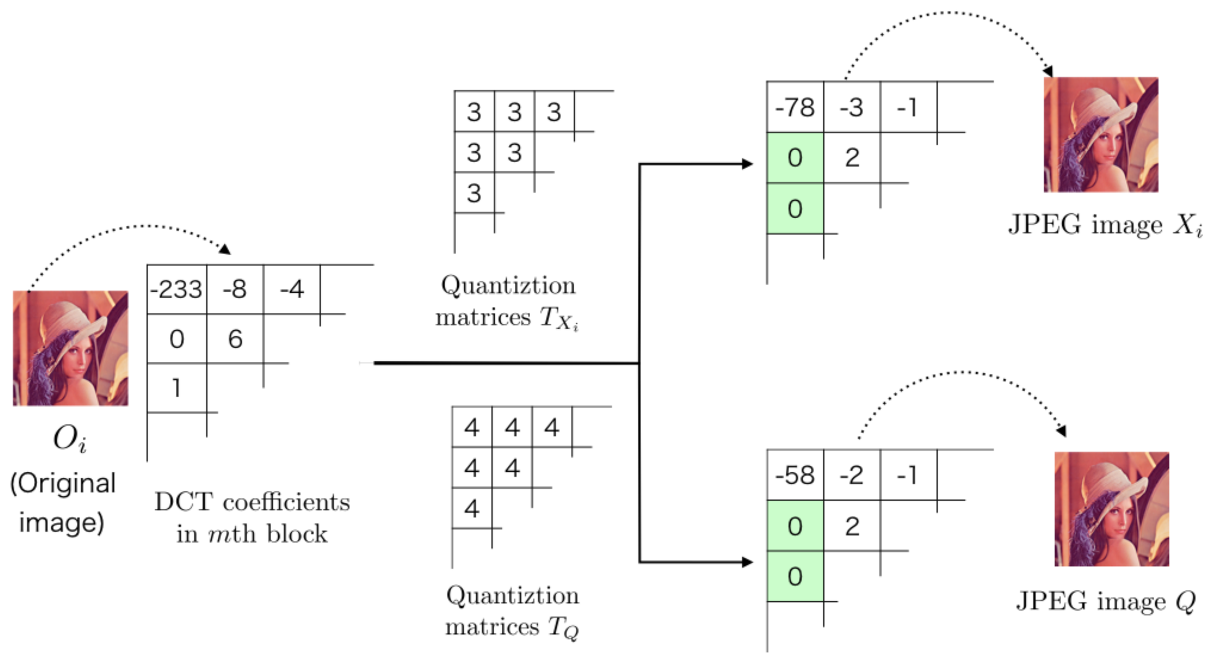
DCT coefficients in mth block — Quantization matrices $\mathbf{T}$ — Quantized DCT coefficients in mth block

Scenario

- Images uploaded to SNs and downloaded ones are identified.
- Features have to provide **no visible information** due to privacy concerns or copyright protection.



Property of DCT coefficients



- When X_i and Q are generated from the same original image O_i,

$$\begin{cases} Q(m,n) = 0, & \text{for } T_{X_i}(n) \le T_Q(n) \text{ and } X_i(m,n) = 0, \\ X_i(m,n) = 0, & \text{for } T_{X_i}(n) \ge T_Q(n) \text{ and } Q(m,n) = 0, \end{cases}$$

where $\forall m \in \{0, \cdots, M\}$ and $\forall n \in \{0, \cdots, N\}$.

- X_i and Q are generated from the different original images, when

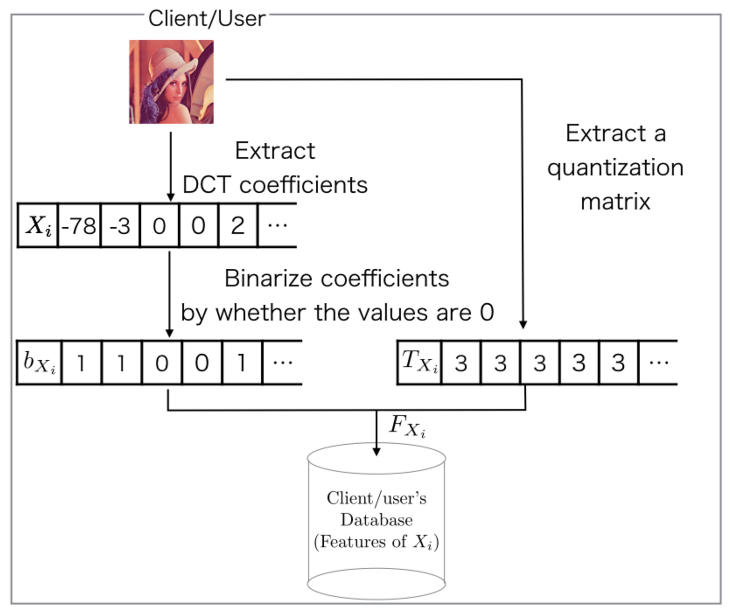
$$\begin{cases} Q(m,n) \neq 0, & \text{for } T_{X_i}(n) \le T_Q(n) \text{ and } X_i(m,n) = 0, \\ X_i(m,n) \neq 0, & \text{for } T_{X_i}(n) \ge T_Q(n) \text{ and } Q(m,n) = 0, \end{cases} \quad (1)$$

where $\exists m \in \{0, \cdots, M\}$ and $\exists n \in \{0, \cdots, N\}$.

- Proposed scheme uses Eq.(1) for identification.

Proposed Scheme

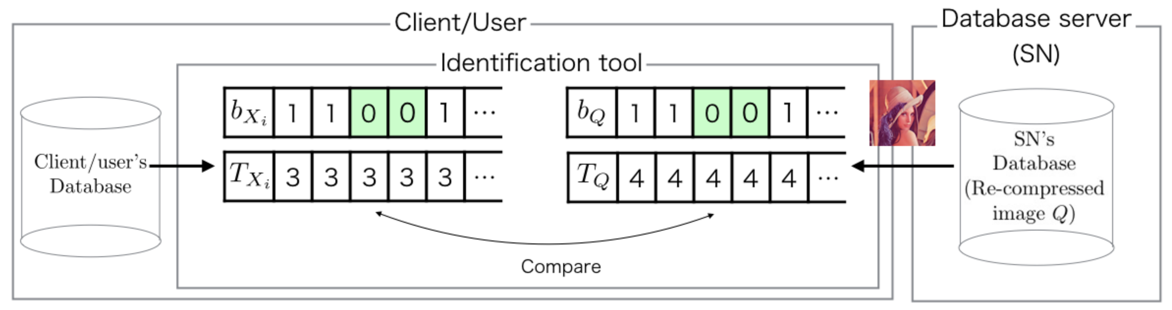
- Feature extraction process
  - Quantization matrices and positions of zero values are used as features.
  - The set of features F_{X_i} provides no visible information.



- Identification process
  - Eq.(1) and features are used.
  - Client/user confirms whether Eq.(1) is satisfied at each position.
    - not satisfied at all positions ⇒ Two images have **same original image**.
    - satisfied at a position ⇒ Two images have **different original images**.

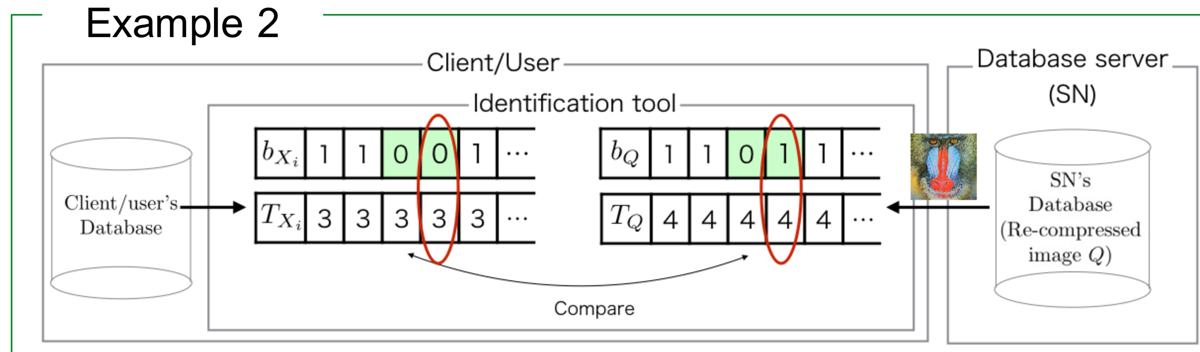
$$\begin{cases} Q(m,n) \neq 0, & \text{for } T_{X_i}(n) \le T_Q(n) \text{ and } X_i(m,n) = 0, \\ X_i(m,n) \neq 0, & \text{for } T_{X_i}(n) \ge T_Q(n) \text{ and } Q(m,n) = 0, \end{cases} \quad (1)$$

Example 1



In this case, client/user compares up to last values.
↓
Client/user judges that two images have the **same original image**.

Example 2



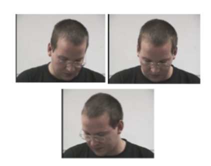
In this case, client/user compares up to 4th values.
↓
Client/user judges that two images have the **different original images.**

Simulation

- Simulation condition
  - Dataset :Head Posed Image Database (186 images in "Person01")
  - Encoder: Independent JPEG Group(IJG)
  - Identification was performed 186x744 times for each database.

Coding conditions

JPEG images		Quality factors	$\mathbf{T}_0$
Images stored as features in	DB_1	$QF_{X_i} = 50$	IJG
	DB_2	$QF_{X_i} = 75$	IJG
	DB_3	$QF_{X_i} = 50$	HVS
Query images		$QF_Q = 40, 60, 85, 95$	IJG



Example of images (size 288x384)

16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

IJG

16	16	16	16	17	18	21	24
16	16	16	16	17	19	22	25
16	16	17	18	20	22	25	29
16	16	18	21	24	27	31	36
17	17	20	24	30	35	41	47
18	19	22	27	35	44	54	65
21	22	25	31	41	54	70	88
24	25	29	36	47	65	88	115

HVS

- Simulation result
  - Measurement

$$TPR = \frac{TP}{TP+FN}, \; FPR = \frac{FP}{FP+TN}$$

TPR=100% means that there were **no false negative matches**.

Querying performance for images with $\mathrm{QF}_Q = 60$

scheme	database	TPR[%]	FPR[%]
proposed	DB_1	100	0
	DB_2	100	0
FCS-based[1]	DB_1	100	0
	DB_2	0	0
image hashing[2]	DB_1	98.92	0.03
	DB_2	97.85	0.04

Querying performance for all query images

scheme	database	TPR[%]	FPR[%]
proposed	DB_1	100	0
	DB_2	100	0
	DB_3	100	0
FCS-based[1]	DB_1	75	0
	DB_2	50	0
	DB_3	71.23	0
image hashing[2]	DB_1	98.79	0.03
	DB_2	99.33	0.03
	DB_3	98.52	0.03

- **Only proposed scheme provided no false negative matches.**
  - Image hashing-based one did in all cases.
  - FCS-based one did under $QF_{X_i} < QF_Q$ or different $\mathbf{T}_0$

Conclusion

- Our proposed identification scheme for JPEG images
  - uses quantization matrices and the positions of zero values as features.
    ⇒They do not provide **no visible information**.
  - uses the features and the property of DCT coefficients for identification.
    ⇒The use of them allows us to provide **no false negative matches in principle**.
  - outperforms the querying performance.

Reference

[1] K. Iida and H. Kiya, "Secure and Robust Identification Based on Fuzzy Commitment Scheme for JPEG Images," in *Proc. IEEE BMSB*, June, 2016.

[2] Y. Li and P. Wang, "Robust image hashing based on low-rank and sparse decomposition," in *Proc. IEEE ICASSP*, March, 2017.