In-Depth DCT Coefficient Distribution Analysis for First Quantization Estimation

S. Battiato[1], O. Giudice[1], F. Guarnera[1], **G. Puglisi**[2]

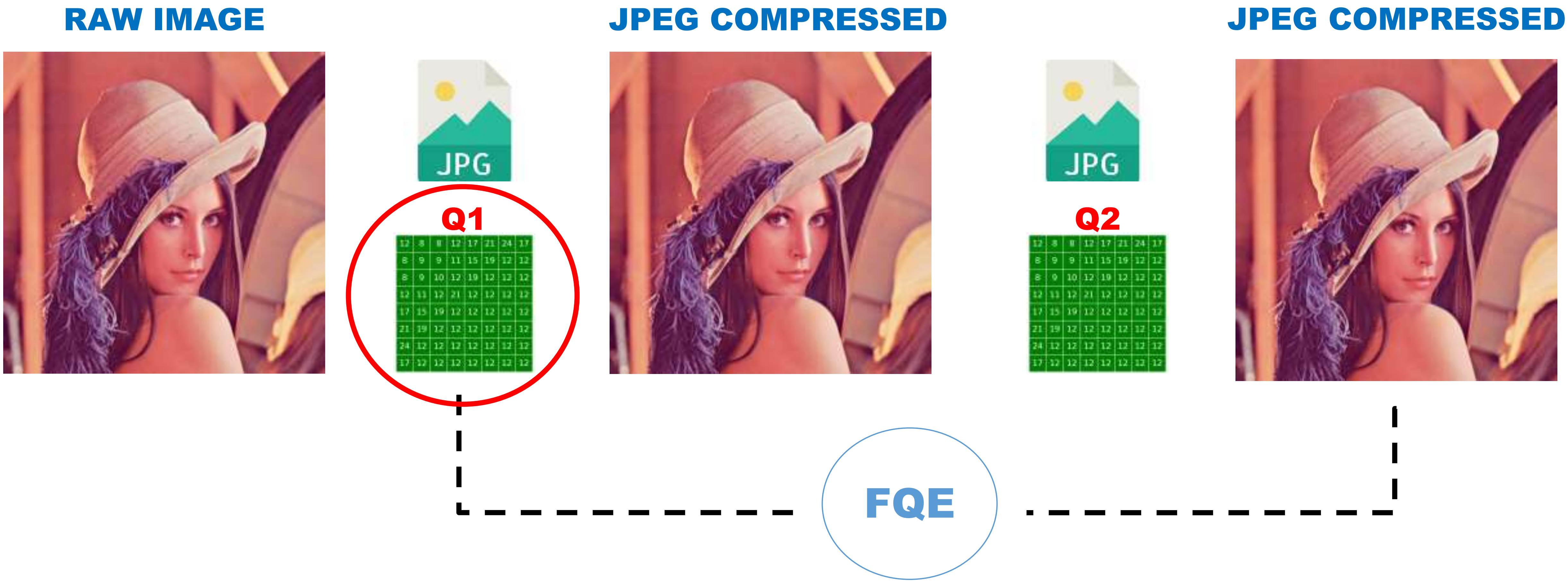
[1]University of Catania, [2]University of Cagliari

UNIVERSITÀ degli STUDI di CATANIA

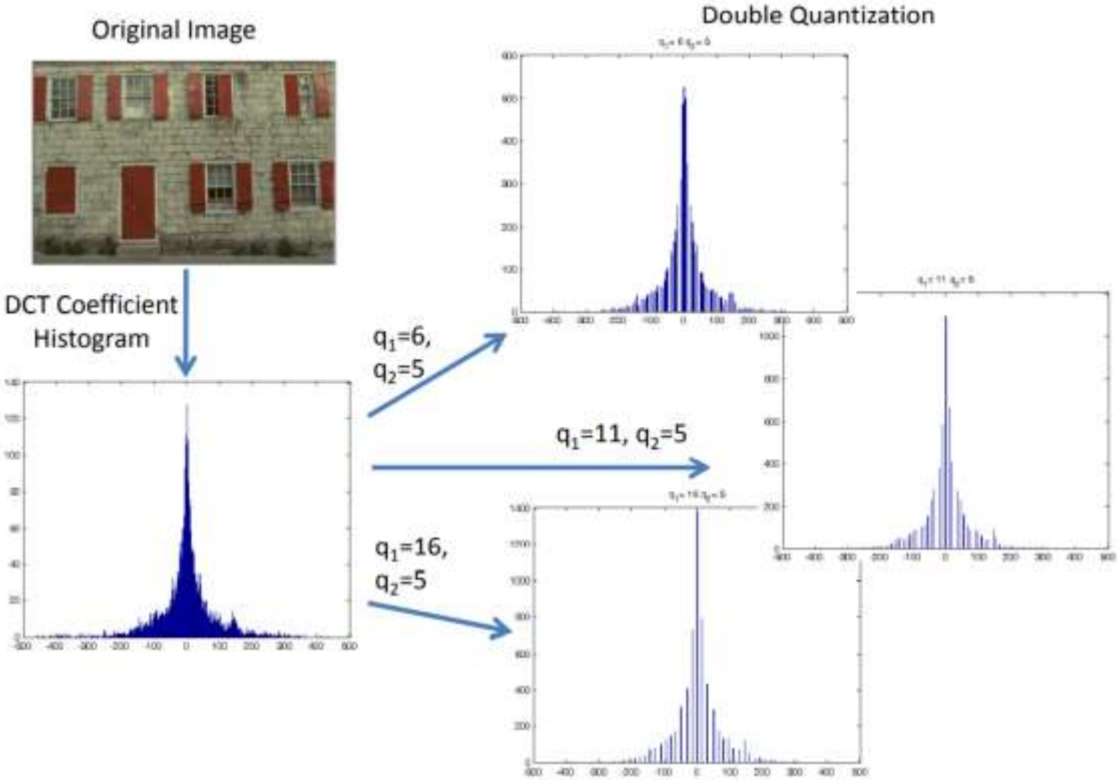
UNIVERSITÀ degli STUDI di CAGLIARI

iCtLab Digital Forensics

FQE: First Quantization Estimation



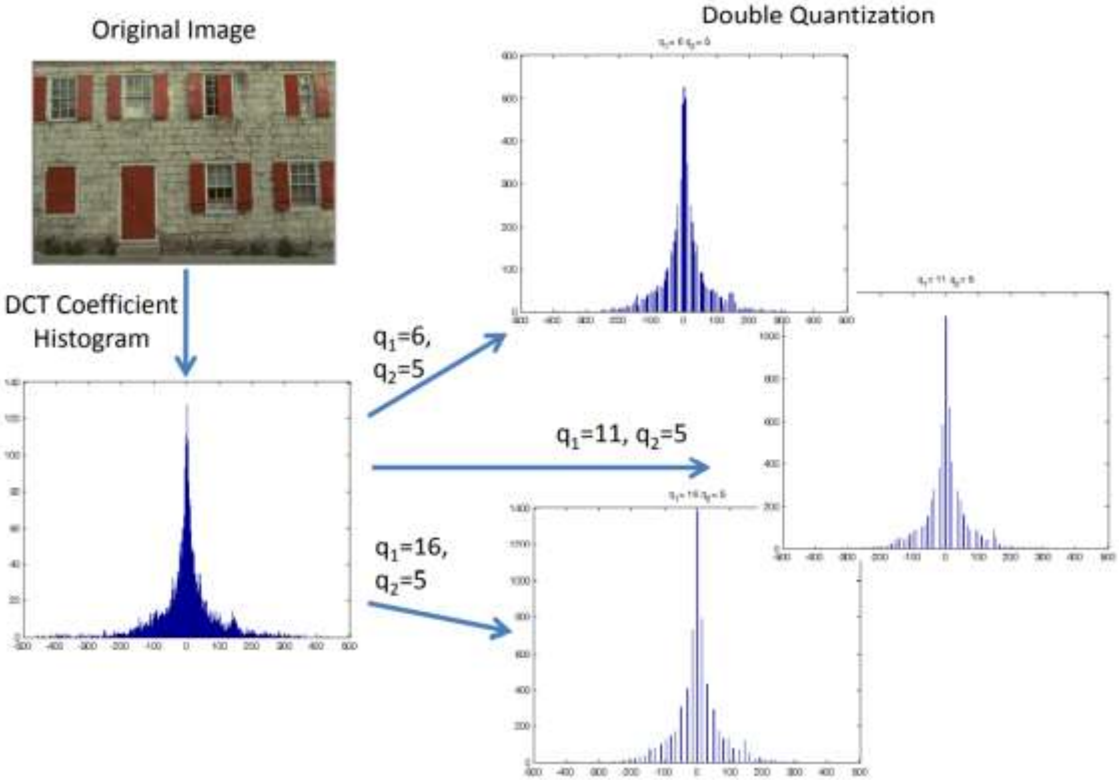
State-of-the-art



Consecutive quantizations introduced periodic artifacts into the histogram of DCT coefficients.

T. Bianchi and A. Piva, “Image forgery localization via block-grained analysis of JPEG artifacts,” Proc. of IEEE Trans. on Information Forensics and Security, vol. 7, no. 3, p. 1003, 2012.
F. Galvan, G. Puglisi, A. R. Bruna, and S. Battiato, “First quantization matrix estimation from double compressed JPEG images,” IEEE Trans. on Information Forensics and Security, vol. 9, no. 8, pp. 1299–1310, 2014.
N. Dalmia and M. Okade, “First quantization matrix estimation for double compressed JPEG images utilizing novel dct histogram selection strategy,” in Proc. of the Tenth Indian Conference on Computer Vision, Graphics and Image Processing, 2016, pp. 1–8.
Y. Niu, B. Tondi, Y. Zhao, and M. Barni, “Primary quantization matrix estimation of double compressed JPEG images via CNN,” IEEE Signal Processing Letters, vol. 27, pp. 191–195, 2020.

State-of-the-art

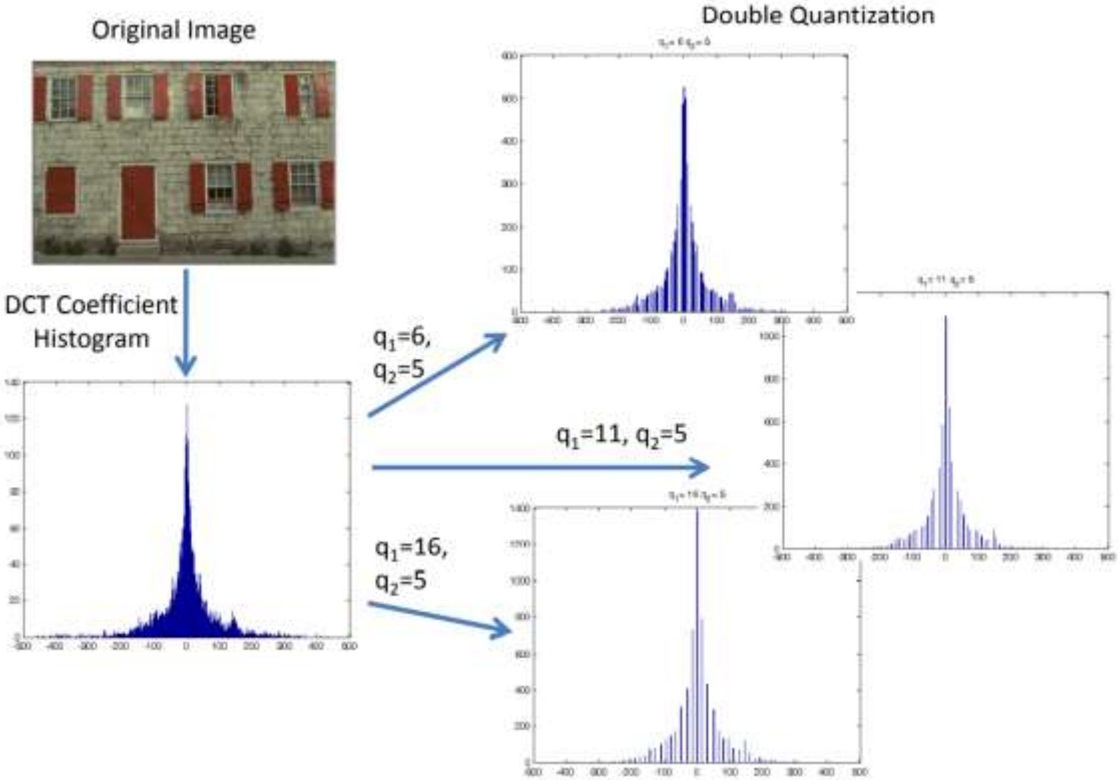


Statistical approaches

Consecutive quantizations introduced periodic artifacts into the histogram of DCT coefficients.

T. Bianchi and A. Piva, "Image forgery localization via block-grained analysis of JPEG artifacts," Proc. of IEEE Trans. on Information Forensics and Security, vol. 7, no. 3, p. 1003, 2012.
F. Galvan, G. Puglisi, A. R. Bruna, and S. Battiato, "First quantization matrix estimation from double compressed JPEG images," IEEE Trans. on Information Forensics and Security, vol. 9, no. 8, pp. 1299–1310, 2014.
N. Dalmia and M. Okade, "First quantization matrix estimation for double compressed JPEG images utilizing novel dct histogram selection strategy," in Proc. of the Tenth Indian Conference on Computer Vision, Graphics and Image Processing, 2016, pp. 1–8.
Y. Niu, B. Tondi, Y. Zhao, and M. Barni, "Primary quantization matrix estimation of double compressed JPEG images via CNN," IEEE Signal Processing Letters, vol. 27, pp. 191–195, 2020.

State-of-the-art



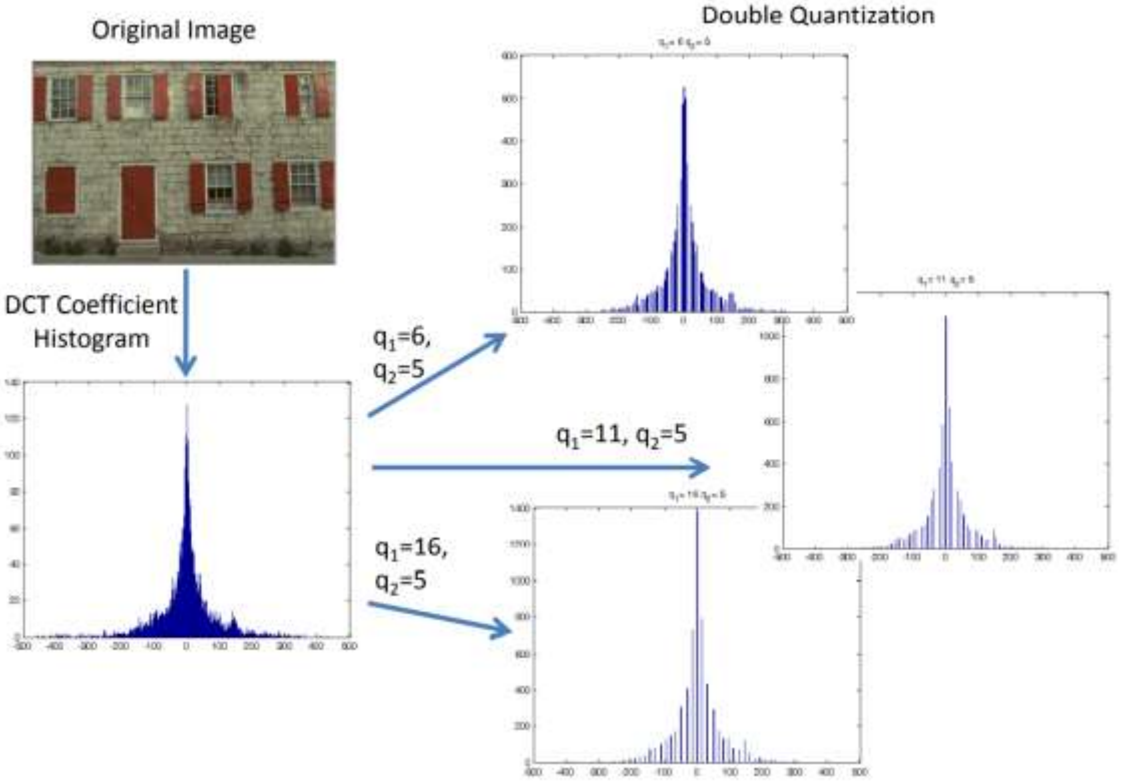
Statistical approaches

They usually provide satisfactory results only at specific combinations between first and second compression factors.

Consecutive quantizations introduced periodic artifacts into the histogram of DCT coefficients.

T. Bianchi and A. Piva, "Image forgery localization via block-grained analysis of JPEG artifacts," Proc. of IEEE Trans. on Information Forensics and Security, vol. 7, no. 3, p. 1003, 2012.
F. Galvan, G. Puglisi, A. R. Bruna, and S. Battiato, "First quantization matrix estimation from double compressed JPEG images," IEEE Trans. on Information Forensics and Security, vol. 9, no. 8, pp. 1299–1310, 2014.
N. Dalmia and M. Okade, "First quantization matrix estimation for double compressed JPEG images utilizing novel dct histogram selection strategy," in Proc. of the Tenth Indian Conference on Computer Vision, Graphics and Image Processing, 2016, pp. 1–8.
Y. Niu, B. Tondi, Y. Zhao, and M. Barni, "Primary quantization matrix estimation of double compressed JPEG images via CNN," IEEE Signal Processing Letters, vol. 27, pp. 191–195, 2020.

State-of-the-art



Consecutive quantizations introduced periodic artifacts into the histogram of DCT coefficients.

Statistical approaches

Machine learning approaches

They usually provide satisfactory results only at specific combinations between first and second compression factors.

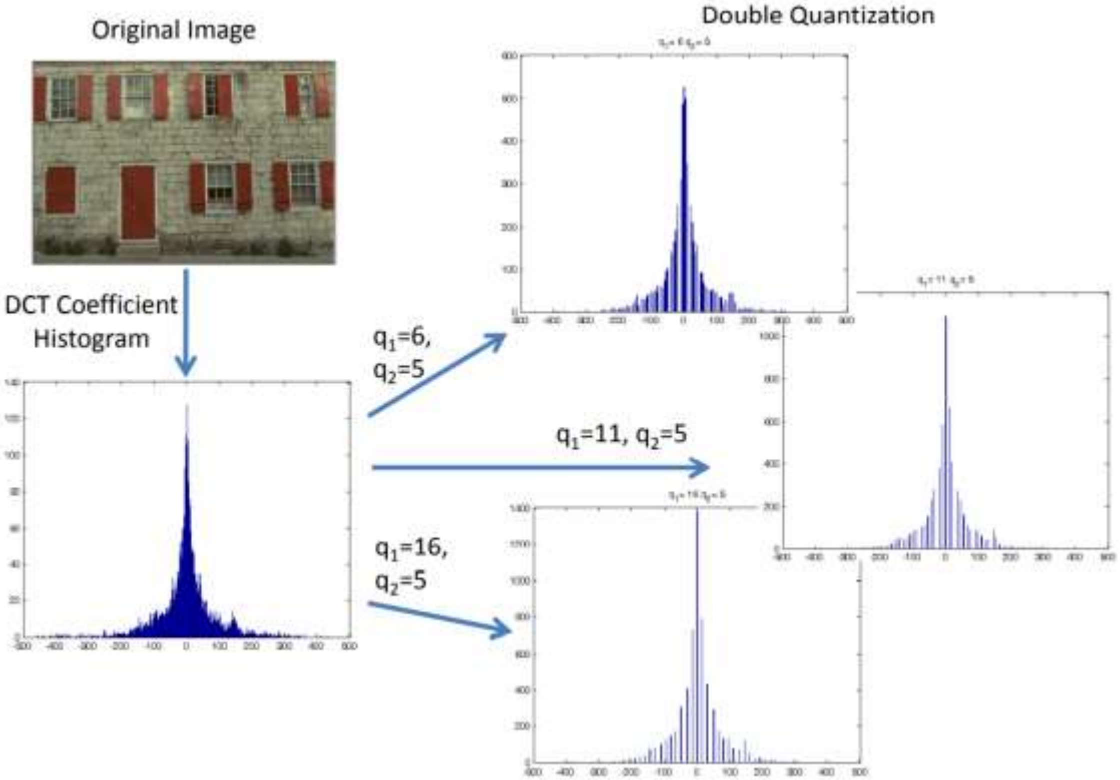
T. Bianchi and A. Piva, “Image forgery localization via block-grained analysis of JPEG artifacts,” Proc. of IEEE Trans. on Information Forensics and Security, vol. 7, no. 3, p. 1003, 2012.

F. Galvan, G. Puglisi, A. R. Bruna, and S. Battiato, “First quantization matrix estimation from double compressed JPEG images,” IEEE Trans. on Information Forensics and Security, vol. 9, no. 8, pp. 1299–1310, 2014.

N. Dalmia and M. Okade, “First quantization matrix estimation for double compressed JPEG images utilizing novel dct histogram selection strategy,” in Proc. of the Tenth Indian Conference on Computer Vision, Graphics and Image Processing, 2016, pp. 1–8.

Y. Niu, B. Tondi, Y. Zhao, and M. Barni, “Primary quantization matrix estimation of double compressed JPEG images via CNN,” IEEE Signal Processing Letters, vol. 27, pp. 191–195, 2020.

State-of-the-art



Consecutive quantizations introduced periodic artifacts into the histogram of DCT coefficients.

Statistical approaches

They usually provide satisfactory results only at specific combinations between first and second compression factors.

Machine learning approaches

They could suffer from overfitting.

T. Bianchi and A. Piva, “Image forgery localization via block-grained analysis of JPEG artifacts,” Proc. of IEEE Trans. on Information Forensics and Security, vol. 7, no. 3, p. 1003, 2012.
F. Galvan, G. Puglisi, A. R. Bruna, and S. Battiato, “First quantization matrix estimation from double compressed JPEG images,” IEEE Trans. on Information Forensics and Security, vol. 9, no. 8, pp. 1299–1310, 2014.
N. Dalmia and M. Okade, “First quantization matrix estimation for double compressed JPEG images utilizing novel dct histogram selection strategy,” in Proc. of the Tenth Indian Conference on Computer Vision, Graphics and Image Processing, 2016, pp. 1–8.
Y. Niu, B. Tondi, Y. Zhao, and M. Barni, “Primary quantization matrix estimation of double compressed JPEG images via CNN,” IEEE Signal Processing Letters, vol. 27, pp. 191–195, 2020.

Training Dataset Generation

D. Dang-Nguyen, C. Pasquini, V. Conotter, and G. Boato, “Raise: a raw images dataset for digital image forensics,” in Proc. of the 6th ACM Multimedia Systems Conference, 2015, pp. 219–224.

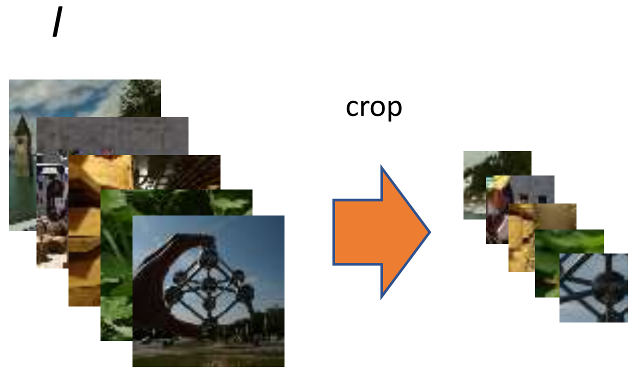
Training Dataset Generation

I

8156 high resolution
images in TIFF format
(RAISE dataset)

D. Dang-Nguyen, C. Pasquini, V. Conotter, and G. Boato, “Raise: a raw images dataset for digital image forensics,” in Proc. of the 6th ACM Multimedia Systems Conference, 2015, pp. 219–224.

Training Dataset Generation



8156 high resolution images in TIFF format (RAISE dataset)

64×64 patches extracted from the center

D. Dang-Nguyen, C. Pasquini, V. Conotter, and G. Boato, "Raise: a raw images dataset for digital image forensics," in Proc. of the 6th ACM Multimedia Systems Conference, 2015, pp. 219–224.

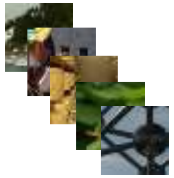
Training Dataset Generation

$$\begin{bmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} 2 & \cdots & 2 \\ \vdots & \ddots & \vdots \\ 2 & \cdots & 2 \end{bmatrix} \begin{bmatrix} 3 & \cdots & 3 \\ \vdots & \ddots & \vdots \\ 3 & \cdots & 3 \end{bmatrix} \begin{bmatrix} 4 & \cdots & 4 \\ \vdots & \ddots & \vdots \\ 4 & \cdots & 4 \end{bmatrix} \cdots \begin{bmatrix} q1_{max} & \cdots & q1_{max} \\ \vdots & \ddots & \vdots \\ q1_{max} & \cdots & q1_{max} \end{bmatrix}$$

Fist JPEG compression with constant matrices

I

crop

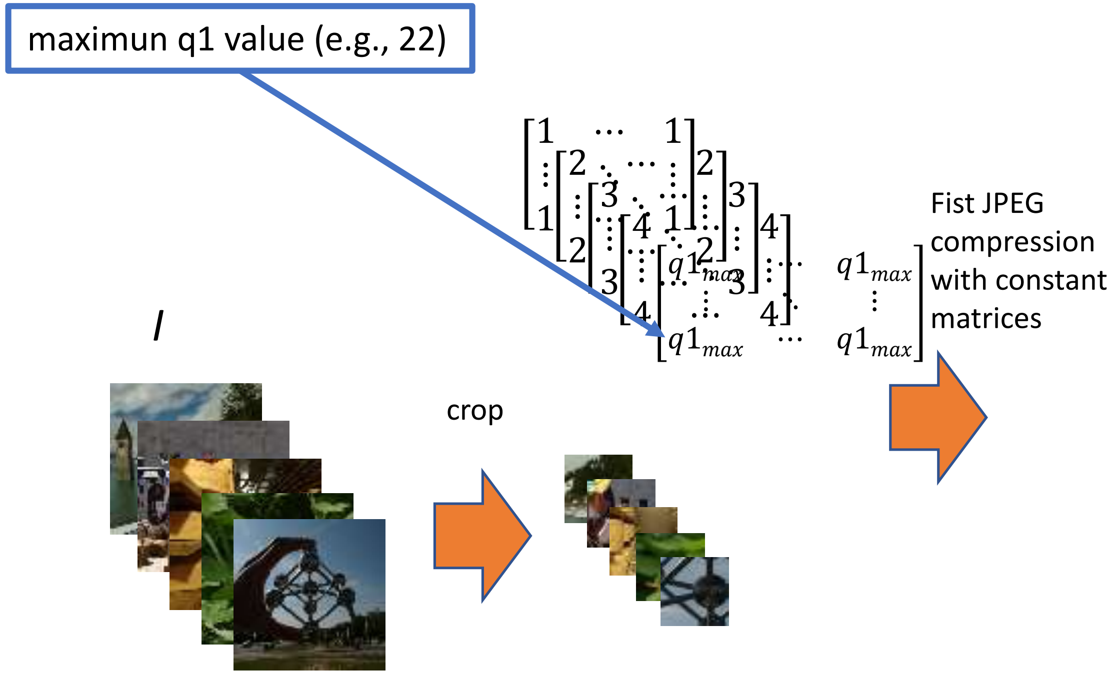


8156 high resolution images in TIFF format (RAISE dataset)

64×64 patches extracted from the center

D. Dang-Nguyen, C. Pasquini, V. Conotter, and G. Boato, "Raise: a raw images dataset for digital image forensics," in Proc. of the 6th ACM Multimedia Systems Conference, 2015, pp. 219–224.

Training Dataset Generation

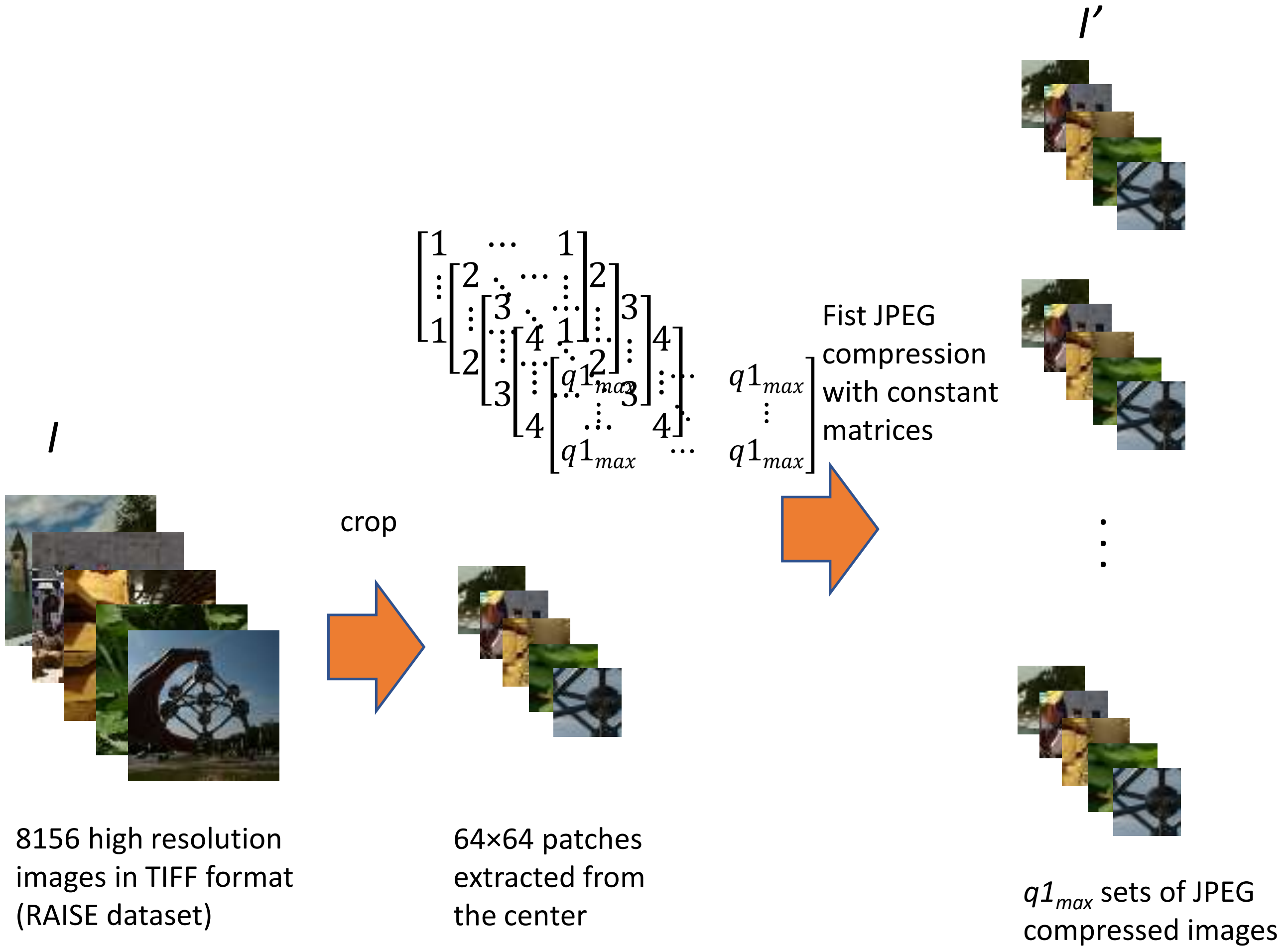


8156 high resolution images in TIFF format (RAISE dataset)

64×64 patches extracted from the center

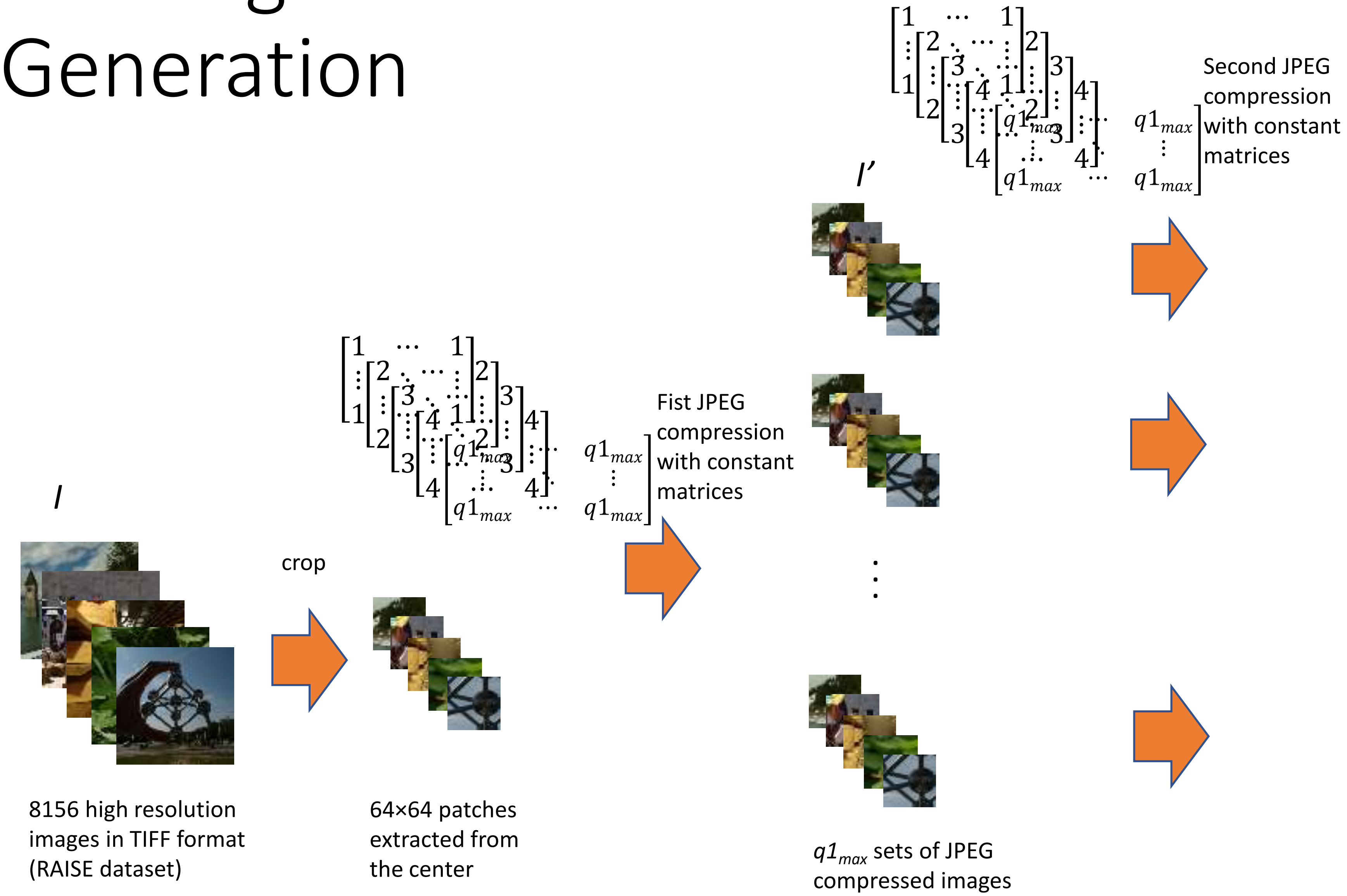
D. Dang-Nguyen, C. Pasquini, V. Conotter, and G. Boato, “Raise: a raw images dataset for digital image forensics,” in Proc. of the 6th ACM Multimedia Systems Conference, 2015, pp. 219–224.

Training Dataset Generation



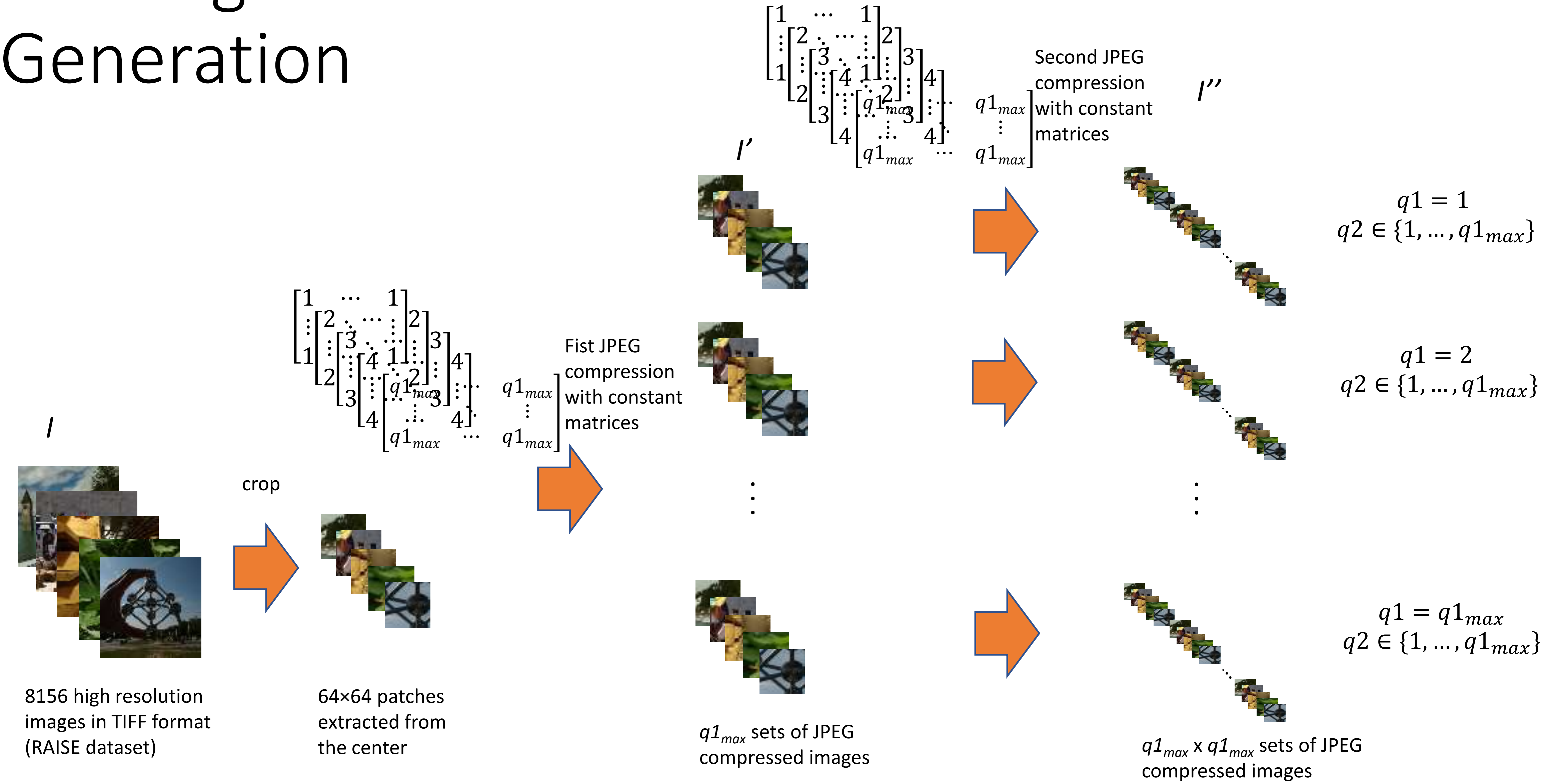
D. Dang-Nguyen, C. Pasquini, V. Conotter, and G. Boato, “Raise: a raw images dataset for digital image forensics,” in Proc. of the 6th ACM Multimedia Systems Conference, 2015, pp. 219–224.

Training Dataset Generation



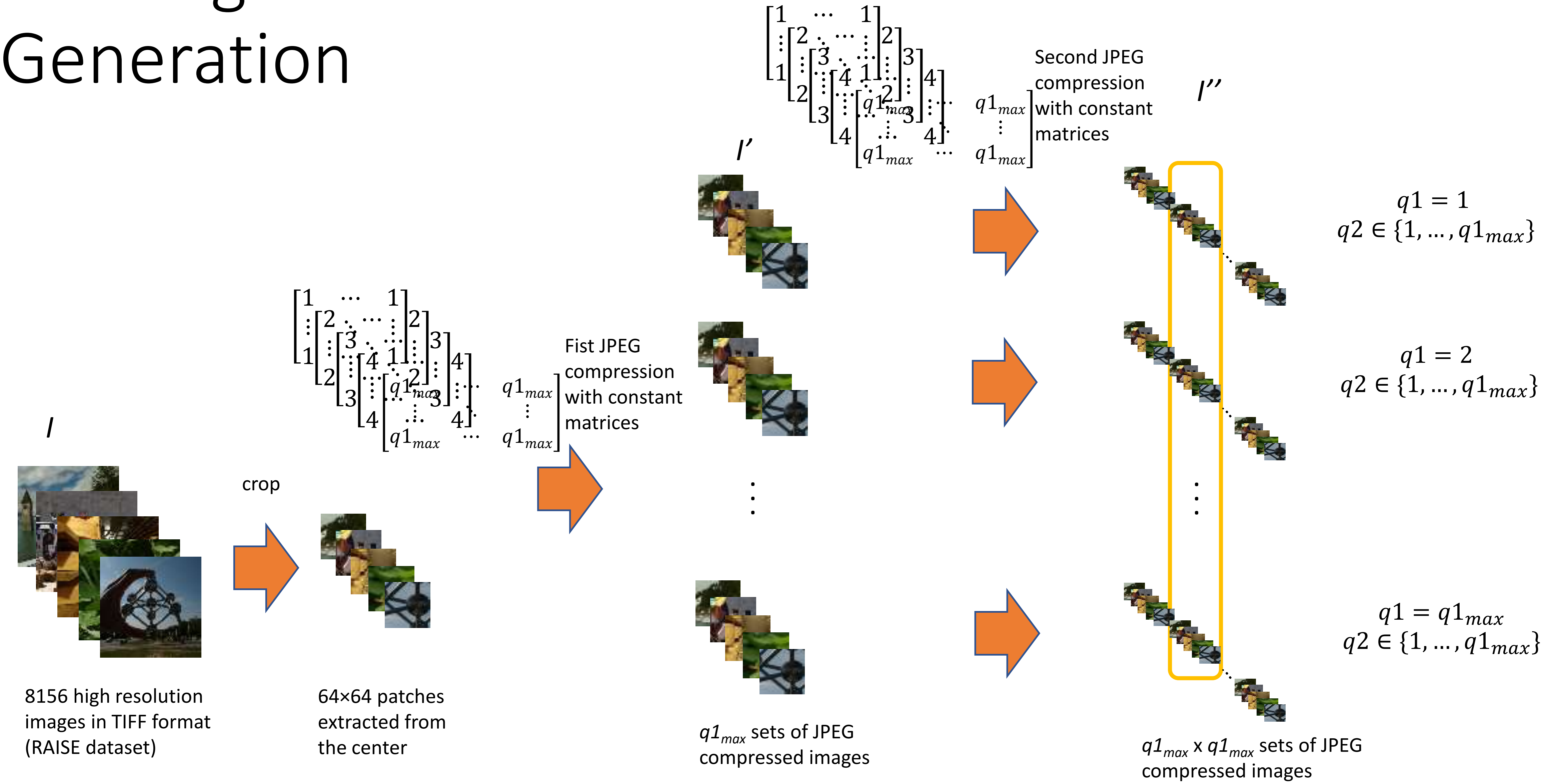
D. Dang-Nguyen, C. Pasquini, V. Conotter, and G. Boato, "Raise: a raw images dataset for digital image forensics," in Proc. of the 6th ACM Multimedia Systems Conference, 2015, pp. 219–224.

Training Dataset Generation



D. Dang-Nguyen, C. Pasquini, V. Conotter, and G. Boato, "Raise: a raw images dataset for digital image forensics," in Proc. of the 6th ACM Multimedia Systems Conference, 2015, pp. 219–224.

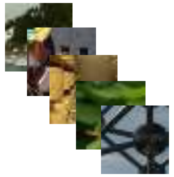
Training Dataset Generation

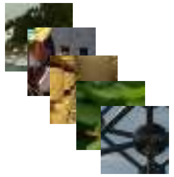


D. Dang-Nguyen, C. Pasquini, V. Conotter, and G. Boato, “Raise: a raw images dataset for digital image forensics,” in Proc. of the 6th ACM Multimedia Systems Conference, 2015, pp. 219–224.

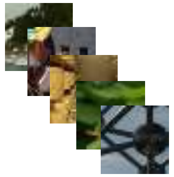
Training Dataset Generation

I''



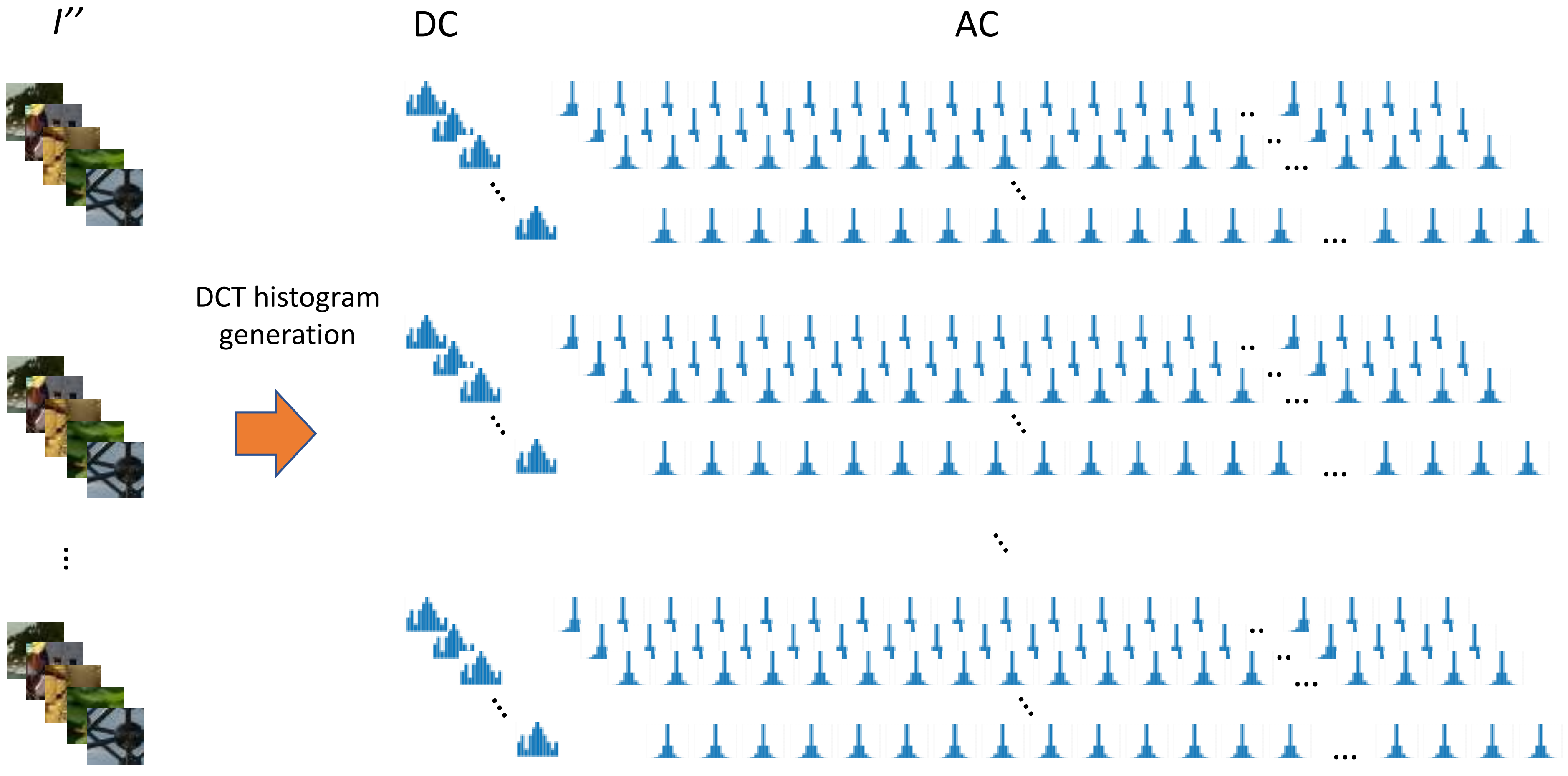


⋮



Double compressed images with
$q1 \in \{1, \dots, q1_{max}\}$
$q2 = q2_i$

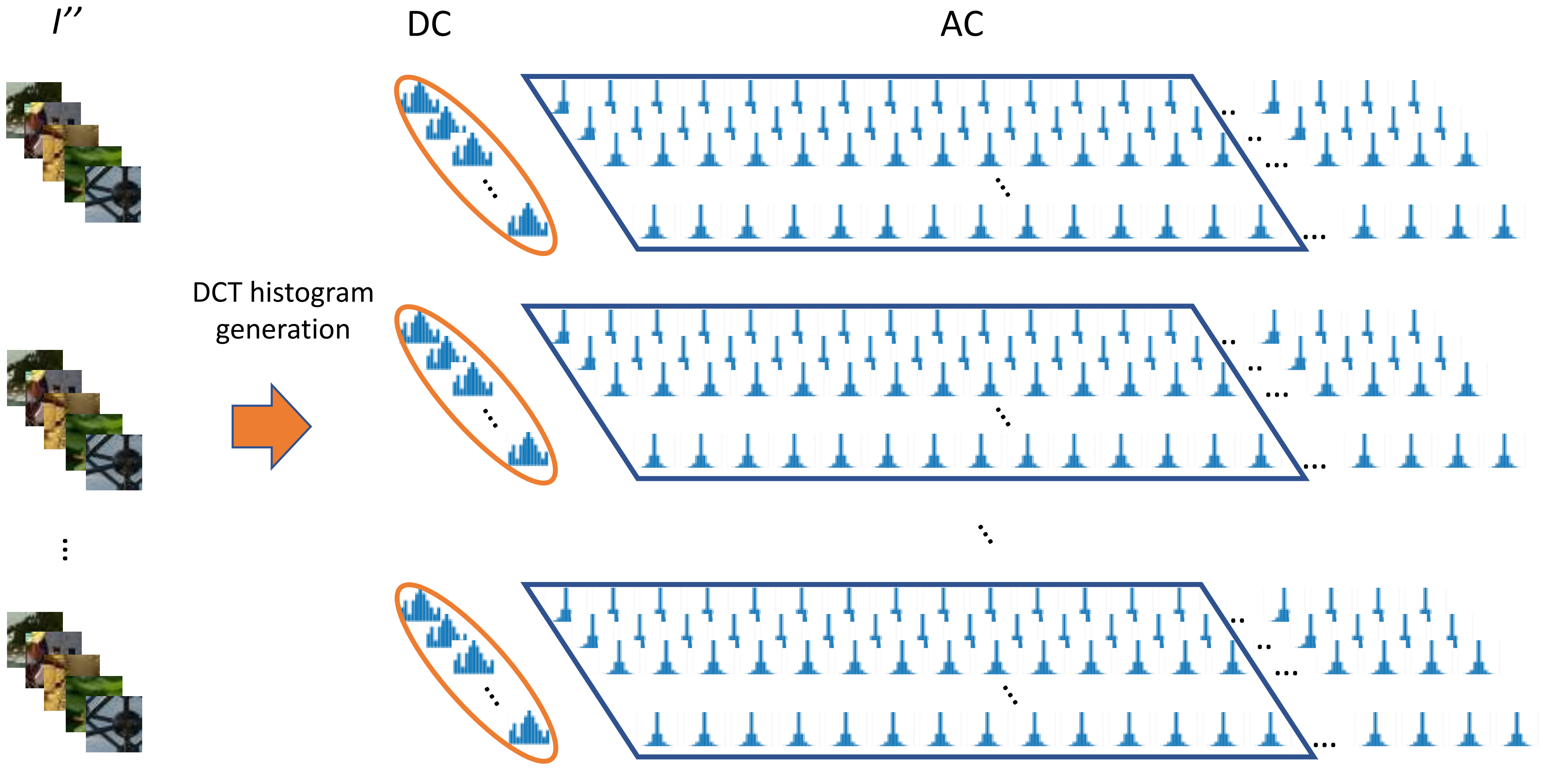
Training Dataset Generation



Double compressed images with
$q1 \in \{1, \ldots, q1_{max}\}$
$q2 = q2_i$

DCT histograms

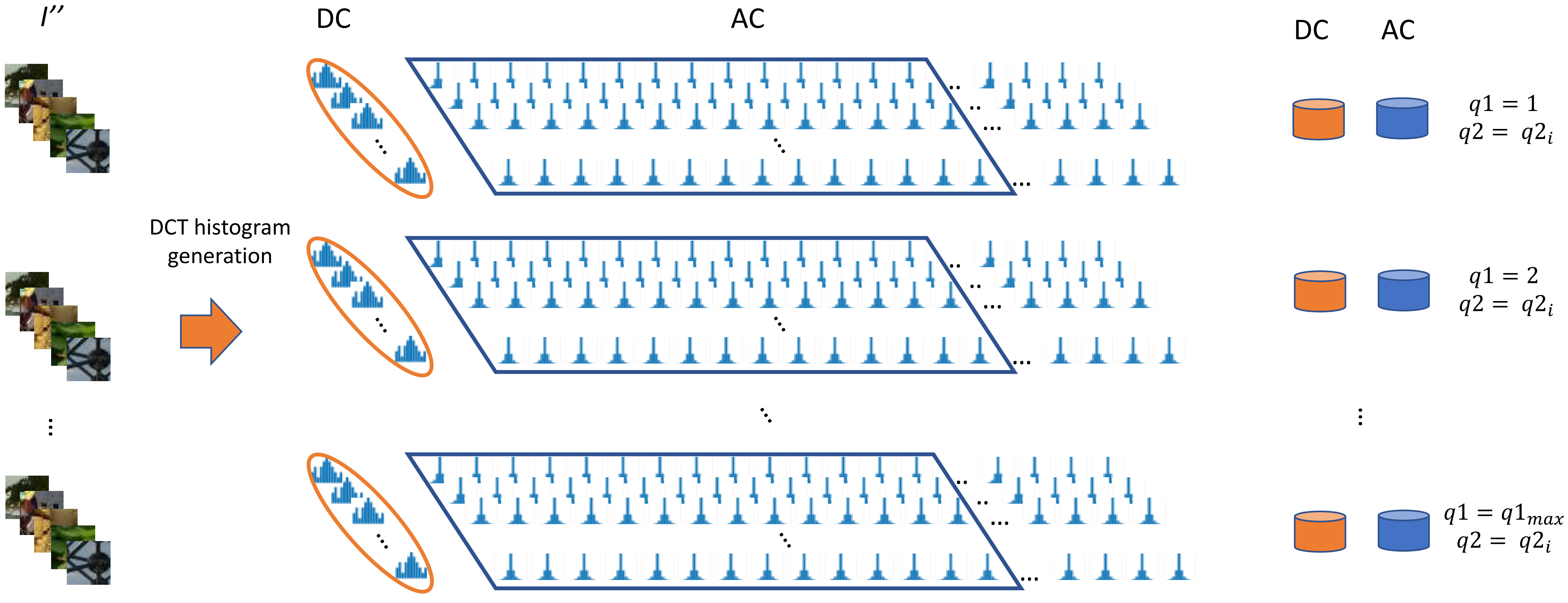
Training Dataset Generation



Double compressed images with
$q1 \in \{1, \dots, q1_{max}\}$
$q2 = q2_i$

DCT histograms

Training Dataset Generation



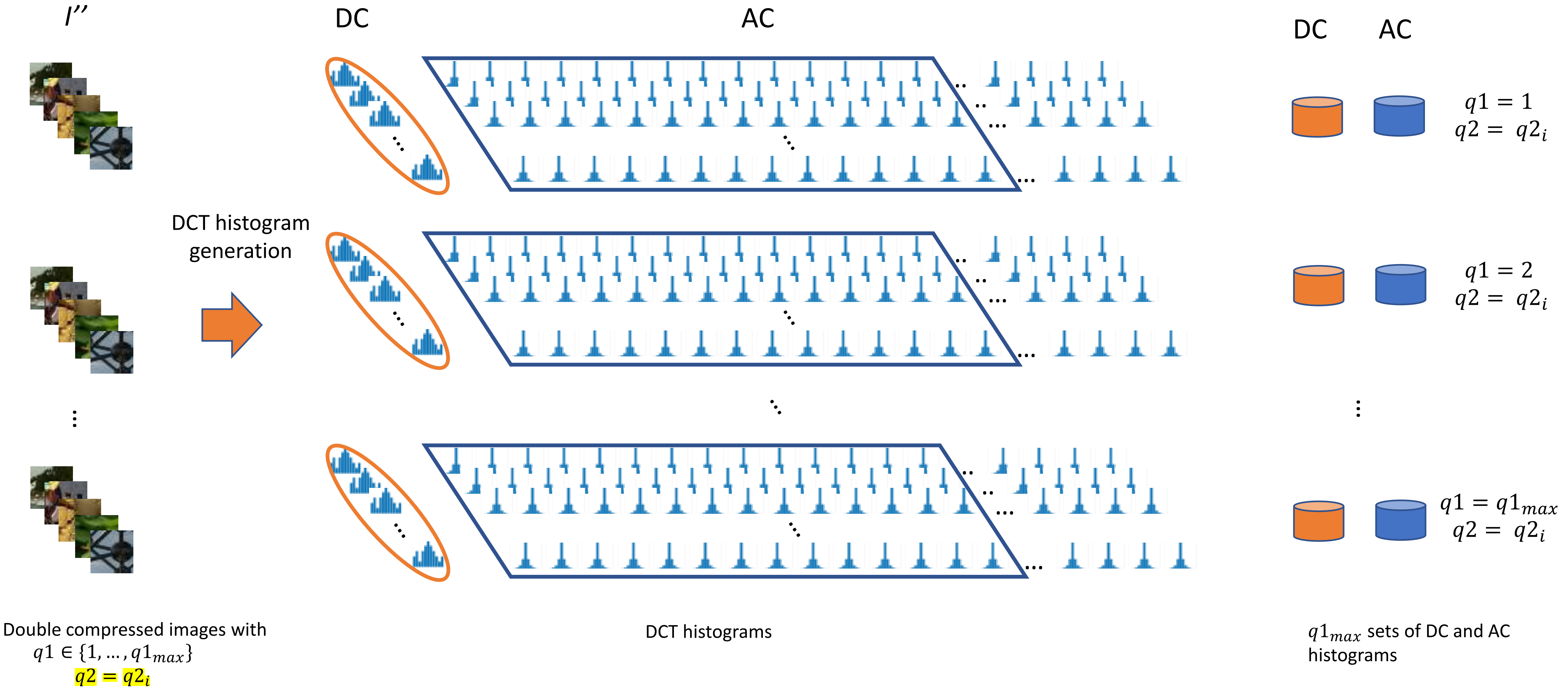
Double compressed images with
$q1 \in \{1, \ldots, q1_{max}\}$
$q2 = q2_i$

DCT histograms

$q1_{max}$ sets of DC and AC histograms

Training Dataset Generation

Considering all the q2 in the range $\{1, \ldots, q1_{max}\}$, $q1_{max} \times q1_{max}$ sets of DC and AC histograms are generated.



Algorithm 1 The Proposed FQE Technique

Input: double compressed image I''
Output: $\{q1_1, q1_2, \ldots, q1_k\}$

Initialization : k, $q1_{max}$

for $i = 1$ to k **do**
  h_i : distribution of i-th DCT coefficient
  if $(i = 1)$ **then**
    $D : DC_{dset}$
    m : median value of h_i
  else
    $D : AC_{dset}$
    β : β fitted on Laplacian h_i
  end if
  $q2_i$: quantization factor of Q_2 for i-th DCT
  for $j = 1$ to $q1_{max}$ **do**
    $D_{j,q2_i}$: sub-dataset $(q1, q2)$ with $q1 = j$, $q2 = q2_i$
    $D_{j,q2_i}(m, \beta)$: sub-range with most similar m, β
    $d_{i,j}$: lower χ^2 distance between h_i and $D_{j,q2_i}$
  end for
  $q1_i : \arg\min_{\{d_{i,j}\}}$, $j \in \{1, 2, \ldots, q1_{max}\}$
end for
regularize($\{q1_1, q1_2, \ldots, q1_k\}$)
return $\{q1_1, q1_2, \ldots, q1_k\}$

Algorithm 1 The Proposed FQE Technique

Input: double compressed image I''

Output: $\{q1_1, q1_2, \ldots, q1_k\}$

Initialization : k, $q1_{max}$

for $i = 1$ to k **do**
  h_i : distribution of i-th DCT coefficient
  if $(i = 1)$ **then**
    $D : DC_{dset}$
    m : median value of h_i
  else
    $D : AC_{dset}$
    β : β fitted on Laplacian h_i
  end if
  $q2_i$: quantization factor of Q_2 for i-th DCT
  for $j = 1$ to $q1_{max}$ **do**
    $D_{j,q2_i}$: sub-dataset $(q1, q2)$ with $q1 = j$, $q2 = q2_i$
    $D_{j,q2_i}(m, \beta)$: sub-range with most similar m, β
    $d_{i,j}$: lower χ^2 distance between h_i and $D_{j,q2_i}$
  end for
  $q1_i : \arg\min_{\{d_{i,j}\}}$, $j \in \{1, 2, \ldots, q1_{max}\}$
end for
regularize($\{q1_1, q1_2, \ldots, q1_k\}$)
return $\{q1_1, q1_2, \ldots, q1_k\}$

I''

Algorithm 1 The Proposed FQE Technique

Input: double compressed image I''

Output: $\{q1_1, q1_2, \ldots, q1_k\}$

Initialization : k, $q1_{max}$

for $i = 1$ to k **do**

  h_i : distribution of i-th DCT coefficient

  if $(i = 1)$ **then**

    $D : DC_{dset}$

    m : median value of h_i

  else

    $D : AC_{dset}$

    β : β fitted on Laplacian h_i

  end if

  $q2_i$: quantization factor of Q_2 for i-th DCT

  for $j = 1$ to $q1_{max}$ **do**

    $D_{j,q2_i}$: sub-dataset $(q1, q2)$ with $q1 = j$, $q2 = q2_i$

    $D_{j,q2_i}(m, \beta)$: sub-range with most similar m, β

    $d_{i,j}$: lower χ^2 distance between h_i and $D_{j,q2_i}$

  end for

  $q1_i$: $\arg\min_{\{d_{i,j}\}}$, $j \in \{1, 2, \ldots, q1_{max}\}$

end for

regularize($\{q1_1, q1_2, \ldots, q1_k\}$)

return $\{q1_1, q1_2, \ldots, q1_k\}$

I''

$$\begin{bmatrix} & \cdots & \\ \vdots & \ddots & \vdots \\ & \cdots & \end{bmatrix}$$

First quantization matrix Q_1

Algorithm 1 The Proposed FQE Technique

Input: double compressed image I''

Output: $\{q1_1, q1_2, \ldots, q1_k\}$

Initialization : k, $q1_{max}$

number of quantization factors to be estimated (e.g., 15)

maximun q1 value (e.g., 22)

for $i = 1$ to k **do**
    h_i : distribution of i-th DCT coefficient
    if $(i = 1)$ **then**
        $D : DC_{dset}$
        m : median value of h_i
    else
        $D : AC_{dset}$
        β : β fitted on Laplacian h_i
    end if
    $q2_i$: quantization factor of Q_2 for i-th DCT
    for $j = 1$ to $q1_{max}$ **do**
        $D_{j,q2_i}$: sub-dataset $(q1, q2)$ with $q1 = j$, $q2 = q2_i$
        $D_{j,q2_i}(m, \beta)$: sub-range with most similar m, β
        $d_{i,j}$: lower χ^2 distance between h_i and $D_{j,q2_i}$
    end for
    $q1_i$: $\arg\min_{\{d_{i,j}\}}$, $j \in \{1, 2, \ldots, q1_{max}\}$
end for
regularize($\{q1_1, q1_2, \ldots, q1_k\}$)
return $\{q1_1, q1_2, \ldots, q1_k\}$

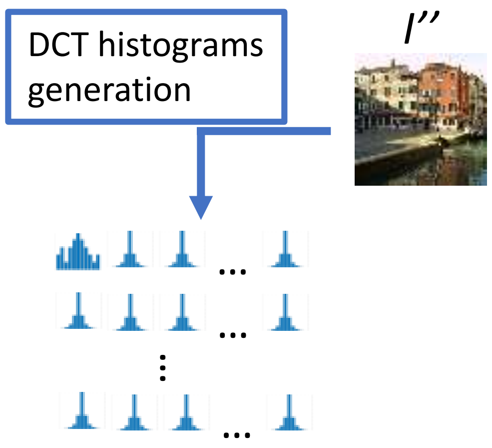
I''

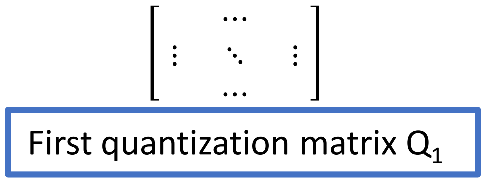
$$\begin{bmatrix} & \cdots & \\ \vdots & \ddots & \vdots \\ & \cdots & \end{bmatrix}$$

First quantization matrix Q_1

Algorithm 1 The Proposed FQE Technique

Input: double compressed image I''
Output: $\{q1_1, q1_2, \ldots, q1_k\}$
Initialization : k, $q1_{max}$
for $i = 1$ to k **do**
  h_i : distribution of i-th DCT coefficient
  if $(i = 1)$ **then**
    $D : DC_{dset}$
    m : median value of h_i
  else
    $D : AC_{dset}$
    β : β fitted on Laplacian h_i
  end if
  $q2_i$: quantization factor of Q_2 for i-th DCT
  for $j = 1$ to $q1_{max}$ **do**
    $D_{j,q2_i}$: sub-dataset $(q1, q2)$ with $q1 = j$, $q2 = q2_i$
    $D_{j,q2_i}(m, \beta)$: sub-range with most similar m, β
    $d_{i,j}$: lower χ^2 distance between h_i and $D_{j,q2_i}$
  end for
  $q1_i$: $\arg\min_{\{d_{i,j}\}}$, $j \in \{1, 2, \ldots, q1_{max}\}$
end for
regularize($\{q1_1, q1_2, \ldots, q1_k\}$)
return $\{q1_1, q1_2, \ldots, q1_k\}$



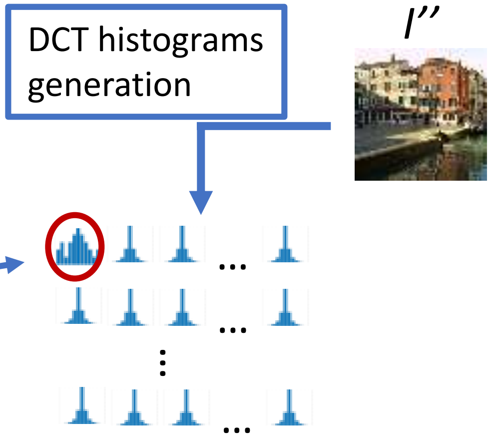


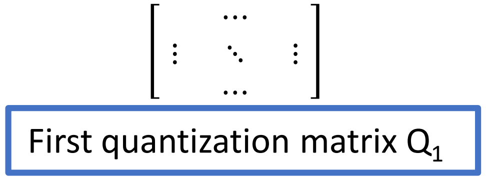
Algorithm 1 The Proposed FQE Technique

Input: double compressed image I''
Output: $\{q1_1, q1_2, \ldots, q1_k\}$

Initialization : k, $q1_{max}$

```
for i = 1 to k do
    h_i : distribution of i-th DCT coefficient
    if (i = 1) then
        D : DC_dset
        m : median value of h_i
    else
        D : AC_dset
        β : β fitted on Laplacian h_i
    end if
    q2_i : quantization factor of Q_2 for i-th DCT
    for j = 1 to q1_max do
        D_{j,q2_i} : sub-dataset (q1, q2) with q1 = j, q2 = q2_i
        D_{j,q2_i}(m, β) : sub-range with most similar m, β
        d_{i,j} : lower χ^2 distance between h_i and D_{j,q2_i}
    end for
    q1_i : arg min_{d_{i,j}}, j ∈ {1, 2, ..., q1_max}
end for
regularize({q1_1, q1_2, ..., q1_k})
return {q1_1, q1_2, ..., q1_k}
```





Algorithm 1 The Proposed FQE Technique

Input: double compressed image I''

Output: $\{q1_1, q1_2, \dots, q1_k\}$

Initialization : k, $q1_{max}$

for $i = 1$ to k **do**

  h_i : distribution of i-th DCT coefficient

  if $(i = 1)$ **then**

    $D : DC_{dset}$

    m : median value of h_i

  else

    $D : AC_{dset}$

    β : β fitted on Laplacian h_i

  end if

  $q2_i$: quantization factor of Q_2 for i-th DCT

  for $j = 1$ to $q1_{max}$ **do**

    $D_{j,q2_i}$: sub-dataset $(q1, q2)$ with $q1 = j$, $q2 = q2_i$

    $D_{j,q2_i}(m, \beta)$: sub-range with most similar m, β

    $d_{i,j}$: lower χ^2 distance between h_i and $D_{j,q2_i}$

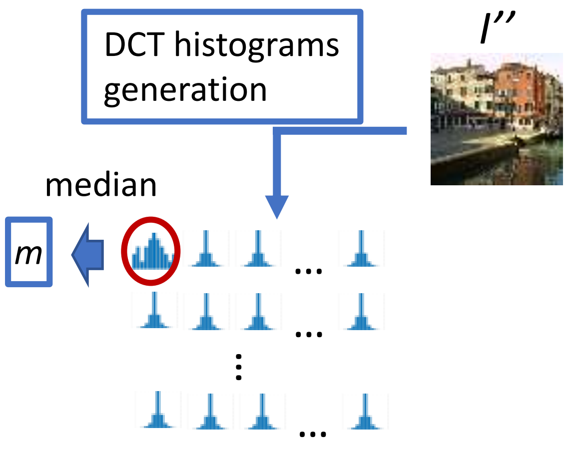
  end for

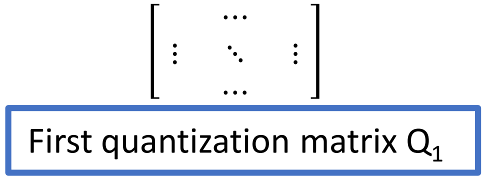
  $q1_i$: $\arg\min_{\{d_{i,j}\}}$, $j \in \{1, 2, \dots, q1_{max}\}$

end for

regularize($\{q1_1, q1_2, \dots, q1_k\}$)

return $\{q1_1, q1_2, \dots, q1_k\}$

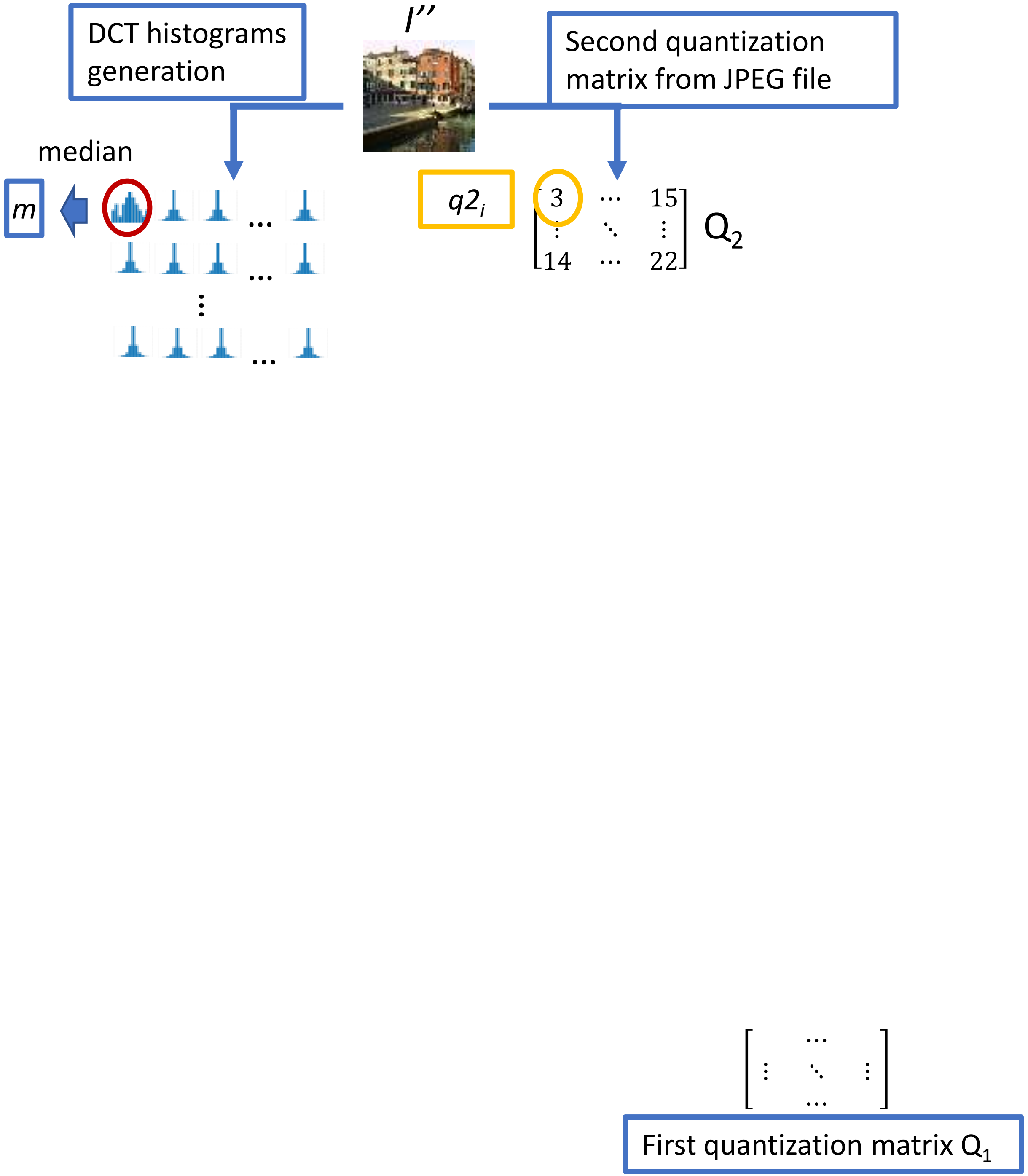




Algorithm 1 The Proposed FQE Technique

Input: double compressed image I''
Output: $\{q1_1, q1_2, \ldots, q1_k\}$
Initialization : k, $q1_{max}$

for $i = 1$ to k **do**
  h_i : distribution of i-th DCT coefficient
  if $(i = 1)$ **then**
    $D : DC_{dset}$
    m : median value of h_i
  else
    $D : AC_{dset}$
    β : β fitted on Laplacian h_i
  end if
  $q2_i$: quantization factor of Q_2 for i-th DCT
  for $j = 1$ to $q1_{max}$ **do**
    $D_{j,q2_i}$: sub-dataset $(q1, q2)$ with $q1 = j$, $q2 = q2_i$
    $D_{j,q2_i}(m, \beta)$: sub-range with most similar m, β
    $d_{i,j}$: lower χ^2 distance between h_i and $D_{j,q2_i}$
  end for
  $q1_i$: $\arg\min_{\{d_{i,j}\}}$, $j \in \{1, 2, \ldots, q1_{max}\}$
end for
regularize($\{q1_1, q1_2, \ldots, q1_k\}$)
return $\{q1_1, q1_2, \ldots, q1_k\}$



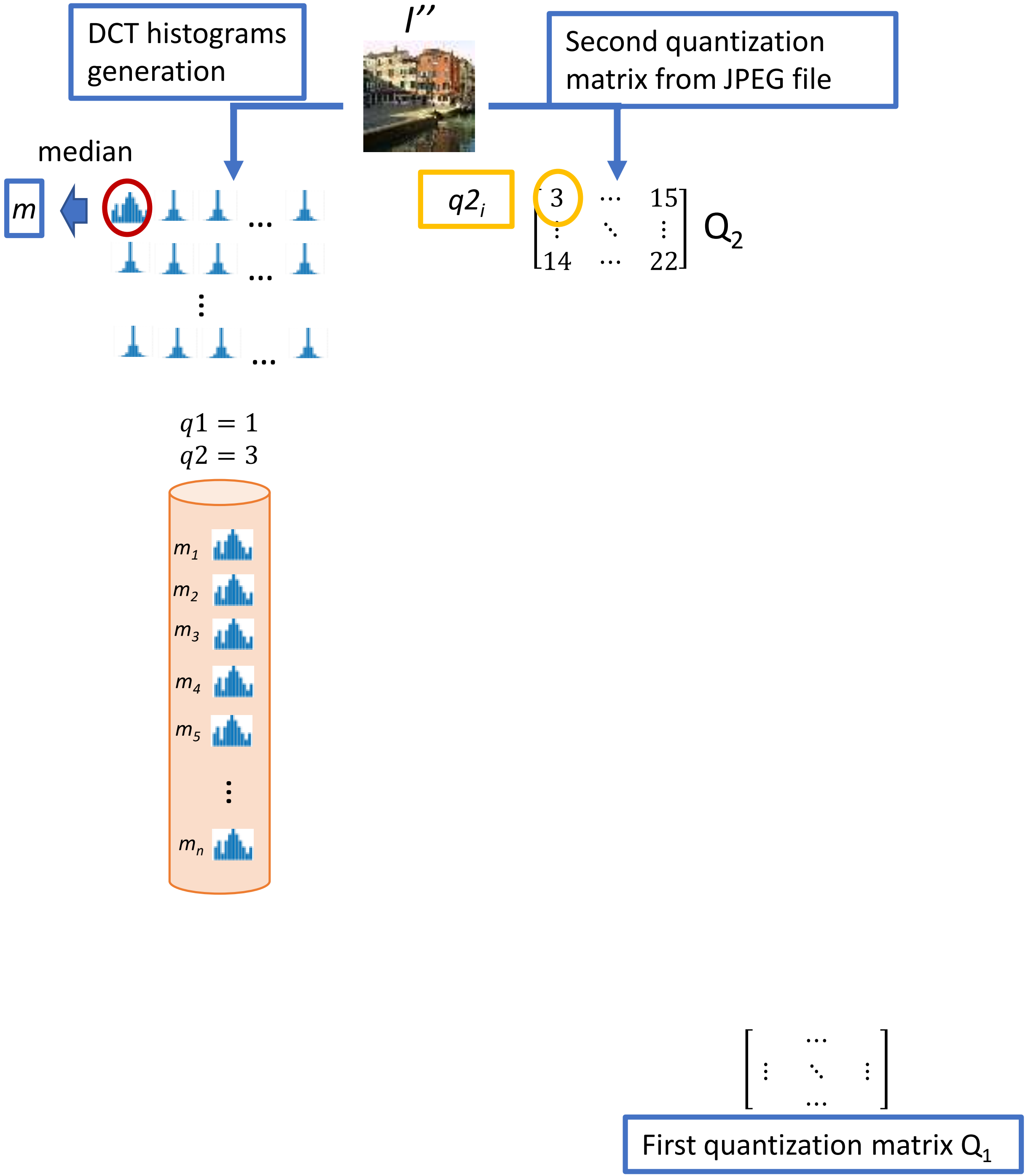
Algorithm 1 The Proposed FQE Technique

Input: double compressed image I''

Output: $\{q1_1, q1_2, \ldots, q1_k\}$

Initialization : k, $q1_{max}$

for $i = 1$ to k **do**
  h_i : distribution of i-th DCT coefficient
  if $(i = 1)$ **then**
    $D : DC_{dset}$
    m : median value of h_i
  else
    $D : AC_{dset}$
    β : β fitted on Laplacian h_i
  end if
  $q2_i$: quantization factor of Q_2 for i-th DCT
  for $j = 1$ to $q1_{max}$ **do**
    $D_{j,q2_i}$: sub-dataset $(q1, q2)$ with $q1 = j$, $q2 = q2_i$
    $D_{j,q2_i}(m, \beta)$: sub-range with most similar m, β
    $d_{i,j}$: lower χ^2 distance between h_i and $D_{j,q2_i}$
  end for
  $q1_i$: $\arg\min_{\{d_{i,j}\}}$, $j \in \{1, 2, \ldots, q1_{max}\}$
end for
regularize($\{q1_1, q1_2, \ldots, q1_k\}$)
return $\{q1_1, q1_2, \ldots, q1_k\}$

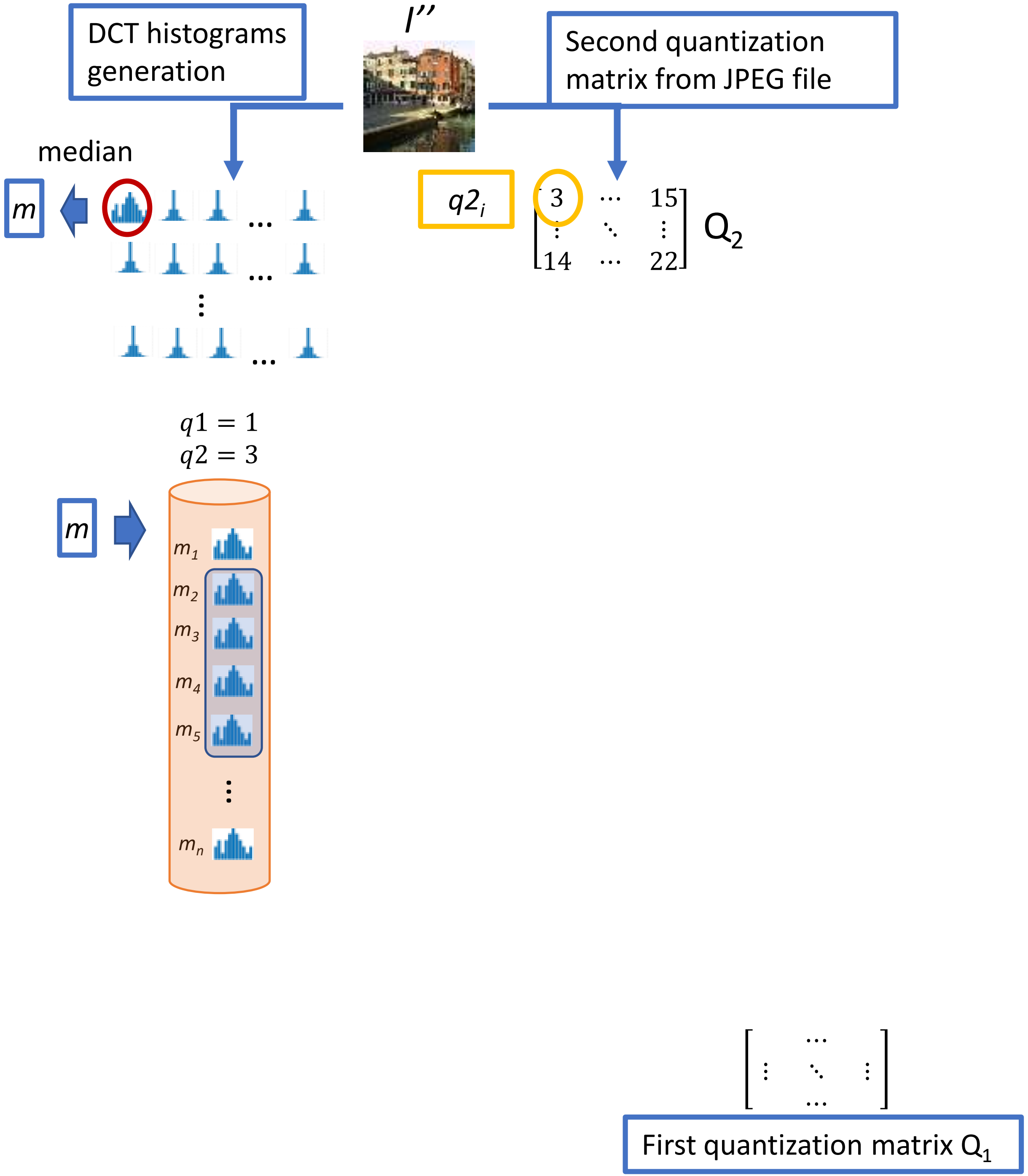


Algorithm 1 The Proposed FQE Technique

Input: double compressed image I''
Output: $\{q1_1, q1_2, \ldots, q1_k\}$

    Initialization : k, $q1_{max}$

for $i = 1$ to k **do**
    h_i : distribution of i-th DCT coefficient
    if $(i = 1)$ **then**
        $D : DC_{dset}$
        m : median value of h_i
    else
        $D : AC_{dset}$
        β : β fitted on Laplacian h_i
    end if
    $q2_i$: quantization factor of Q_2 for i-th DCT
    for $j = 1$ to $q1_{max}$ **do**
        $D_{j,q2_i}$: sub-dataset $(q1, q2)$ with $q1 = j$, $q2 = q2_i$
        $D_{j,q2_i}(m, \beta)$: sub-range with most similar m, β
        $d_{i,j}$: lower χ^2 distance between h_i and $D_{j,q2_i}$
    end for
    $q1_i$: $\arg\min_{\{d_{i,j}\}}$, $j \in \{1, 2, \ldots, q1_{max}\}$
end for
regularize($\{q1_1, q1_2, \ldots, q1_k\}$)
return $\{q1_1, q1_2, \ldots, q1_k\}$

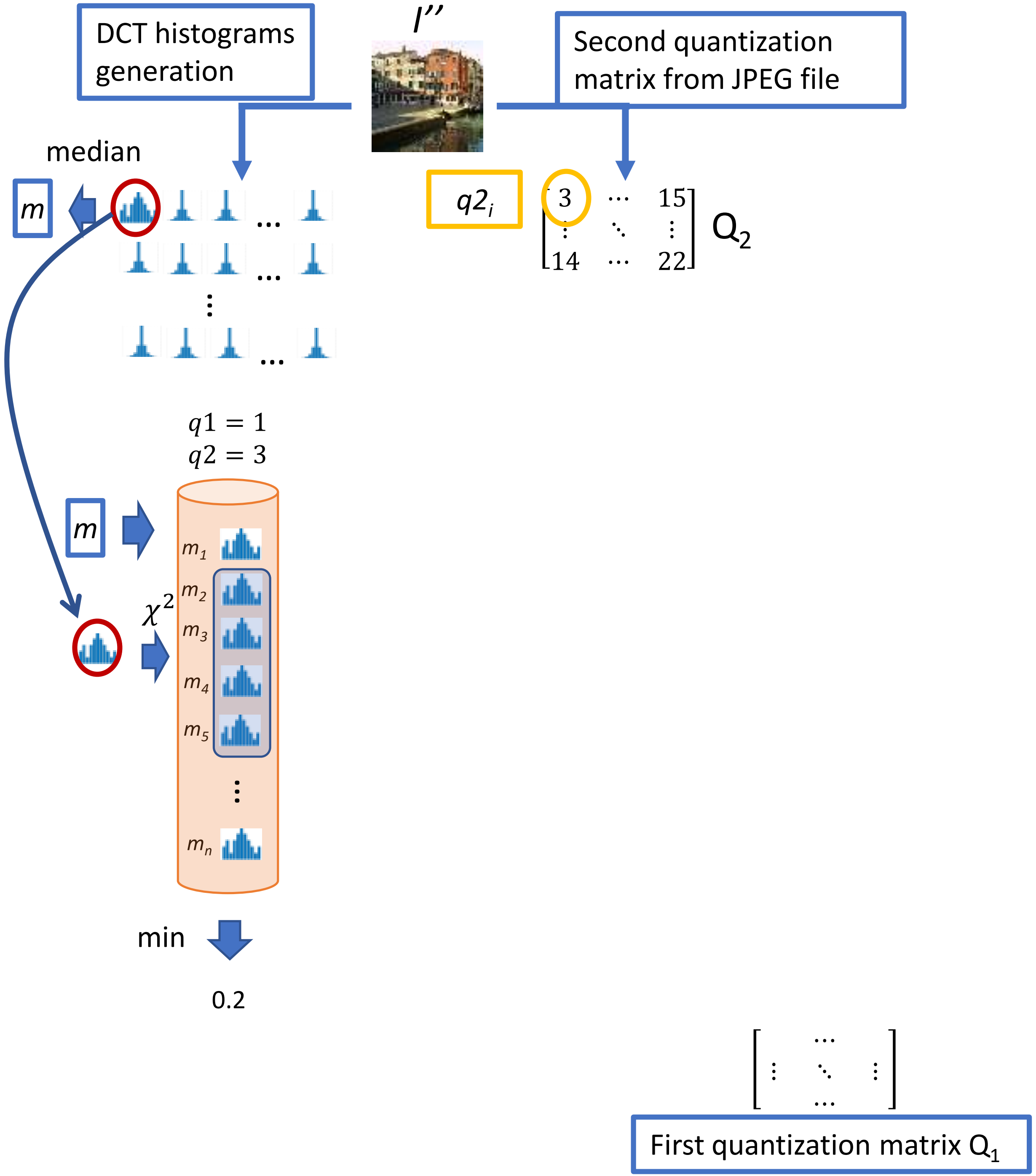


Algorithm 1 The Proposed FQE Technique

Input: double compressed image I''

Output: $\{q1_1, q1_2, \ldots, q1_k\}$

```
    Initialization : k, q1_max
for i = 1 to k do
    h_i : distribution of i-th DCT coefficient
    if (i = 1) then
        D : DC_dset
        m : median value of h_i
    else
        D : AC_dset
        β : β fitted on Laplacian h_i
    end if
    q2_i : quantization factor of Q_2 for i-th DCT
    for j = 1 to q1_max do
        D_{j,q2_i} : sub-dataset (q1, q2) with q1 = j, q2 = q2_i
        D_{j,q2_i}(m, β) : sub-range with most similar m, β
        d_{i,j} : lower χ^2 distance between h_i and D_{j,q2_i}
    end for
    q1_i : arg min_{d_{i,j}}, j ∈ {1, 2, ..., q1_max}
end for
regularize({q1_1, q1_2, ..., q1_k})
return {q1_1, q1_2, ..., q1_k}
```

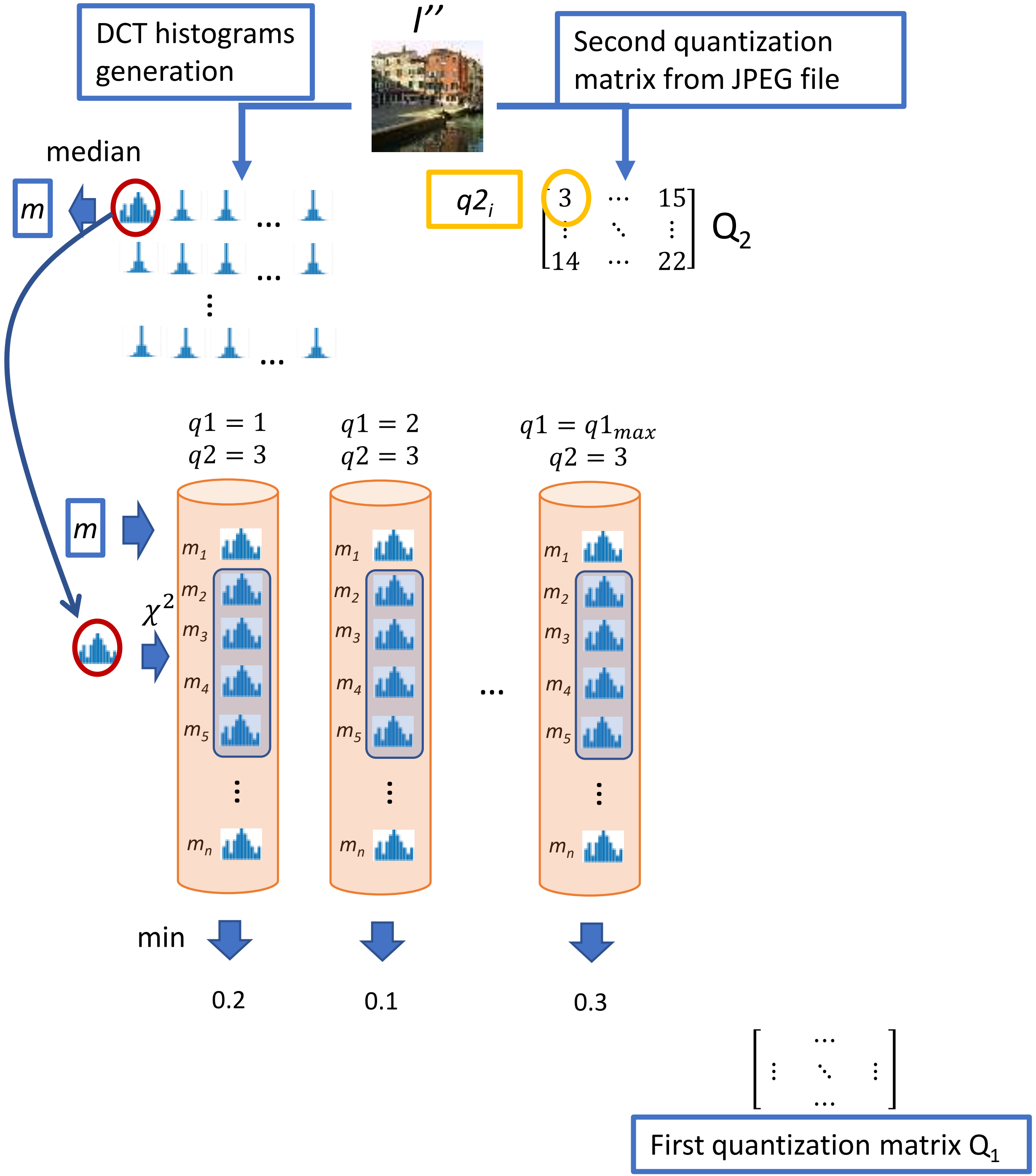


Algorithm 1 The Proposed FQE Technique

Input: double compressed image I''
Output: $\{q1_1, q1_2, \ldots, q1_k\}$

Initialization : k, $q1_{max}$

for $i = 1$ to k **do**
    h_i : distribution of i-th DCT coefficient
    if $(i = 1)$ **then**
        $D : DC_{dset}$
        m : median value of h_i
    else
        $D : AC_{dset}$
        β : β fitted on Laplacian h_i
    end if
    $q2_i$: quantization factor of Q_2 for i-th DCT
    for $j = 1$ to $q1_{max}$ **do**
        $D_{j,q2_i}$: sub-dataset $(q1, q2)$ with $q1 = j$, $q2 = q2_i$
        $D_{j,q2_i}(m, \beta)$: sub-range with most similar m, β
        $d_{i,j}$: lower χ^2 distance between h_i and $D_{j,q2_i}$
    end for
    $q1_i$: $\arg\min_{\{d_{i,j}\}}$, $j \in \{1, 2, \ldots, q1_{max}\}$
end for
regularize($\{q1_1, q1_2, \ldots, q1_k\}$)
return $\{q1_1, q1_2, \ldots, q1_k\}$



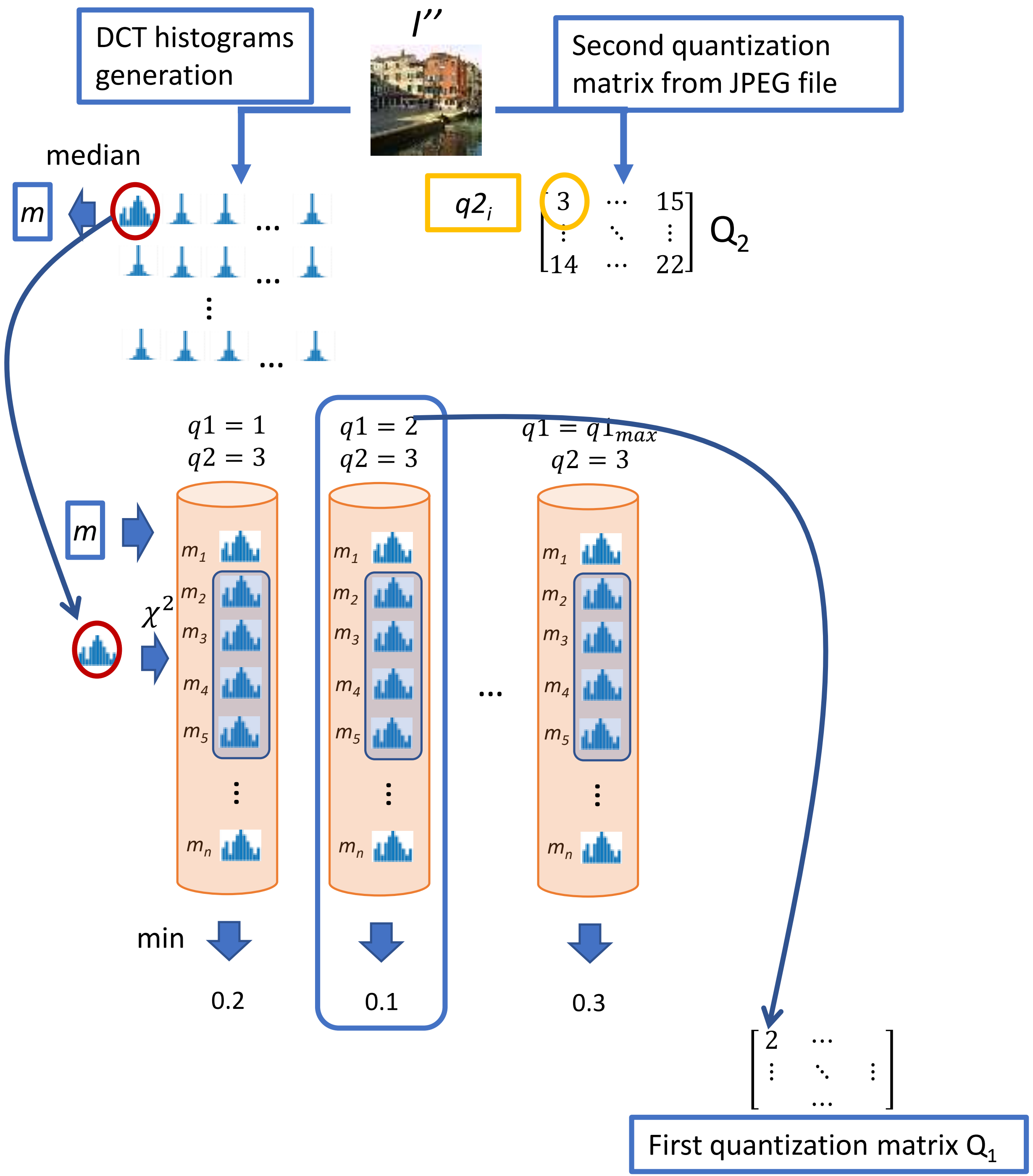
Algorithm 1 The Proposed FQE Technique

Input: double compressed image I''

Output: $\{q1_1, q1_2, \ldots, q1_k\}$

Initialization : k, $q1_{max}$

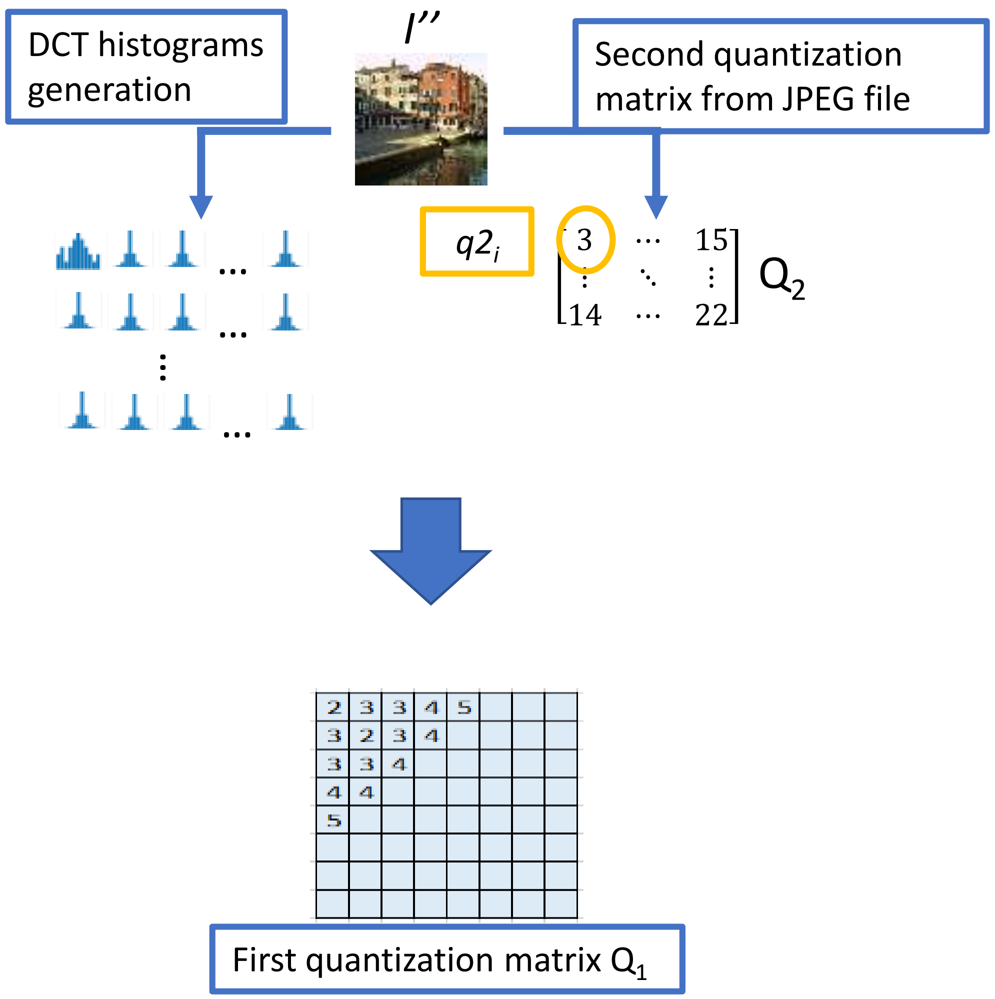
for $i = 1$ to k **do**
  h_i : distribution of i-th DCT coefficient
  if $(i = 1)$ **then**
    $D : DC_{dset}$
    m : median value of h_i
  else
    $D : AC_{dset}$
    β : β fitted on Laplacian h_i
  end if
  $q2_i$: quantization factor of Q_2 for i-th DCT
  for $j = 1$ to $q1_{max}$ **do**
    $D_{j,q2_i}$: sub-dataset $(q1, q2)$ with $q1 = j$, $q2 = q2_i$
    $D_{j,q2_i}(m, \beta)$: sub-range with most similar m, β
    $d_{i,j}$: lower χ^2 distance between h_i and $D_{j,q2_i}$
  end for
  $q1_i : \arg\min_{\{d_{i,j}\}}, j \in \{1, 2, \ldots, q1_{max}\}$
end for
regularize($\{q1_1, q1_2, \ldots, q1_k\}$)
return $\{q1_1, q1_2, \ldots, q1_k\}$



Algorithm 1 The Proposed FQE Technique

Input: double compressed image I''
Output: $\{q1_1, q1_2, \ldots, q1_k\}$
Initialization : k, $q1_{max}$

for $i = 1$ to k **do**
  h_i : distribution of i-th DCT coefficient
  if $(i = 1)$ **then**
    $D : DC_{dset}$
    m : median value of h_i
  else
    $D : AC_{dset}$
    β : β fitted on Laplacian h_i
  end if
  $q2_i$: quantization factor of Q_2 for i-th DCT
  for $j = 1$ to $q1_{max}$ **do**
    $D_{j,q2_i}$: sub-dataset $(q1, q2)$ with $q1 = j$, $q2 = q2_i$
    $D_{j,q2_i}(m, \beta)$: sub-range with most similar m, β
    $d_{i,j}$: lower χ^2 distance between h_i and $D_{j,q2_i}$
  end for
  $q1_i$: $\arg\min_{\{d_{i,j}\}}$, $j \in \{1, 2, \ldots, q1_{max}\}$
end for
regularize($\{q1_1, q1_2, \ldots, q1_k\}$)
return $\{q1_1, q1_2, \ldots, q1_k\}$

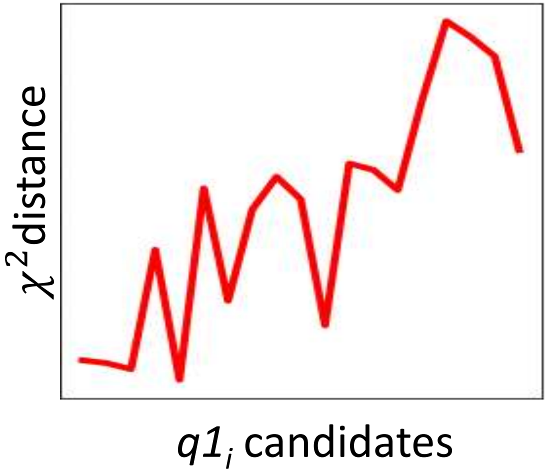


Regularization

- Sometimes, the information contained in h_i does not clearly allow the discrimination among the possible $q1_i$ candidates.

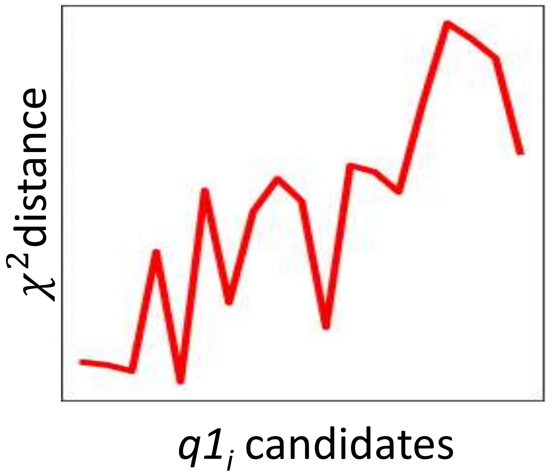
Regularization

- Sometimes, the information contained in h_i does not clearly allow the discrimination among the possible $q1_i$ candidates.



Regularization

- Sometimes, the information contained in h_i does not clearly allow the discrimination among the possible $q1_i$ candidates.



- A strong minimum is not always present at varying of $q1_i$ candidates.

Regularization

- Data coming from neighbors DCT coefficients can be exploited.

J. Park, D. Cho, W. Ahn, and H. Lee, “Double JPEG detection in mixed JPEG quality factors using deep convolutional neural network,” in The European Conference on Computer Vision (ECCV), September 2018.

Regularization

- Data coming from neighbors DCT coefficients can be exploited.

- More specifically, we start from the empirical hypothesis that a generic $q1_i$ value is usually close to $q1_{i-1}$ and $q1_{i+1}$.

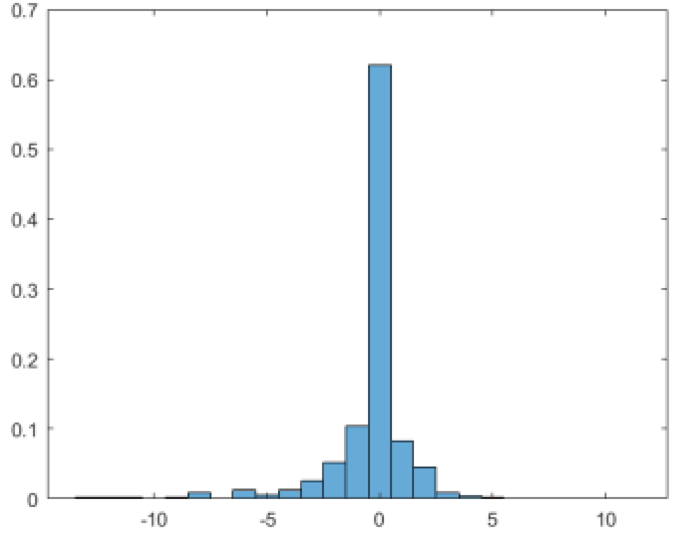
J. Park, D. Cho, W. Ahn, and H. Lee, "Double JPEG detection in mixed JPEG quality factors using deep convolutional neural network," in The European Conference on Computer Vision (ECCV), September 2018.

Regularization

- Data coming from neighbors DCT coefficients can be exploited.

- More specifically, we start from the empirical hypothesis that a generic $q1_i$ value is usually close to $q1_{i-1}$ and $q1_{i+1}$.

Distribution of $q1_i$-$q1_{i+1}$built considering custom tables from Park et al. and $q1_{max}$<22.



J. Park, D. Cho, W. Ahn, and H. Lee, “Double JPEG detection in mixed JPEG quality factors using deep convolutional neural network,” in The European Conference on Computer Vision (ECCV), September 2018.

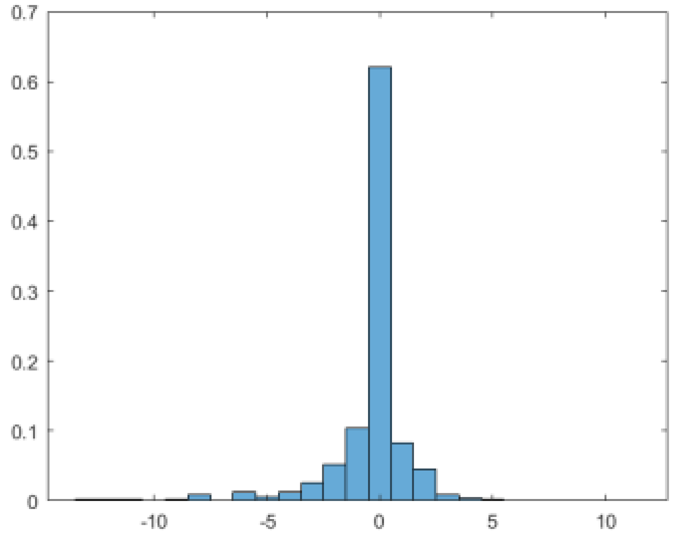
Regularization

- Data coming from neighbors DCT coefficients can be exploited.

- More specifically, we start from the empirical hypothesis that a generic $q1_i$ value is usually close to $q1_{i-1}$ and $q1_{i+1}$.

- Instead of estimating each coefficient independently, three consecutive elements in zig-zag order are considered.

Distribution of $q1_i$-$q1_{i+1}$built considering custom tables from Park et al. and $q1_{max}$<22.



J. Park, D. Cho, W. Ahn, and H. Lee, “Double JPEG detection in mixed JPEG quality factors using deep convolutional neural network,” in The European Conference on Computer Vision (ECCV), September 2018.

Regularization

- Considering $q1_{max}$ as the maximum $q1$ value to be estimated, $q1_{max} \times q1_{max} \times q1_{max}$ first quantization factors combinations are taken into account.

Regularization

- Considering $q1_{max}$ as the maximum $q1$ value to be estimated, $q1_{max} \times q1_{max} \times q1_{max}$ first quantization factors combinations are taken into account.

- A proper score S is then obtained by a weighted average between a data term (C_{data}) and a regularization term (C_{reg}) as follows:

$$S = wC_{data} + (1 - w)C_{reg} \quad w \epsilon [0,1]$$

Regularization

- Considering $q1_{max}$ as the maximum $q1$ value to be estimated, $q1_{max} \times q1_{max} \times q1_{max}$ first quantization factors combinations are taken into account.

- A proper score S is then obtained by a weighted average between a data term (C_{data}) and a regularization term (C_{reg}) as follows:

$$S = wC_{data} + (1 - w)C_{reg} \quad w\epsilon[0,1]$$

$$C_{reg1} = \frac{|c_i - c_{i+1}| + |c_i - c_{i-1}|}{2}$$

$$C_{reg2} = \frac{|c_i - c_{i+1}| + |c_i - c_{i-1}|}{2\sqrt{c_i}}$$

$$C_{reg3} = \frac{|c_i - c_{i+1}| + |c_i - c_{i-1}|}{2c_i}$$

Regularization

- Considering $q1_{max}$ as the maximum $q1$ value to be estimated, $q1_{max} \times q1_{max} \times q1_{max}$ first quantization factors combinations are taken into account.

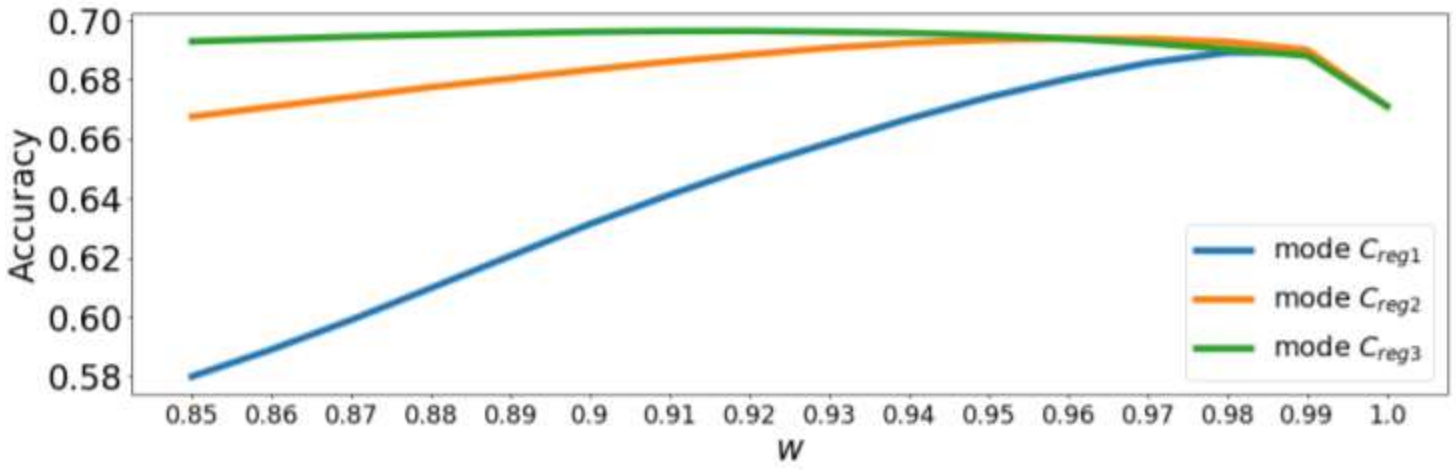
- A proper score S is then obtained by a weighted average between a data term (C_{data}) and a regularization term (C_{reg}) as follows:

$$S = wC_{data} + (1-w)C_{reg} \quad w\epsilon[0,1]$$

$$C_{reg1} = \frac{|c_i - c_{i+1}| + |c_i - c_{i-1}|}{2}$$

$$C_{reg2} = \frac{|c_i - c_{i+1}| + |c_i - c_{i-1}|}{2\sqrt{c_i}}$$

$$C_{reg3} = \frac{|c_i - c_{i+1}| + |c_i - c_{i-1}|}{2c_i}$$



Regularization

- Considering $q1_{max}$ as the maximum $q1$ value to be estimated, $q1_{max} \times q1_{max} \times q1_{max}$ first quantization factors combinations are taken into account.

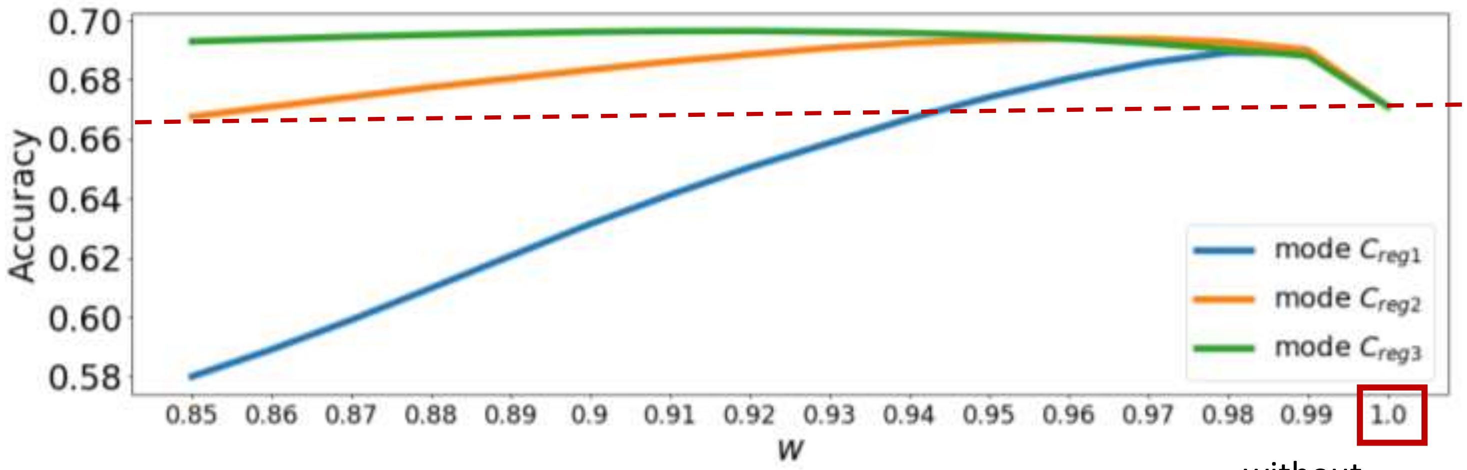
- A proper score S is then obtained by a weighted average between a data term (C_{data}) and a regularization term (C_{reg}) as follows:

$$S = wC_{data} + (1-w)C_{reg} \quad w\epsilon[0,1]$$

$$C_{reg1} = \frac{|c_i - c_{i+1}| + |c_i - c_{i-1}|}{2}$$

$$C_{reg2} = \frac{|c_i - c_{i+1}| + |c_i - c_{i-1}|}{2\sqrt{c_i}}$$

$$C_{reg3} = \frac{|c_i - c_{i+1}| + |c_i - c_{i-1}|}{2c_i}$$



without regularization

Regularization

- Considering $q1_{max}$ as the maximum $q1$ value to be estimated, $q1_{max} \times q1_{max} \times q1_{max}$ first quantization factors combinations are taken into account.

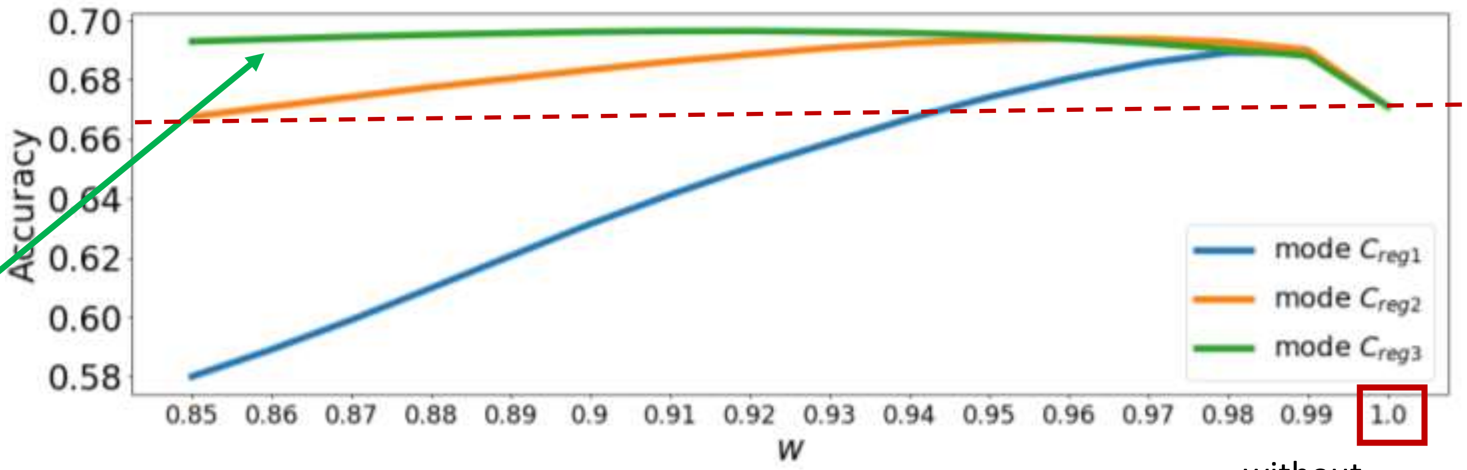
- A proper score S is then obtained by a weighted average between a data term (C_{data}) and a regularization term (C_{reg}) as follows:

$$S = wC_{data} + (1-w)C_{reg} \quad w\epsilon[0,1]$$

$$C_{reg1} = \frac{|c_i - c_{i+1}| + |c_i - c_{i-1}|}{2}$$

$$C_{reg2} = \frac{|c_i - c_{i+1}| + |c_i - c_{i-1}|}{2\sqrt{c_i}}$$

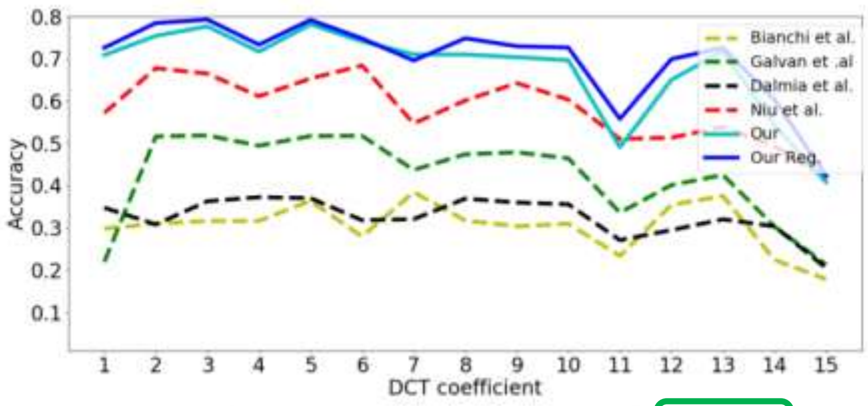
$$\boxed{C_{reg3} = \frac{|c_i - c_{i+1}| + |c_i - c_{i-1}|}{2c_i}}$$



without regularization

Experimental results (comparisons)

Standard Matrices

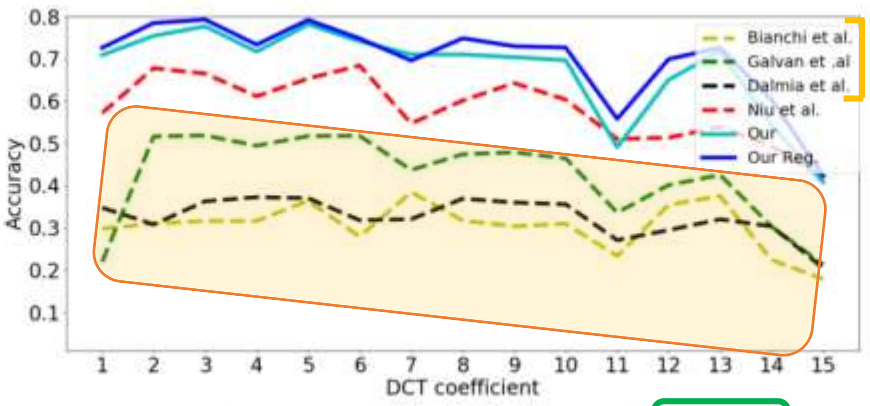


a. $QF_1 \in \{55, 60, 65, 70, 75, 80, 85, 90, 95, 98\}$, $\boxed{QF_2 = 90}$

T. Bianchi and A. Piva, "Image forgery localization via block-grained analysis of JPEG artifacts," Proc. of IEEE Trans. on Information Forensics and Security, vol. 7, no. 3, p. 1003, 2012.
F. Galvan, G. Puglisi, A. R. Bruna, and S. Battiato, "First quantization matrix estimation from double compressed JPEG images," IEEE Trans. on Information Forensics and Security, vol. 9, no. 8, pp. 1299–1310, 2014.
N. Dalmia and M. Okade, "First quantization matrix estimation for double compressed JPEG images utilizing novel dct histogram selection strategy," in Proc. of the Tenth Indian Conference on Computer Vision, Graphics and Image Processing, 2016, pp. 1–8.
Y. Niu, B. Tondi, Y. Zhao, and M. Barni, "Primary quantization matrix estimation of double compressed JPEG images via CNN," IEEE Signal Processing Letters, vol. 27, pp. 191–195, 2020.

Experimental results (comparisons)

Standard Matrices



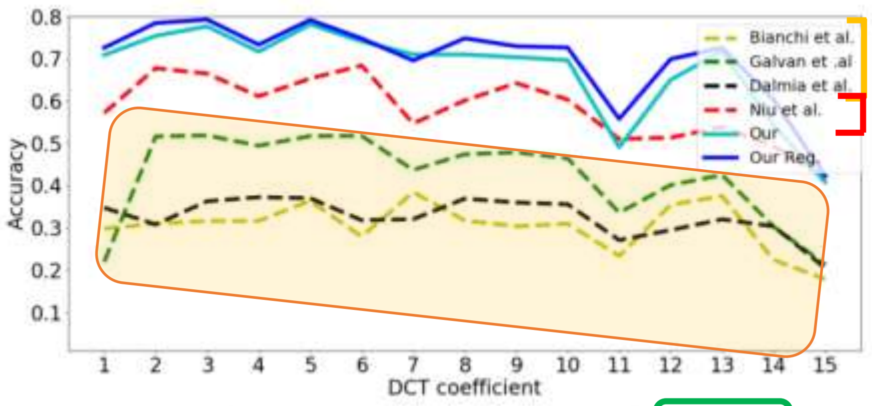
Classical approaches

a. $QF_1 \in \{55, 60, 65, 70, 75, 80, 85, 90, 95, 98\}$, $\boxed{QF_2 = 90}$

T. Bianchi and A. Piva, "Image forgery localization via block-grained analysis of JPEG artifacts," Proc. of IEEE Trans. on Information Forensics and Security, vol. 7, no. 3, p. 1003, 2012.
F. Galvan, G. Puglisi, A. R. Bruna, and S. Battiato, "First quantization matrix estimation from double compressed JPEG images," IEEE Trans. on Information Forensics and Security, vol. 9, no. 8, pp. 1299–1310, 2014.
N. Dalmia and M. Okade, "First quantization matrix estimation for double compressed JPEG images utilizing novel dct histogram selection strategy," in Proc. of the Tenth Indian Conference on Computer Vision, Graphics and Image Processing, 2016, pp. 1–8.
Y. Niu, B. Tondi, Y. Zhao, and M. Barni, "Primary quantization matrix estimation of double compressed JPEG images via CNN," IEEE Signal Processing Letters, vol. 27, pp. 191–195, 2020.

Experimental results (comparisons)

Standard Matrices



a. $QF_1 \in \{55, 60, 65, 70, 75, 80, 85, 90, 95, 98\}$, $\boxed{QF_2 = 90}$

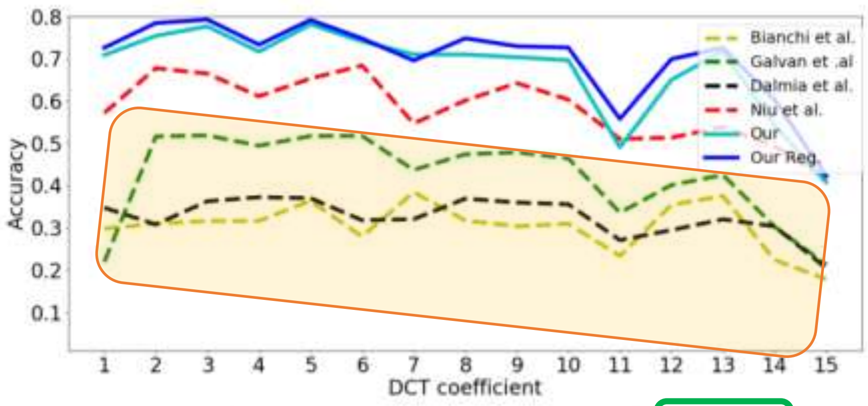
Classical approaches

CNN based approach

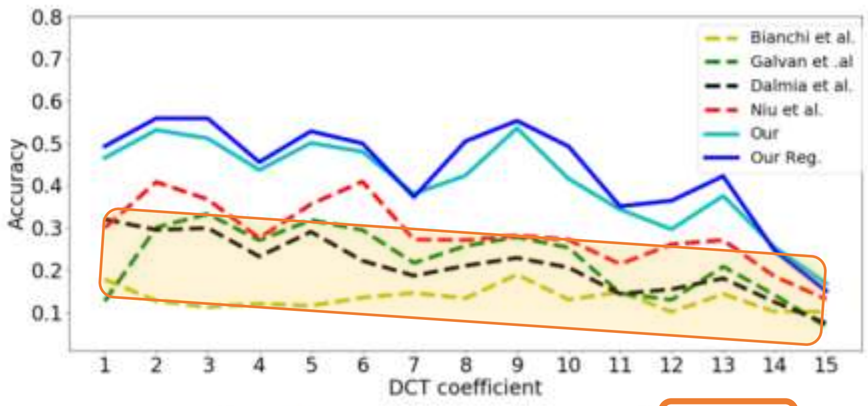
T. Bianchi and A. Piva, "Image forgery localization via block-grained analysis of JPEG artifacts," Proc. of IEEE Trans. on Information Forensics and Security, vol. 7, no. 3, p. 1003, 2012.
F. Galvan, G. Puglisi, A. R. Bruna, and S. Battiato, "First quantization matrix estimation from double compressed JPEG images," IEEE Trans. on Information Forensics and Security, vol. 9, no. 8, pp. 1299–1310, 2014.
N. Dalmia and M. Okade, "First quantization matrix estimation for double compressed JPEG images utilizing novel dct histogram selection strategy," in Proc. of the Tenth Indian Conference on Computer Vision, Graphics and Image Processing, 2016, pp. 1–8.
Y. Niu, B. Tondi, Y. Zhao, and M. Barni, "Primary quantization matrix estimation of double compressed JPEG images via CNN," IEEE Signal Processing Letters, vol. 27, pp. 191–195, 2020.

Experimental results (comparisons)

Standard Matrices



a. $QF_1 \in \{55, 60, 65, 70, 75, 80, 85, 90, 95, 98\}$, $\boxed{QF_2 = 90}$

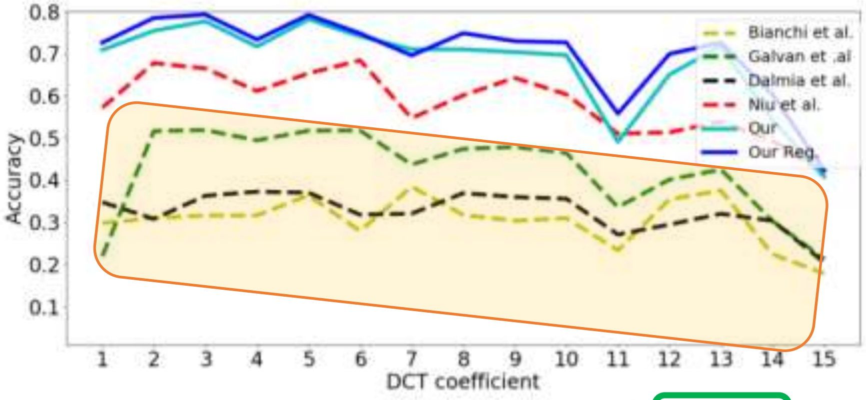


b. $QF_1 \in \{55, 60, 65, 70, 75, 80, 85, 90, 95, 98\}$, $\boxed{QF_2 = 80}$

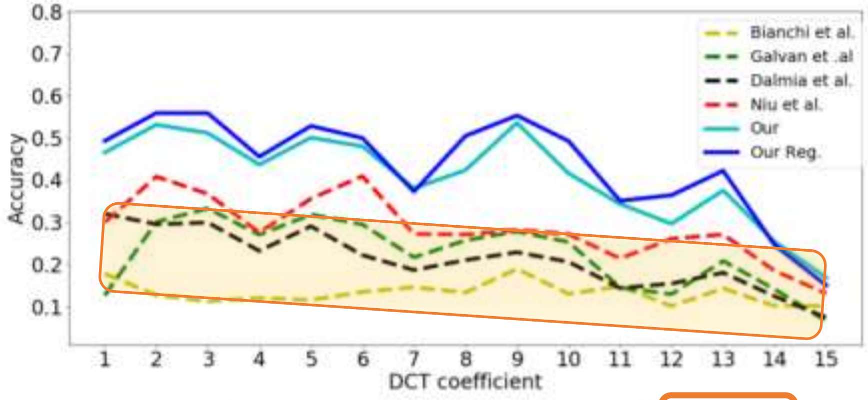
T. Bianchi and A. Piva, "Image forgery localization via block-grained analysis of JPEG artifacts," Proc. of IEEE Trans. on Information Forensics and Security, vol. 7, no. 3, p. 1003, 2012.
F. Galvan, G. Puglisi, A. R. Bruna, and S. Battiato, "First quantization matrix estimation from double compressed JPEG images," IEEE Trans. on Information Forensics and Security, vol. 9, no. 8, pp. 1299–1310, 2014.
N. Dalmia and M. Okade, "First quantization matrix estimation for double compressed JPEG images utilizing novel dct histogram selection strategy," in Proc. of the Tenth Indian Conference on Computer Vision, Graphics and Image Processing, 2016, pp. 1–8.
Y. Niu, B. Tondi, Y. Zhao, and M. Barni, "Primary quantization matrix estimation of double compressed JPEG images via CNN," IEEE Signal Processing Letters, vol. 27, pp. 191–195, 2020.

Experimental results (comparisons)

Standard Matrices

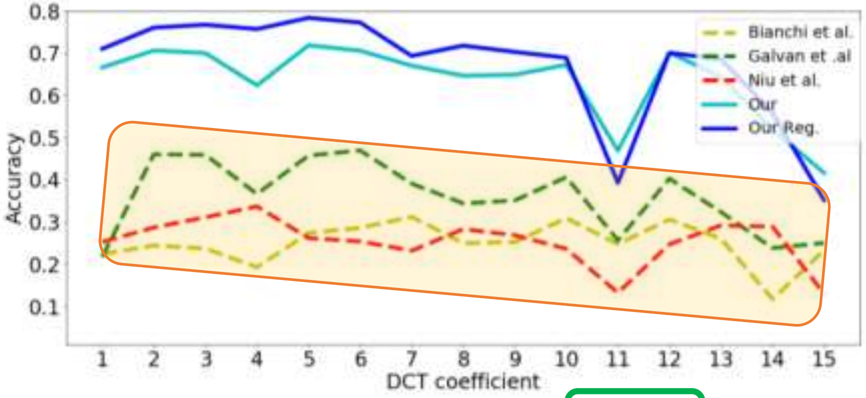


a. $QF_1 \in \{55, 60, 65, 70, 75, 80, 85, 90, 95, 98\}$, $\boxed{QF_2 = 90}$

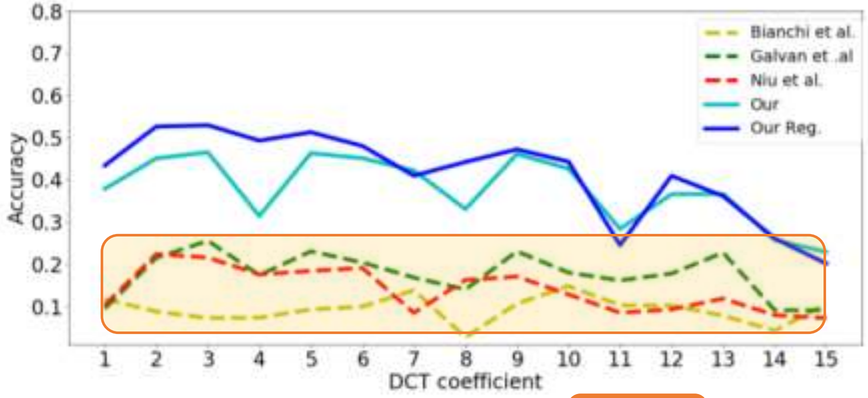


b. $QF_1 \in \{55, 60, 65, 70, 75, 80, 85, 90, 95, 98\}$, $\boxed{QF_2 = 80}$

Photoshop Custom Matrices



c. $Q_1 \in \{5, 6, 7, 8, 9, 10, 11, 12\}$, $\boxed{QF_2 = 90}$



d. $Q_1 \in \{5, 6, 7, 8, 9, 10, 11, 12\}$, $\boxed{QF_2 = 80}$

T. Bianchi and A. Piva, "Image forgery localization via block-grained analysis of JPEG artifacts," Proc. of IEEE Trans. on Information Forensics and Security, vol. 7, no. 3, p. 1003, 2012.
F. Galvan, G. Puglisi, A. R. Bruna, and S. Battiato, "First quantization matrix estimation from double compressed JPEG images," IEEE Trans. on Information Forensics and Security, vol. 9, no. 8, pp. 1299–1310, 2014.
N. Dalmia and M. Okade, "First quantization matrix estimation for double compressed JPEG images utilizing novel dct histogram selection strategy," in Proc. of the Tenth Indian Conference on Computer Vision, Graphics and Image Processing, 2016, pp. 1–8.
Y. Niu, B. Tondi, Y. Zhao, and M. Barni, "Primary quantization matrix estimation of double compressed JPEG images via CNN," IEEE Signal Processing Letters, vol. 27, pp. 191–195, 2020.

Experimental results (custom tables)

Park et al. collected a dataset of JPEG quantization matrices employed in real scenarios.

J. Park, D. Cho, W. Ahn, and H. Lee, "Double JPEG detection in mixed JPEG quality factors using deep convolutional neural network," in The European Conference on Computer Vision (ECCV), September 2018.
D. Dang-Nguyen, C. Pasquini, V. Conotter, and G. Boato, "Raise: a raw images dataset for digital image forensics," in Proc. of the 6th ACM Multimedia Systems Conference, 2015, pp. 219–224.
G. Schaefer and M. Stich, "UCID: An uncompressed color image database," in Storage and Retrieval Methods and Applications for Multimedia 2004, vol. 5307. International Society for Optics and Photonics, 2003, pp. 472–480.

Experimental results (custom tables)

Park et al. collected a dataset of JPEG quantization matrices employed in real scenarios.

To evaluate the generalization capability of the proposed solution a series of tests have been conducted considering these quantization matrices divided into three groups: Low, Mid, High.

J. Park, D. Cho, W. Ahn, and H. Lee, “Double JPEG detection in mixed JPEG quality factors using deep convolutional neural network,” in The European Conference on Computer Vision (ECCV), September 2018.
D. Dang-Nguyen, C. Pasquini, V. Conotter, and G. Boato, “Raise: a raw images dataset for digital image forensics,” in Proc. of the 6th ACM Multimedia Systems Conference, 2015, pp. 219–224.
G. Schaefer and M. Stich, “UCID: An uncompressed color image database,” in Storage and Retrieval Methods and Applications for Multimedia 2004, vol. 5307. International Society for Optics and Photonics, 2003, pp. 472–480.

Experimental results (custom tables)

Park et al. collected a dataset of JPEG quantization matrices employed in real scenarios.

To evaluate the generalization capability of the proposed solution a series of tests have been conducted considering these quantization matrices divided into three groups: Low, Mid, High.

Method	Dataset	Cropped Patch	Low/Low	Low/Mid	Low/High	Mid/Low	Mid/Mid	Mid/High	High/Low	High/Mid	High/High	Mean
Our	RAISE	64 × 64	0.25	0.47	0.79	0.17	0.32	0.82	**0.27**	0.31	0.70	0.46
Our Reg.			**0.30**	**0.53**	**0.81**	**0.22**	**0.37**	**0.84**	0.25	**0.33**	**0.75**	**0.49**
Our	UCID	64 × 64	0.33	0.63	0.93	0.20	0.39	0.90	**0.15**	0.21	0.66	0.49
Our Reg.			**0.36**	**0.65**	**0.96**	**0.23**	**0.42**	**0.91**	0.13	**0.23**	**0.73**	**0.51**
Our	RAISE	128 × 128	0.36	**0.60**	0.85	0.29	0.44	0.87	**0.25**	0.32	0.74	0.52
Our Reg.			**0.41**	0.55	**0.88**	**0.34**	**0.48**	**0.89**	**0.25**	**0.38**	**0.79**	**0.55**
Our	UCID	128 × 128	0.47	0.76	0.96	0.31	0.49	**0.94**	0.18	0.29	0.74	0.57
Our Reg.			**0.50**	**0.79**	**0.96**	**0.35**	**0.51**	0.93	**0.20**	**0.34**	**0.79**	**0.60**
Our	RAISE	256 × 256	0.45	0.69	0.88	0.38	0.52	0.89	0.25	0.36	0.77	0.58
Our Reg.			**0.49**	**0.73**	**0.90**	**0.40**	**0.55**	**0.90**	**0.30**	**0.45**	**0.82**	**0.62**
Our	UCID	256 × 256	0.56	0.83	**0.98**	0.44	0.57	**0.96**	0.23	0.34	0.77	0.63
Our Reg.			**0.60**	**0.85**	0.97	**0.48**	**0.60**	**0.96**	**0.28**	**0.43**	**0.82**	**0.67**
Our	RAISE	512 × 512	**0.50**	0.74	0.91	0.44	0.57	0.91	0.26	0.38	0.77	0.61
Our Reg.			**0.50**	**0.78**	**0.92**	**0.48**	**0.59**	**0.92**	**0.32**	**0.48**	**0.83**	**0.65**
Our	UCID	Full size	0.63	0.86	**0.98**	0.51	**0.62**	**0.96**	0.27	0.39	0.80	0.66
Our Reg.			**0.67**	**0.87**	0.97	**0.56**	**0.62**	**0.96**	**0.37**	**0.49**	**0.85**	**0.71**

J. Park, D. Cho, W. Ahn, and H. Lee, "Double JPEG detection in mixed JPEG quality factors using deep convolutional neural network," in The European Conference on Computer Vision (ECCV), September 2018.
D. Dang-Nguyen, C. Pasquini, V. Conotter, and G. Boato, "Raise: a raw images dataset for digital image forensics," in Proc. of the 6th ACM Multimedia Systems Conference, 2015, pp. 219–224.
G. Schaefer and M. Stich, "UCID: An uncompressed color image database," in Storage and Retrieval Methods and Applications for Multimedia 2004, vol. 5307. International Society for Optics and Photonics, 2003, pp. 472–480.

Experimental results (custom tables)

Park et al. collected a dataset of JPEG quantization matrices employed in real scenarios.

To evaluate the generalization capability of the proposed solution a series of tests have been conducted considering these quantization matrices divided into three groups: Low, Mid, High.

[0,0.33[
[0.33,0.66[
[0.66,1]

Method	Dataset	Cropped Patch	Low/Low	Low/Mid	Low/High	Mid/Low	Mid/Mid	Mid/High	High/Low	High/Mid	High/High	Mean
Our	RAISE	64 × 64	0.25	0.47	0.79	0.17	0.32	0.82	**0.27**	0.31	0.70	0.46
Our Reg.			**0.30**	**0.53**	**0.81**	**0.22**	**0.37**	**0.84**	0.25	**0.33**	**0.75**	**0.49**
Our	UCID	64 × 64	0.33	0.63	0.93	0.20	0.39	0.90	**0.15**	0.21	0.66	0.49
Our Reg.			**0.36**	**0.65**	**0.96**	**0.23**	**0.42**	**0.91**	0.13	**0.23**	**0.73**	**0.51**
Our	RAISE	128 × 128	0.36	**0.60**	0.85	0.29	0.44	0.87	**0.25**	0.32	0.74	0.52
Our Reg.			**0.41**	0.55	**0.88**	**0.34**	**0.48**	**0.89**	**0.25**	**0.38**	**0.79**	**0.55**
Our	UCID	128 × 128	0.47	0.76	0.96	0.31	0.49	**0.94**	0.18	0.29	0.74	0.57
Our Reg.			**0.50**	**0.79**	**0.96**	**0.35**	**0.51**	0.93	**0.20**	**0.34**	**0.79**	**0.60**
Our	RAISE	256 × 256	0.45	0.69	0.88	0.38	0.52	0.89	0.25	0.36	0.77	0.58
Our Reg.			**0.49**	**0.73**	**0.90**	**0.40**	**0.55**	**0.90**	**0.30**	**0.45**	**0.82**	**0.62**
Our	UCID	256 × 256	0.56	0.83	**0.98**	0.44	0.57	**0.96**	0.23	0.34	0.77	0.63
Our Reg.			**0.60**	**0.85**	0.97	**0.48**	**0.60**	**0.96**	**0.28**	**0.43**	**0.82**	**0.67**
Our	RAISE	512 × 512	**0.50**	0.74	0.91	0.44	0.57	0.91	0.26	0.38	0.77	0.61
Our Reg.			**0.50**	**0.78**	**0.92**	**0.48**	**0.59**	**0.92**	**0.32**	**0.48**	**0.83**	**0.65**
Our	UCID	Full size	0.63	0.86	**0.98**	0.51	**0.62**	**0.96**	0.27	0.39	0.80	0.66
Our Reg.			**0.67**	**0.87**	0.97	**0.56**	**0.62**	**0.96**	**0.37**	**0.49**	**0.85**	**0.71**

J. Park, D. Cho, W. Ahn, and H. Lee, “Double JPEG detection in mixed JPEG quality factors using deep convolutional neural network,” in The European Conference on Computer Vision (ECCV), September 2018.
D. Dang-Nguyen, C. Pasquini, V. Conotter, and G. Boato, “Raise: a raw images dataset for digital image forensics,” in Proc. of the 6th ACM Multimedia Systems Conference, 2015, pp. 219–224.
G. Schaefer and M. Stich, “UCID: An uncompressed color image database,” in Storage and Retrieval Methods and Applications for Multimedia 2004, vol. 5307. International Society for Optics and Photonics, 2003, pp. 472–480.

Conclusions

- In this paper, a new technique able to estimate the first quantization factors for JPEG double compressed images was presented, employing a mixed statistical and Machine Learning approach.

Conclusions

- In this paper, a new technique able to estimate the first quantization factors for JPEG double compressed images was presented, employing a mixed statistical and Machine Learning approach.
- One of the main contributions was the way we employed the big amount of data to avoid overfitting: constant matrices allow us to uncouple $\{q1, q2\}$ and the use of m and β made computational times acceptable.

Conclusions

- In this paper, a new technique able to estimate the first quantization factors for JPEG double compressed images was presented, employing a mixed statistical and Machine Learning approach.
- One of the main contributions was the way we employed the big amount of data to avoid overfitting: constant matrices allow us to uncouple {*q1*, *q2*} and the use of m and β made computational times acceptable.
- The presented solution was demonstrated to work for both custom and standard tables thus being generalizable enough to be employed in real-case scenarios.

Conclusions

- In this paper, a new technique able to estimate the first quantization factors for JPEG double compressed images was presented, employing a mixed statistical and Machine Learning approach.
- One of the main contributions was the way we employed the big amount of data to avoid overfitting: constant matrices allow us to uncouple $\{q1, q2\}$ and the use of m and β made computational times acceptable.
- The presented solution was demonstrated to work for both custom and standard tables thus being generalizable enough to be employed in real-case scenarios.
- Experimental tests showed the goodness of the technique outperforming state-of-the-art results.

Conclusions

- In this paper, a new technique able to estimate the first quantization factors for JPEG double compressed images was presented, employing a mixed statistical and Machine Learning approach.
- One of the main contributions was the way we employed the big amount of data to avoid overfitting: constant matrices allow us to uncouple {*q1*, *q2*} and the use of *m* and β made computational times acceptable.
- The presented solution was demonstrated to work for both custom and standard tables thus being generalizable enough to be employed in real-case scenarios.
- Experimental tests showed the goodness of the technique outperforming state-of-the-art results.
- Finally, the use of 1-nn to learn the distribution underlines rooms for improvement of the proposed method.

Thank you!

In-Depth DCT Coefficient Distribution Analysis for First Quantization Estimation

Sebastiano Battiato[1] battiato@dmi.unict.it
Oliver Giudice[1] giudice@dmi.unict.it
Francesco Guarnera[1] francesco.guarnera@unict.it
Giovanni Puglisi[2] puglisi@unica.it

[1]University of Catania, [2]University of Cagliari