

ISO/IEC JTC 1/SC 29/WG 1  
(ITU-T SG16)

## Coding of Still Pictures

### JBIG

Joint Bi-level Image  
Experts Group

### JPEG

Joint Photographic  
Experts Group

**TITLE:** Description of the JPEG AI Verification Model under  
Consideration and associated software integration procedure

**SOURCE:** ICQ Group

**STATUS:** Final

**REQUESTED ACTION:** Distribute

**DISTRIBUTION:** Public

**Contact:**

ISO/IEC JTC 1/SC 29/WG 1 Convener – Prof. Touradj Ebrahimi  
EPFL/STI/IEL/GR-EB, Station 11, CH-1015 Lausanne, Switzerland  
Tel: +41 21 693 2606, Fax: +41 21 693 7600, E-mail: [Touradj.Ebrahimi@epfl.ch](mailto:Touradj.Ebrahimi@epfl.ch)

# 1 Purpose of this Document

This document reports the decisions taken at 96th JPEG meeting regarding the JPEG AI verification model under consideration (VMuC).

## 2 Verification model under consideration

1. The JPEG AI VMuC corresponds to the combination of TEAM14 + TEAM24 proposals. The VMuC architecture is illustrated in Figure 1 and the list of tools are presented in Table 1.
2. The software integration for VMuC will have the following procedure:
  - 2.1. 1st stage: TEAM14 CfP submission minus decoder RNAB, minus autoregressive context, minus Gaussian Mixture Model (GMM). After this stage, RD performance and complexity assessment results will be provided by enabling/disabling RDOQ, ICCI and ACT.
  - 2.2. 2nd stage: TEAM24 is responsible to include the arithmetic coding engine, decoupled architecture (with wavefront processing), block-based skip mode, mask, scale units MS1, MS2, MS3, adaptive offset and adaptive resampling. If possible, evaluate the effect of the arithmetic coding engine change.
  - 2.3. 3rd stage: Verify the implementation of the codec obtained after 2<sup>nd</sup> stage. Verify that only 4 models are still used to support variable rate functionality. Correct any abnormal behavior.
3. Document every step, namely the addition of any tool in the VMuC creation with RD performance assessment (objective) and complexity analysis according to CTTC.
4. The documentation process is only to understand each tool contribution to the overall performance and thus, the final VMuC will be the same independently of the results provided.

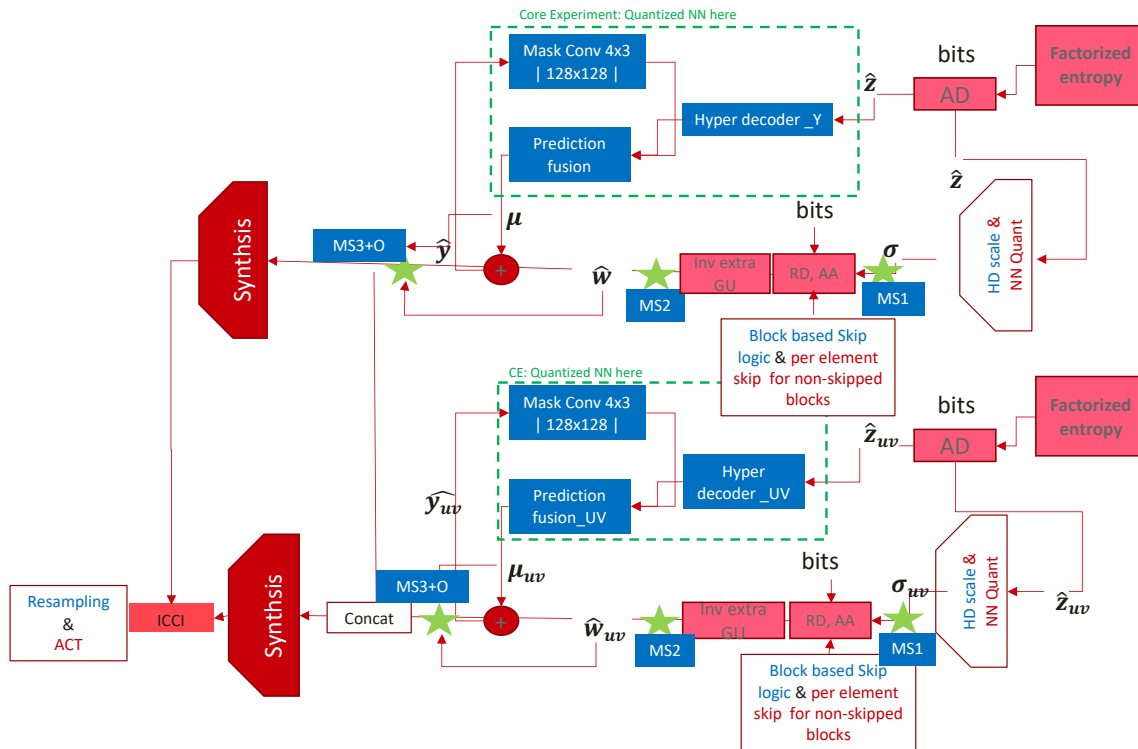


Figure 1 – JPEG AI VMuC high-level architecture.

Table 1 – JPEG AI VMuC list of tools for integration.

<i>Tools to be provided by Team14 that will be integrated in the VMuC</i>	
1.	Colors Separation and Conditional Coding NN architecture with 4 models for all rate points.
2.	Boundary handling for NN-based codec (wg1m 96016, section 2.1.3)
3.	Variable rate support: separate “beta” for Y and UV and extraGU (wg1m96016, section 2.1.5)
4.	Adaptive alphabet (wg1m 96016, section 2.1.7)
5.	RDOQ encoder only tool (wg1m96016 section 2.1.8)
6.	NN quantizing algorithm (wg1m96016 section 2.1.17.2) for device interoperability
7.	Synthesis transform with bug fix (wg1m96083)
8.	Inter Channel Correlation Information filter (ICCI) sub-network (wg1m96016 section 2.1.20)
9.	Overlapping latent space tiles (wg1m96016 section 2.1.21)
10.	Per-element skip logic (wg1m96016 section 2.1.16)
<i>Tools to be provided by Team24 that will be integrated in the VMuC</i>	
1.	Decouple Architecture with wavefront processing
1.1.	Prediction fusion model, context model, hyper scale decoder (wg1m96053 section 3.1)
1.2.	Hyper encoder and hyper decoder (aligned with decouple architecture). Note: in stage 3 quantization of the hyper scale decoder will be performed. Prediction fusion module, context model and hyper decoder module might be quantized depending on the device interoperability needs
1.3.	Wavefront processing (wg1m96053 section 3.2)
2.	Mask, scale and offset operations:
2.1.	Adaptive Quantization (AQ) (wg1m96053 section 3.4)
2.2.	Latent Scaling Before Synthesis (LSBS) (wg1m96053 section 3.5)
2.3.	Latent Domain Adaptive Offset (LDAO) (wg1m96053 section 3.6)
3.	Block-based skip logic (wg1m96053 section 3.7)
4.	Reconstruction resampling (wg1m96053 section 3.8)
5.	Tiling of synthesis transform (wg1m96053 section 3.10). This should be same as wg1m96016 section 2.1.21)
6.	Arithmetic coding engine (wg1m96053 section 3.11)
7.	Latent refiner (wg1m96053 section 5.11). It corresponds to encoder logic.