P-ISSN 2586-9000
E-ISSN 2586-9027
Homepage : *https://tci-thaijo.org/index.php/SciTechAsia* Science & Technology Asia
Vol.xx No.yy xxx - xxx 2026 Page: [1-13]

*Original research article*

# Quadrant Amplitude Representation of Quantum Images Using Hybrid Qubit-Qudit Systems

Received xx Month 2026; Received in revised form xx Month 2026
Accepted xx Month 2026; Available online xx Month 2026

## ABSTRACT

Quantum image representation remains shaped by a basic scaling problem, because higher image resolution usually increases the number of address registers or the cost of controlling a circuit over a larger pixel space. This paper presents QARQI, the Quadrant Amplitude Representation of Quantum Images, as a hybrid qubit-qudit image representation that stores position through two polarity qubits and two finite-dimensional magnitude qudits, while grayscale intensity is encoded in an ancilla qubit through an address-conditioned rotation whose measurement probability recovers the normalized pixel value. The proposed coordinate map assigns every register tuple to one pixel of a $2d \times 2d$ image, the resulting state is shown to be normalized, and the encoding procedure is implemented in Cirq using generalized Hadamard preparation over the qudit registers and controlled $R_y$ rotations over the ancilla. Validation on nine 256-by-256 grayscale images with $d = 128$ used 262,144 to 7,864,320 total measurements per reconstruction and showed monotonic improvement for every image, with mean PSNR rising from 11.4818 dB at 262,144 measurements to 25.7818 dB at 5,111,808 measurements, and with the 7,864,320-measurement reconstructions ranging from 27.1686 to 28.2223 dB. Spatial operation tests on the lake image further showed that horizontal and vertical flips and right-angle rotations can be applied as address-state transformations before measurement, reaching maximum PSNR values between 28.0490 and 28.0956 dB across the operation sweeps.

**Keywords:** Quantum image representation, qudits, amplitude encoding, hybrid systems, complexity analysis

## 1. INTRODUCTION

Processing large, high-dimensional data efficiently is a fundamental challenge in current classical computing. One example is images, which represent many pixels arranged in rows and columns. By nature, classical operations on images must sequentially explore each pixel. Meanwhile, quantum computing provides a new way to encode and process data [1, 2]. This allows

**doi:** xx.xxxxx/scitechasia.xxxx.xx

high-dimensional data to be represented efficiently and processed in parallel. As such, images are a natural target for quantum computing [3].

To use images in quantum computing, images must be encoded into a quantum state, which allows advantages of quantum computing to be fully utilized. However, a central challenge is on how we prepare a classical image as a quantum state in a form that is both compact and recoverable. Early designs established that such encodings are possible, but each introduced tradeoffs. For example: a design could use more quantum data units to store a richer representation of pixel values [4]. On the other hand, a design could use less quantum data units to efficiently represent pixel values, but this increases the complexity of quantum operations needed to achieve same outcome as the previous design [5]. The central problem in quantum image representation is deciding where resolution scaling belongs: in register count or in circuit depth.

Both schemes adopted a binary address register to identify pixel position. This register grows directly with image resolution: a $2^n \times 2^n$ image requires width $2n$ qubits. Indeed when resolution doubles, two additional qubits are needed. It grows so on based on the formula which may become a constraint or bottleneck in quantum systems.

Researchers have also explored qutrits and other qudit systems, where each quantum unit has more than two basis states [6, 7]. This is a useful direction because one qutrit can carry more address information than one qubit. As a result, qutrit image models can reduce the number of registers needed for the same image size. This gives better compression and supports a more compact image representation.

At the same time, this compression still follows a scaling rule. When the image resolution becomes higher, the address space must also become larger. A fixed qutrit design can slow this growth, but the system still needs more address capacity for larger and clearer images. In this view, the dimension of a qudit becomes an important mathematical resource. Instead of only counting how many registers are used, the dimension $d$ can also describe how much image space the representation can support.

This study proposes QARQI (Quadrant Amplitude Representation of Quantum Images), which is a hybrid qubit-qudit image representation. It encodes pixel position through a quadrant-based coordinate system and pixel intensity through the amplitude of an ancilla qubit. Its register structure is fixed at two polarity qubits, two magnitude qudits, and one ancilla qubit. QARQI is simulated and validated using the Cirq quantum simulator.

## 2. RELATED WORK

In the early foundational era of quantum image representations, algorithms and research in the space were trying to find their footing. Such of the first algorithms used tensor methods that expanded across multiple Hilbert space dimensions. Latorre's Real Ket representation encoded pixel position through quadtree-structured indexing, mapping each hierarchical quadrant directly to basis states

$$|\varphi_{2^n \times 2^n}\rangle = \sum_{\substack{i_1, i_2, \ldots, i_n \\ =1,2,\ldots,4}} c_{i_n,\ldots,i_1} |i_n, \ldots, i_1\rangle , \tag{2.1}$$

where the basis indices carry positional information and $c_{i_n,\ldots,i_1}$ stores pixel intensity as real coefficients. Real Ket was a classical algorithm built on matrix product state (MPS) tensor decomposition. Its truncation step extracts lower-rank factors and produces classical image compression. The same tensor form also supports a position-and-intensity description of the image at the level of the representation itself.

To introduce genuinely quantum image encoding, the Flexible Representation of Quantum Images (FRQI) decoupled position from intensity [5]. For a $2^n \times 2^n$ image, the FRQI state is

$$|I\rangle = \frac{1}{2^n} \sum_{xy=0}^{2^{2n}-1} |C_{xy}\rangle \otimes |xy\rangle\,, \tag{2.2}$$

$$|C_{xy}\rangle = \cos\theta_{xy}\,|0\rangle + \sin\theta_{xy}\,|1\rangle\,, \tag{2.3}$$

where $|xy\rangle$ is the computational basis position and $\theta_{xy}$ is the rotation angle encoding pixel intensity. This architecture enables polynomial-time state preparation with one controlled-$R_y$ gate per pixel. Intensity recovery uses repeated measurements to estimate the original value from the rotation-encoded amplitude. The Novel Enhanced Quantum Representation (NEQR) expanded upon FRQI's color accuracy limitations by placing grayscale values directly in the computational basis [4], giving deterministic exact retrieval [8, 9] with direct arithmetic on stored values. In contrast with Real Ket, where intensity appears as a real coefficient and may remain unnormalized, NEQR places intensity inside the quantum register itself and treats the pixel value as a quantum encoding choice.

NEQR represents each pixel intensity as a binary string of length $q$ stored in the computational basis. NEQR then requires an additional $q$ color qubits on top of the $2n$ address qubits for a $2^n \times 2^n$ image, giving a total register count of $2n + q$. For example: 8-bit grayscale uses $q = 8$, so a $1024 \times 1024$ image ($n = 10$) needs $2n + q = 20 + 8 = 28$ qubits; a full 24-bit RGB pixel stored as a single value requires $q = 24$, so the same resolution would need $20 + 24 = 44$ qubits. For higher-precision or multi-channel formats, such as 16–32 bits per channel or HDR, the qubit count rises in the same way.

On near-term hardware, larger $q$ means more physical qubits. And as gates are conditioned on all $2n$ address qubits, the quantum costs needed to conditionally control matching intensity per its given pixel position grows with image resolution. But as qubit counts and error rates make this scaling difficult to sustain on quantum hardware; naturally, gravitating towards reducing the qubits used and gates used aid in allowing images to scale better on quantum hardware.

To that end, the binary approaches share the same address-register scaling especially at the position register. A $2^n \times 2^n$ resolution image requires exactly $2n$ qubits for the position register alone, which creates severe connectivity bottlenecks on NISQ devices (Near-Intermediate Scale Quantum systems) [10]. Ultimately, constraints would always grow with image resolution.

To address these scaling limits, research shifted from binary towards ternary or even finite $d$-ary quantum systems. The mathematical basis for synthesizing arbitrary unitaries in $d$-level systems was established in early work on qudit quantum gates [6]. Building on this framework, later research adapted FRQI to qutrit logic ($d = 3$) in the Flexible Qutrit Representation of Quantum Images (FQRI)[7], so that the register width scales as

$$\frac{2n}{\log_2 3} \approx 0.631\,(2n),$$

which means the qutrit register uses about $63.1\%$ of the binary address width, or equivalently reduces the register count by about $36.9\%$. And like improvements to the FRQI, recent work extended NEQR to qutrit logic using ternary Toffoli gates to reduce circuit depth in the Qutrit Representation of Quantum Image (QTRQ) model [11, 12].

Beyond image representation, qudit-based quantum algorithms have shown promise in optimization and simulation contexts [13, 14]. Photonic experiments with orbital angular momentum qudits demonstrate natural compatibility with high-dimensional encoding [15]. As seen, most qudit-based QIR schemes still treat the dimension $d$ as a fixed constant in the case of the qutrit models mentioned earlier. However, it notably confines the scaling benefit to one compression ratio. Recent surveys note that this constraint prevents the full use of the scaling potential of high-dimensional quantum systems [16] possibly implying better compression or alternative storage methodologies for using higher-dimensions.

Recent quantum computing work has explored mixed-dimensional systems that combine two-level qubits with $d$-level qudits. Photonic experiments have demonstrated hybrid entanglement between qubit and high-dimensional qudit states for information transfer [17, 18]. Trapped-ion platforms have likewise explored simultaneous control of qubit and qudit degrees of freedom.

At the same time, variational quantum algorithms for image loading have emerged as an alternative. They trade exact encoding for shallower circuits [19]. While effective for classification tasks, these approaches are inherently approximate. For applications that require deterministic pixel reconstruction, amplitude encoding with classical preprocessing remains the preferred approach. Beyond that, other recent quantum image processing work shows renewed interest in hybrid encodings that combine fixed-radix qudits with variable-dimensional systems [18].

QARQI can be read as a finite-dimensional qubit-qudit design within quantum image representation. It grows from the steady progression of earlier models and advances a clear objective. The model uses $d$ in ket space as a free finite parameter and uses mixed-dimensional qubit-qudit structure so scaling moves into higher-dimensional but finite Hilbert spaces. It thereby reframes the cost which would be inherited from prior image encodings regarding growth related to image resolution that was previously carried by register width is now redirected toward Hilbert-space dimensionality.

## 3. METHODS

### 3.1 Registers and indexing

Standard quantum images use a binary raster scan, with pixels indexed sequentially from 0 to $N-1$. QARQI instead adopts a quadrant-based coordinate system. We can consider the system to be similar to a Cartesian grid where the center of the image is the origin.

However, the design was actually inspired by classical quadrature amplitude modulation (QAM) constellation diagrams which is also a bit similar to Cartesian grids. In QAM each symbol is described by a polarity (the sign about the in-phase and quadrature axes) and a magnitude (the distance from the origin). We adopt this terminology for the QARQI model's components: the binary qubits are called "polarity qubits" because they select the sign (left/right, bottom/top) of the coordinate, and the high-dimensional qudits are called "magnitude qudits" because they encode the offset from the axis within the chosen quadrant.

The registers are divided into two

functional groups:

- **Polarity Qubits** $(i, j)$**:** Two binary qubits determine the quadrant. $i$ selects the Left/Right half, and $j$ selects the Bottom/Top half.
- **Magnitude Qudits** $(k, \ell)$**:** Two high-dimensional qudits determine the absolute distance from the center. $k$ is the horizontal distance, and $\ell$ is the vertical distance.

Fig. 1 shows this structure. The polarity qubits select the quadrant and the magnitude qudits select the offset within that quadrant. The register addresses $M = 4d^2$ positions. The mapping from the register state to the linear pixel index $m$ is piecewise:

$$x = \begin{cases} d - 1 - k, & i = 0, \\ d + k, & i = 1, \end{cases}$$

$$y = \begin{cases} d + \ell, & j = 0, \\ d - 1 - \ell, & j = 1, \end{cases}$$

$$m(i, j, k, \ell) = y\,(2d) + x. \qquad (3.1)$$

The zero offsets $k = 0$ and $\ell = 0$ address the cells adjacent to the origin. The four quadrant address ranges are disjoint for every $d \geq 1$: $x \in \{0, \dots, d-1\}$ when $i = 0$ and $x \in \{d, \dots, 2d-1\}$ when $i = 1$, while $y \in \{0, \dots, d-1\}$ when $j = 1$ and $y \in \{d, \dots, 2d-1\}$ when $j = 0$.

**Proposition 1.** *For every finite dimension $d \geq 1$, the mapping in Eq.* (3.1) *assigns each register state $(i, j, k, \ell)$ to exactly one pixel in the $2d \times 2d$ image grid, and every pixel in that grid is reached.*

*Proof.* For the horizontal coordinate, $i = 0$ gives $x = d - 1 - k$, so $x \in \{0, \dots, d-1\}$. The value of $k$ is recovered as $k = d - 1 - x$. For $i = 1$, $x = d + k$, so $x \in \{d, \dots, 2d-1\}$. The value of $k$ is recovered as $k = x - d$. These two ranges are disjoint and together cover all possible horizontal coordinates. The same argument applies to $y$. When $j = 1$, $y = d - 1 - \ell$ covers $\{0, \dots, d-1\}$, and when $j = 0$, $y = d + \ell$ covers $\{d, \dots, 2d-1\}$. Thus each register tuple gives one pair $(x, y)$, and each pair $(x, y)$ has one register tuple. Since $m = y(2d) + x$ is the usual row-major index on this grid, each register tuple also gives one pixel index $m$. □

**Example 1.** (Coordinate mapping with a $4 \times 4$ image.) Let $d = 2$ and consider the labeled grayscale image in which each cell shows its pixel value and row-major index:

| | | | |
|---|---|---|---|
| $12_0$ | $40_1$ | $75_2$ | $90_3$ |
| $18_4$ | $55_5$ | $96_6$ | $120_7$ |
| $22_8$ | $64_9$ | $128_{10}$ | $180_{11}$ |
| $30_{12}$ | $80_{13}$ | $160_{14}$ | $220_{15}$ |

Fig. 1 shows the same $d = 2$ coordinate layout. For the highlighted address state $|i, j, k, \ell\rangle = |1, 0, 1, 0\rangle$, the polarity qubits select the $+x, -y$ quadrant. The magnitude qudits fix the offset inside that quadrant. Substituting into Eq. (3.1):

$$x = d + k = 2 + 1 = 3,$$
$$y = d + \ell = 2 + 0 = 2, \qquad (3.2)$$

yielding

$$m = y\,(2d) + x = 2 \cdot 4 + 3 = 11. \qquad (3.3)$$

The selected pixel is therefore the entry at $(x, y) = (3, 2)$, whose value in the example image is $I_{11} = 180$.

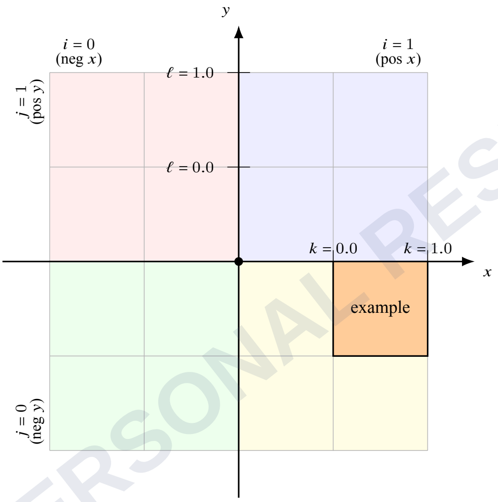



**Fig. 1.** Example with $d = 2$. The highlighted cell is $(1, 0, 1, 0)$ in the $+x, -y$ quadrant, and its row-major index is $m = 11$.

### 3.2 State definition

The quantum state is built in two stages: initialization and encoding. First, all registers are set to the zero state:

$$|\Psi_0\rangle = |0\rangle^{\otimes(n+1)} \tag{3.4}$$

$$= |0, 0, 0, 0\rangle_{\text{index}} \otimes |0\rangle_c \,. \tag{3.5}$$

Next, Hadamard gates and their qudit equivalents create a uniform superposition over the index register. The state $|\Psi_1\rangle$ contains every pixel location simultaneously, with the ancilla still in $|0\rangle$. Let $\mathrm{a} = (i, j, k, \ell)$ denote one address tuple. Here $H_d$ denotes the $d$-dimensional discrete Fourier transform, with matrix entries $(H_d)_{ab} = d^{-1/2} e^{2\pi i ab/d}$ for $a, b \in \{0, \ldots, d-1\}$.

$$|\Psi_1\rangle = \frac{1}{2d} \sum_{\mathrm{a}} |\mathrm{a}\rangle \otimes |0\rangle_c \,. \tag{3.6}$$

### 3.3 Classical normalization of pixel data

Classical pixels are integers $I_m \in [0, I_{\max}]$, while quantum amplitudes are normalized to unit norm. Each pixel is mapped to the unit interval $[0, 1]$ and then to a rotation angle $\theta_m$ such that the ancilla measurement probability matches the normalized brightness:

$$\nu_m = \frac{I_m}{I_{\max}} \in [0, 1], \tag{3.7}$$

$$\sin^2 \theta_m = \nu_m. \tag{3.8}$$

Solving for $\theta_m$ yields the rotation parameter:

$$\theta_m = \arcsin\left(\sqrt{\nu_m}\right). \tag{3.9}$$

In contrast to QAM encoding, where zero magnitude represents a special constellation point, pixel intensity is encoded separately as an ancilla amplitude via rotations. The ancilla does not use zero magnitude as a pixel code. Instead, the ancilla's reference basis ($|0\rangle$) functions as the rotation reference, so that intensity appears through rotations of the ancilla ($\cos\theta\,|0\rangle + \sin\theta\,|1\rangle$).

This geometric transformation allows the physical rotation of the qubit state vector to represent the grayscale value. When an $R_y(2\theta_m)$ gate is applied to an ancilla initialized at $|0\rangle$, it produces:

$$R_y(2\theta_m)\,|0\rangle = \cos(\theta_m)\,|0\rangle + \sin(\theta_m)\,|1\rangle \,. \tag{3.10}$$

The coefficient of $|1\rangle$ becomes $\sin(\theta_m)$, and its squared magnitude recovers the original intensity $\nu_m$ during measurement.

### 3.4 Encoding unitary

The final step is to imprint the image data. The circuit applies a global unitary $R$ formed by controlled rotations, with one projector $\Pi_{ijk\ell}$ for each address $(i, j, k, \ell)$. For each address, the projector is $\Pi_{ijk\ell} = |i, j, k, \ell\rangle \langle i, j, k, \ell| \otimes I_c$ on the full register space.

$$R = \prod_{m=0}^{M-1} \mathrm{C}_{\Pi_m}\big(R_y(2\theta_m)\big). \tag{3.11}$$

Because each projector $\Pi_m$ selects a unique basis state, different address projectors are orthogonal:

$$\Pi_m \Pi_{m'} = 0 \quad \text{for } m \neq m'. \tag{3.12}$$

The controlled rotations therefore commute with one another. Algorithm 1 gives the full procedure.

---

**Algorithm 1** QARQI Encoding Protocol

---

**Require:** Image $I$ $(N \times N)$, Ancilla qubit $c$

**Classical Preprocessing:**
**for** each pixel $m = 0$ to $M - 1$ **do**
  $v_m \leftarrow I_m / I_{\max}$
  $\theta_m \leftarrow \arcsin(\sqrt{v_m})$
**end for**
**Quantum Initialization:**
Initialize $|\Psi_0\rangle \leftarrow |0\rangle_i |0\rangle_j |0\rangle_k |0\rangle_\ell \otimes |0\rangle_c$
Apply $H$ to qubits $i, j$
Apply $H_d$ to qudits $k, \ell$
Result: Uniform Superposition $|\Psi_1\rangle$
**Amplitude Encoding:**
**for** each pixel $m$ **do**
  Apply $C_{\Pi_m}(R_y(2\theta_m))$ to ancilla $c$
**end for**
**return** Final State $|\Psi_{\text{QARQI}}\rangle$

---

Let

$$|C_m\rangle = \cos\theta_m |0\rangle_c + \sin\theta_m |1\rangle_c . \tag{3.13}$$

For compact notation, let $m_{\mathrm{a}} = m(i, j, k, \ell)$. The result is the final QARQI state, where every position index is entangled with an ancilla amplitude representing its specific intensity:

$$|\Psi_{\text{QARQI}}\rangle = \frac{1}{2d} \sum_{\mathrm{a}} |\mathrm{a}\rangle \otimes |C_{m_{\mathrm{a}}}\rangle . \tag{3.14}$$

Fig. 2 gives the full QARQI state for the $4 \times 4$ example. The expansion lists all 16 address states produced when $d = 2$, and each address is paired with the carrier state $|C_m\rangle$ for the corresponding row-major pixel. This makes the role of the full state explicit: the address register identifies the pixel location, while the ancilla amplitude stores that pixel's normalized intensity through $\theta_m$.

**Proposition 2.** *The QARQI state in Eq.* (3.14) *is normalized.*

*Proof.* From Eq. (3.14), with $\mathrm{a} = (i, j, k, \ell)$, $\mathrm{a}' = (i', j', k', \ell')$, and $m_{\mathrm{a}} = m(i, j, k, \ell)$ denoting the pixel index assigned by the coordinate map. We write $|C_{\mathrm{a}}\rangle \equiv |C_{m_{\mathrm{a}}}\rangle$ and that $|\Psi\rangle \equiv |\Psi_{\text{QARQI}}\rangle$. Then

$$\langle\Psi|\Psi\rangle = \frac{1}{4d^2} \sum_{\mathrm{a}} \sum_{\mathrm{a}'} \langle \mathrm{a}|\mathrm{a}'\rangle \langle C_{\mathrm{a}}|C_{\mathrm{a}'}\rangle . \tag{3.15}$$

Here $\sum_{\mathrm{a}}$ abbreviates $\sum_{i,j=0}^{1} \sum_{k,\ell=0}^{d-1}$. The address basis is orthonormal, so

$$\langle i, j, k, \ell | i', j', k', \ell' \rangle = \delta_{ii'} \delta_{jj'} \delta_{kk'} \delta_{\ell\ell'} . \tag{3.16}$$

Hence all cross-address terms vanish, and the double sum collapses to

$$\langle\Psi|\Psi\rangle = \frac{1}{4d^2} \sum_{i,j=0}^{1} \sum_{k,\ell=0}^{d-1} \langle C_{\mathrm{a}}|C_{\mathrm{a}}\rangle . \tag{3.17}$$

For each pixel state,

$$\langle C_m|C_m\rangle = \big(\cos\theta_m \langle 0|_c + \sin\theta_m \langle 1|_c\big) \times \big(\cos\theta_m |0\rangle_c + \sin\theta_m |1\rangle_c\big) \tag{3.18}$$

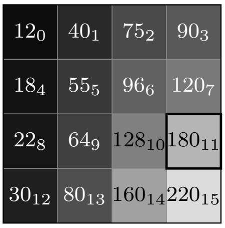
$$= \cos^2\theta_m \langle 0|0\rangle + \sin^2\theta_m \langle 1|1\rangle \tag{3.19}$$

$$= \cos^2\theta_m + \sin^2\theta_m = 1. \tag{3.20}$$

Therefore

$$\langle\Psi|\Psi\rangle = \frac{1}{4d^2} \sum_{i=0}^{1} \sum_{j=0}^{1} \sum_{k=0}^{d-1} \sum_{\ell=0}^{d-1} 1 \tag{3.21}$$

$$= \frac{4d^2}{4d^2} \tag{3.22}$$




$$
\begin{aligned}
|\Psi_{4\times4}\rangle = \frac{1}{4}\big(&|0111\rangle|C_0\rangle + |0101\rangle|C_1\rangle + |1101\rangle|C_2\rangle + |1111\rangle|C_3\rangle \\
&+ |0110\rangle|C_4\rangle + |0100\rangle|C_5\rangle + |1100\rangle|C_6\rangle + |1110\rangle|C_7\rangle \\
&+ |0010\rangle|C_8\rangle + |0000\rangle|C_9\rangle + |1000\rangle|C_{10}\rangle + |1010\rangle|C_{11}\rangle \\
&+ |0011\rangle|C_{12}\rangle + |0001\rangle|C_{13}\rangle + |1001\rangle|C_{14}\rangle + |1011\rangle|C_{15}\rangle\big).
\end{aligned}
$$

**Fig. 2.** Expanded QARQI state for $d = 2$. The left panel shows pixel values with row-major indices. The right panel shows the matching address-state expansion, where $|C_m\rangle = \cos\theta_m |0\rangle_c + \sin\theta_m |1\rangle_c$.

$$= 1. \tag{3.23}$$

□

Conditioned on a measured address $(i, j, k, \ell)$, the ancilla outcome recovers the normalized intensity in expectation:

$$
\begin{aligned}
\Pr(c = 1 \mid i, j, k, \ell) &= \sin^2\theta_{m(i,j,k,\ell)} \\
&= \nu_{m(i,j,k,\ell)}.
\end{aligned} \tag{3.24}
$$

### 3.5 Simulation Implementation

The simulator follows the same register model as the formal construction: two polarity qubits, two $d$-level magnitude qudits, and one ancilla qubit. In the implementation these are ordered as $(b_0, b_1, x_0, x_1, h)$, where $h$ is the ancilla; this corresponds to the paper notation $(i, j, k, \ell, c)$.

For the validation runs, $N = 256$ and $d = 128$. Initialization prepares the uniform address superposition using Hadamard gates on the polarity qubits and dimension-$d$ Fourier transforms on the magnitude qudits. Intensity upload then applies an address-indexed ancilla rotation with angle $2\arcsin\sqrt{\nu_m}$, so the measured ancilla probability follows Eq. (3.24).

Readout measures all registers together, giving tuples $(i, j, k, \ell, h)$ with $h \in \{0, 1\}$. Outcomes are grouped by address $(i, j, k, \ell)$, and the local hit ratio estimates

$$\hat{p} = \frac{n_{\text{hit}}}{n_{\text{hit}} + n_{\text{miss}}}, \tag{3.25}$$

which is mapped back to grayscale by $I \approx 255\,\hat{p}$.

The sampling budget is reported in total shots. For an $N \times N$ image with $d = N/2$, there are $N^2$ addresses and two ancilla outcomes per address, so

$$\text{BIN\_COUNT} = 2N^2. \tag{3.26}$$

For $N = 256$, this gives 131,072 bins. The simulator also records a normalized shots-per-bin value, denoted spb, and the total shots for one run are

$$K_{\text{total}} = \text{spb} \times \text{BIN\_COUNT}. \tag{3.27}$$

In this study, spb is swept from 2 to 60, which corresponds to 262,144 to 7,864,320 total shots per reconstruction. The intermediate reported levels of spb = 19 and spb = 39 correspond to 2,490,368 and 5,111,808 total shots, respectively.

### 3.6 Operations: Flips and Rotations

The simulation also applies spatial transformations before measurement. The image is not classically transformed before upload, and the reconstructed image is not reordered after measurement. Each operation is applied to the encoded quantum state as a permutation of the address basis, while the ancilla readout rule remains unchanged.

Let $\mathrm{a} = (i, j, k, \ell)$ and let $T_g$ denote the address transformation for an operation $g$. The corresponding unitary $U_g$ acts on basis states by

$$U_g |\mathrm{a}\rangle |c\rangle = |T_g(\mathrm{a})\rangle |c\rangle. \tag{3.28}$$

For the $2d \times 2d$ row-major image grid used in Eq. (3.1), the spatial transformations are

$$\begin{aligned} H(x,y) &= (2d-1-x,\ y), \\ V(x,y) &= (x,\ 2d-1-y), \\ R_{180}(x,y) &= (2d-1-x,\ 2d-1-y), \\ R_{90}(x,y) &= (2d-1-y,\ x), \\ R_{270}(x,y) &= (y,\ 2d-1-x), \end{aligned} \tag{3.29}$$

where $R_{90}$ and $R_{270}$ denote clockwise rotations by 90° and 270°, respectively. Substituting these image-coordinate maps into the QARQI coordinate rule gives the address-register maps

$$\begin{aligned} T_H(i,j,k,\ell) &= (1-i,\ j,\ k,\ \ell), \\ T_V(i,j,k,\ell) &= (i,\ 1-j,\ k,\ \ell), \\ T_{180}(i,j,k,\ell) &= (1-i,\ 1-j,\ k,\ \ell), \\ T_{90}(i,j,k,\ell) &= (j,\ 1-i,\ \ell,\ k), \\ T_{270}(i,j,k,\ell) &= (1-j,\ i,\ \ell,\ k). \end{aligned} \tag{3.30}$$

Thus a horizontal flip applies an $X$ gate to the column-polarity qubit, a vertical flip applies an $X$ gate to the row-polarity qubit, and a 180° rotation applies both polarity flips:

$$\begin{aligned} U_H &= X_i \otimes I_{j,k,\ell,c}, \\ U_V &= I_i \otimes X_j \otimes I_{k,\ell,c}, \\ U_{180} &= X_i \otimes X_j \otimes I_{k,\ell,c}. \end{aligned} \tag{3.31}$$

The 90° and 270° rotations are implemented as finite-basis permutation unitaries over the address registers:

$$\begin{aligned} U_{90}\,|i,j,k,\ell,c\rangle &= |j,1-i,\ell,k,c\rangle, \\ U_{270}\,|i,j,k,\ell,c\rangle &= |1-j,i,\ell,k,c\rangle. \end{aligned} \tag{3.32}$$

The lake image is used as the reference image for the spatial-operation validation. The recorded runs apply the operation before measurement, with no classical preprocessing of the uploaded image and no postprocessing reordering of the reconstructed output. For the horizontal-flip case, the high-shot correlation with the classical target reached 0.988161 at 7,864,320 total shots. The same run reported that the ancilla marginal distribution was unchanged after the flip operation, while the address probabilities were permuted.

## 4. RESULTS

The shot-budget sweep was executed from 262,144 to 7,864,320 total shots per run on nine 256-by-256 test images: cameraman, house, jetplane, lake, lena_gray_512, livingroom, mandril_gray, pirate, and walkbridge. Reconstruction quality was evaluated using MSE and PSNR. Fig. 3 shows the best reconstruction from the sweep for each image.

At the lowest shot level, reconstructions are coarse, as expected from shot-based sampling. As total shots increase, PSNR rises and MSE falls for every image, with no reversals across the sweep. The mean PSNR is 11.4818 dB at 262,144 shots, 22.6000 dB at 2,490,368 shots, and 25.7818 dB at 5,111,808 shots. The per-image results show that the trend is similar across the nine test images.

The highest shot budget produced the strongest reconstructions for every image in the set. At 7,864,320 total shots, the final PSNR values ranged from 27.1686 dB for mandril_gray to 28.2223 dB for jetplane, while the corresponding MSE values ranged from 97.9158 to 124.8015. The remaining images followed the same pattern, with cameraman at 27.9502 dB, house at 27.7629 dB, lake at 28.0522 dB, lena_gray_512 at 27.4440 dB, livingroom at 27.3079 dB, pirate at 27.4744 dB, and walkbridge at 27.6319 dB.


**Fig. 3.** Best validation reconstructions for the nine 256-by-256 images. Each image reached its strongest reconstruction at 7,864,320 total shots.

**Table 1.** Best validation metrics at 7,864,320 total shots.

| Image | MSE | PSNR (dB) |
|---|---|---|
| cameraman | 104.2472 | 27.9502 |
| house | 108.8417 | 27.7629 |
| jetplane | 97.9158 | 28.2223 |
| lake | 101.8265 | 28.0522 |
| lena_gray_512 | 117.1332 | 27.4440 |
| livingroom | 120.8632 | 27.3079 |
| mandril_gray | 124.8015 | 27.1686 |
| pirate | 116.3159 | 27.4744 |
| walkbridge | 112.1747 | 27.6319 |

The same sampling setup was used to evaluate spatial operations. The lake image was transformed by horizontal flip, vertical flip, and right-angle rotations before measurement. Fig. 4 shows the reconstructed output at 7,864,320 total shots together with the corresponding shot-budget trend for each operation.

The transformed circuits follow the same accuracy-cost pattern as the base reconstruction. PSNR increases and MSE decreases as total shots grow for each operation. In the spatial-operation validation, the maximum PSNR across the operation sweeps ranged from 28.0490 dB for the 180° rotation to 28.0956 dB for the 90° rotation.

## 5. DISCUSSION

Existing quantum image representations often inherit assumptions from qubit-based systems. As it currently stands, information is fundamentally expressed through binary basis states. During the formulation of QARQI, it became apparent that the quantum state itself need not require a strictly binary interpretation. While indeed the computational basis of a qubit remains limited to states corresponding to $|0\rangle$ and $|1\rangle$, the overall normalized quantum state is capable of representing much more information through higher-dimensional systems as well.

This realization became particularly relevant when examining such a finite-dimensional qudit system. In such systems, the address structure may still rely on a small number of control registers, while the encoded information occupies the aforementioned higher-dimensional state spaces. While it could be said that the growth of the representation is not necessarily driven by the number of registers alone, it could also be said that a significant portion of the information capacity is transferred into the dimensionality of the quantum states themselves that this paper thereby proposes. However, this raises several open questions regarding how such information would be physically encoded, manipulated, and retrieved in practical implementations. Although the mathematical representation remains well-defined, the operational aspects of accessing and controlling higher-dimensional information require further investigation.

Another area that remains open concerns gate realization. The current implementation primarily relies on generalized Hadamard operations and multi-controlled rotation gates. Such gates were naturally provided in the simulation software used and not much discussion focused on that particular aspect. Simply put, most if not all the mixed dimensional quantum simulation software already provide such gate realizations and generalizations natively and abstractly. While these operators are sufficient for demonstrating the encoding procedure at face value, the underlying decomposition of such gates in finite-dimensional systems was not explored in detail. The cost of these operations may depend not only on the number of control registers. It could also be possible that the cost depend on the dimensionality of the target system or the dimensionality required by higher resolution images. If transitions between different levels require unique operations, the implementation cost could scale differently than initially expected. At present, this remains an unresolved question and would require a dedicated analysis of gate complexity and physical realization.

During simulation, the implementation also appeared to exhibit favorable memory behavior compared to several quantum image representations explored during the early stages of learning quantum image processing. This observation was primarily qualitative and should not be interpreted as a formal performance claim. The reduced number of required registers naturally led to smaller simulated state representations, creating the impression of improved memory efficiency. However, a rigorous evaluation would require a formal analysis of register complexity, state-

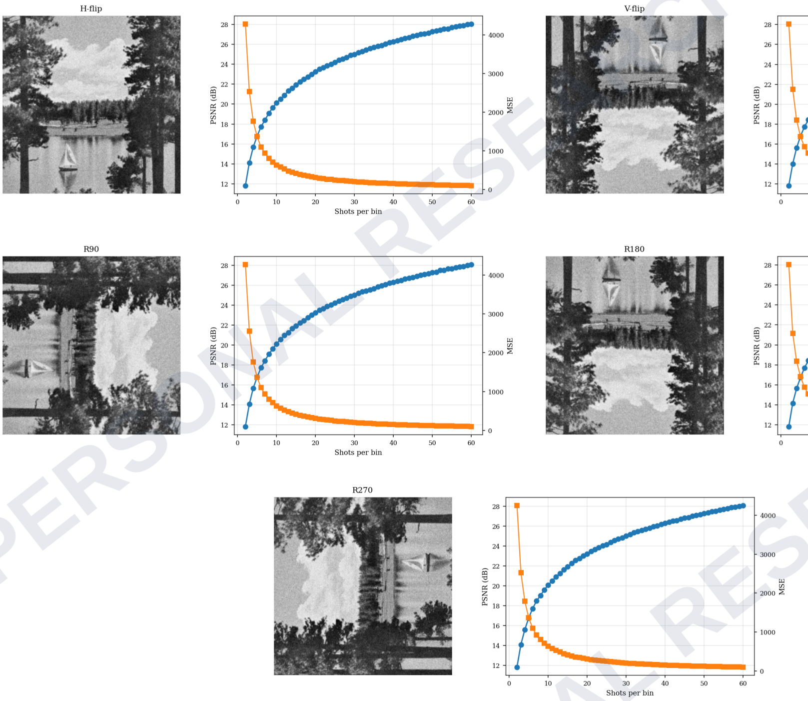



**Fig. 4.** Spatial-operation composites for the lake image. Each composite pairs the operation reconstruction at 7,864,320 total shots with the corresponding shot-budget sweep.

space growth, gate complexity, and execution cost. Such an analysis was outside the scope of the present work.

The principal contribution of QARQI produces a valid normalized quantum state and that the encoding process itself is genuinely quantum in nature. The framework offers an alternative direction for quantum image representation, whereby it explores dimensionality as a resource apart from relying solely on increasing register. As quantum hardware continues to develop beyond conventional qubit system, particularly in platforms capable of supporting higher-dimensional quantum states, such as orbital angular momentum systems, the exploration of finite-dimensional image representations may become increasingly relevant. The present work serves as an initial step toward investigating how these higher-dimensional quantum systems can be utilized for image encoding and processing.

## References

[1] Nielsen MA, Chuang IL. Quantum Computation and Quantum Information. 10th ed. Cambridge University Press; 2010.

[2] Babbush R, Low GH, McClean JR, McClain J, Chan GKL, Aspuru-Guzik A. Encoding electronic spectra in quantum circuits with linear T complexity. Physical Review X. 2018;8(4):041015.

[3] Latorre JI. Image compression and entanglement. arXiv preprint arXiv:quant-ph/0510031. 2005.

[4] Zhang Y, Lu K, Gao Y, Wang M. NEQR: a novel enhanced quantum representation of digital images. Quantum Information Processing. 2013;12(8):2833-60.

[5] Le PQ, Dong F, Hirota K. A flexible representation of quantum images for polynomial preparation, image compression, and processing operations. Quantum Information Processing. 2011;10(1):63-84.

[6] Muthukrishnan A, Stroud Jr CR. Multivalued logic gates for quantum computation. Physical Review A. 2000;62(5):052309.

[7] Song X, He G, et al. Flexible representation of quantum images with qutrits. In: International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP); 2017. p. 125-8.

[8] Zhang Y, Lu K. INEQR: An improved novel enhanced quantum representation of digital images. Quantum Information Processing. 2015;14(5):1673-87.

[9] Sun B, Iliyasu AM, Yan F, Hirota K. Watermarking quantum images based on NEQR. Quantum Information Processing. 2013;12(11):3411-30.

[10] Preskill J. Quantum Computing in the NISQ era and beyond. Quantum. 2018;2:79.

[11] Dong H, Lu D, Li C. QTRQ: A novel qutrit representation of quantum images. Quantum Information Processing. 2022;21(3):108.

[12] Monfared AT, Ciriani V, Haghparast M. Qutrit representation of quantum images: new quantum ternary circuit design. Quantum Information Processing. 2024.

[13] Wang Y, Hu Z, Sanders BC, Kais S. Qudits and high-dimensional quantum computing. Frontiers in Physics. 2020;8:589504.

[14] Erhard M, Krenn M, Zeilinger A. Advances in high-dimensional quantum entanglement. Nature Reviews Physics. 2020;2(7):365-81.

[15] Luo YH, et al. Quantum teleportation in high dimensions. Physical Review Letters. 2019;123(7):070505.

[16] Khandelwal A, Chandra MG. Quantum Image Representation Methods Using Qutrits. arXiv preprint arXiv:220709096. 2022. Available from: `https://arxiv.org/abs/2207.09096`.

[17] Liu J, Niu MY, Zhang ea. Hybrid quantum-classical architectures for image processing. IEEE Transactions on Quantum Engineering. 2023;4:1-12.

[18] Yan F, Iliyasu AM, Venegas-Andraca SE. A survey of quantum image representations and their applications in the NISQ era. ACM Computing Surveys. 2024;56(3):1-35.

[19] Cerezo M, Arrasmith A, Babbush R, et al. Variational quantum algorithms. Nature Reviews Physics. 2021;3(9):625-44.