Article

A Hybrid Quantum Image-Matching Algorithm

Guoqiang Shu, Zheng Shan, Shiqin Di, Xiaodong Ding and Congcong Feng

Article

A Hybrid Quantum Image-Matching Algorithm

Guoqiang Shu [ID], Zheng Shan *, Shiqin Di, Xiaodong Ding and Congcong Feng

State Key Laboratory of Mathematical Engineering and Advanced Computing, Zhengzhou 450000, China
* Correspondence: zzzhengming@163.com

Abstract: Image matching is an important research topic in computer vision and image processing. However, existing quantum algorithms mainly focus on accurate matching between template pixels, and are not robust to changes in image location and scale. In addition, the similarity calculation of the matching process is a fundamentally important issue. Therefore, this paper proposes a hybrid quantum algorithm, which uses the robustness of SIFT (scale-invariant feature transform) to extract image features, and combines the advantages of quantum exponential storage and parallel computing to represent data and calculate feature similarity. Finally, the quantum amplitude estimation is used to extract the measurement results and realize the quadratic acceleration of calculation. The experimental results show that the matching effect of this algorithm is better than the existing classical architecture. Our hybrid algorithm broadens the application scope and field of quantum computing in image processing.

Keywords: image matching; amplitude encoding; quantum computing; amplitude estimation; quadratic acceleration

1. Introduction

Image matching is one of the key technologies of digital photogrammetry in machine learning fields, such as information processing, pattern recognition, computer vision, etc. [1–4]. Image matching is known as the comparison of two images, and it can be roughly divided into three categories: the first category is based on image grayscale [5], the second category is related to image features [6,7], and the third category is in view of the understanding and interpretation of images. However, image matching faces some obstacles, such as object rotation, scaling and translation, illumination and noise, etc. Dong et al. proposed the DSP-SIFT (domain-size pooling scale-invariant feature transform) method that improves descriptor matching performance by pooling feature block size and gradient direction [8,9]. The local features and spatial layout are integrated and encoded into the edges of a graph whose nodes represent potential correspondences for discovering all common visual patterns. All strongly connected subgraphs correspond to large local maxima of a quadratic function on simplex using a replicator equation and through a systematic way of initialization. This was demonstrated to be robust to outliers, and able to discover all common visual patterns [10]. The method ICF (identifying point correspondences by correspondence function) was proposed for rejecting mismatches from given putative point correspondences. This introduces the correspondence function for capturing the relationships between corresponding points and an algorithm IECF (iteratively estimate correspondence function) based on the diagnostic technique and SVM. The method rejects the mismatches by checking whether they are consistent with the estimated correspondence functions, and demonstrated excellent performance [11]. The locality preserving matching (LPM) was proposed to seek reliable correspondences between two feature sets and the principle of which is to maintain the local neighborhood structures of those potential true matches. It achieves better or favorably competitive performance in accuracy while intensively cutting the time cost by more than two orders of magnitude [12]. The method (GMS) grid-based motion statistics was proposed to separate

true correspondences from false ones at high speed, which incorporates the smoothness constraint into a statistic framework for separation and uses a grid-based implementation for fast calculation. GMS is robust to various challenging image changes [13]. There may leave several open questions about which method would be a suitable choice for specific applications with respect to different scenarios and task requirements and how to design better image-matching methods with superior performance inaccuracy, robustness and efficiency. A work analyzed them from the perspective of theory, experiment, etc. [14]. A method motion-consistency driven matching (MCDM) for mismatch removal was proposed, which formulates the matching problem into a probabilistic graphical model to infer the state of each node. The final inference is cast into an integer quadratic programming problem, and the solution is obtained by using an efficient optimization technique based on the FrankWolfe algorithm. The strong generalization ability as well as high accuracy, which outperforms state-of-the-art methods, was demonstrated in the work [15]. The multi-scale attention-based network (called MSA-Net) feature-matching method was proposed to solve the limited effectiveness and robustness when applied to different scenarios. By introducing a multi-scale attention block to enhance the robustness to outliers, it is able to effectively infer the probability of correspondences being inliers with fewer parameters, achieving a 11.7% improvement at error threshold 5 compared to the state-of-the-art method on the YFCC100M dataset [16].

Furthermore, the similarity calculation methods of features are diverse. For example, the correlation function, covariance function, structural similarity index method, etc. [17,18]. Moreover, high accuracy is necessary for image matching, and thus the amount of calculation is huge, so it is difficult to meet the requirements of real-time operation, and the matching rate will increase in areas with poor information frequency. Fortunately, advances in quantum information science and machine learning have led to the natural emergence of quantum machine learning, a field that bridges the two, aiming to revolutionize information technology [19–21]. The core of its interest lies in either taking advantage of quantum mechanics to achieve machine learning that surpasses the classical pendant in terms of computational complexity or to entirely be able to apply such techniques on quantum data. An important application of machine learning is pattern recognition in big data analysis [22].

There have been several references that discuss the quantum image matching. A pattern-matching method based on the Grover algorithm is proposed, and approximate amplitude coding and inversion test operation are used. Experiments show that the algorithm is effective; however, the dimension requirements for querying images are low [23]. Daniel Curtis et al. considered a problem of locating a template as a sub image of a larger image [24]. Algorithms that search for a pattern within a larger dataset was proposed; the complexity of the theory is $O(\sqrt{n})$ [25]. They used Fourier transform to calculate the correlation, and the computational complexity was found to be less than the classical correlation algorithm. Yang et al. gave a quantum gray-scale image matching scheme [26]. By mapping the template image with the quantum reference image directly, the scheme is given to calculate the difference of all pixels in two images. Furthermore, Luo et al. proposed a quantum fuzzy matching algorithm based on the pixel-mapping difference, which shows that quantum is exponential parallel acceleration [27]. Jiang et al. investigated the quantum image matching in depth [28,29]. They proposed a quantum matching scheme based on Grover algorithm to find out a small image in a big image. Unfortunately, the paper only matched one pixel and without more than one pixel, the algorithm will cause an error.Therefore, an improved version was presented which takes full advantage of the whole matched area to locate a small image in a big image soon after that [30]. Zhou et al. proposed a quantum image similarity analysis method based on the combination of quantum image pixel representation and amplitude amplification, which demonstrates the strong parallel ability and secondary acceleration advantages of quantum compared with the classical violent element search method [31]. In addition, Ashley proposed a quantum

pattern matching algorithm based on pixel, which is super-polynomially faster than the classical algorithm [32].

Previous research work is mainly based on the image pixel matching method. This requires high consistency between template images and reference images, and greatly limits the expansion of matching applications. Therefore, this paper proposes a hybrid quantum image-matching algorithm based on feature similarity. The classical SIFT algorithm is used to extract the feature attributes of the image and effectively solve the problem of matching error caused by image deformation and other factors. At the same time, we combine the exponential storage and parallel computing capabilities of quantum to calculate the similarity of features, which improves the efficiency of computing and reduces the consumption of resources greatly. Finally, the quantum amplitude estimation algorithm is applied to bring quadratic acceleration to the measurement process of the results.

The article is organized as follows. Section 2 introduces the hybrid quantum algorithm, including the classical feature extraction method, quantum image representation, quantum similarity calculation and amplitude estimation algorithm. In Section 3, two datasets are selected to test the effectiveness of the proposed algorithm, and we analyze the experimental results compared with the classical algorithm. Finally, the conclusions are drawn in Section 4.

2. Materials and Methods

2.1. Feature Extraction and Matching

The SIFT detector extracts a number of meaningful descriptive features [33–35]. These features are invariant to scaling, rotation, and illumination. Such points are generally presented in high-contrast areas, possibly on object edges. One significant quality of these features is that the relative positions between them should not change from one image to another.

The main stage is scale–space extrema extraction are as follows. The first stage is scale–space extrema extraction. In this stage, the interest points, which are scale and rotation invariant are searched, and are obtained with the difference of Gaussian (DoG) function. Then, the key point localization and filtering are performed. In this stage, the location and scale for resultant interest points are found, and the key points are selected, which are robust to image distortion. The next stage is the orientation assignment in which one or more orientation is assigned to each key point location based on the local image-gradient directions. Finally, the feature description is performed. The local image gradients are measured at the selected scale in the neighborhood of a key point, and the 128D feature descriptor is obtained (Figure 1).
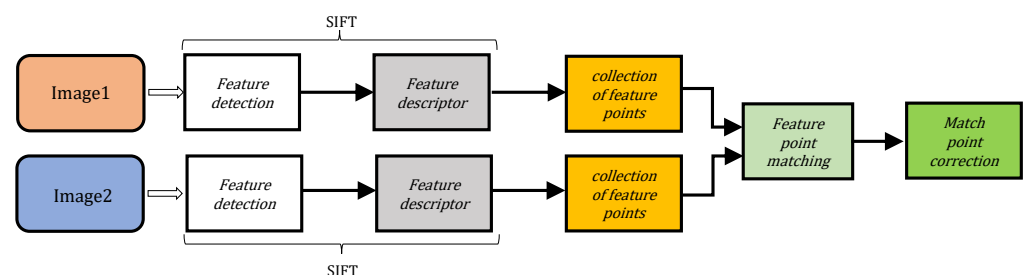

Figure 1. Diagram of SIFT feature extraction.

Quantum is characterized by parallelism, as is classical computing. For the same number of bits n, the data representation and processing capability of the quantum $O(2^n)$ is beyond the classical computing capability $O(n)$. Hence, we propose to analyze the similarity of features with quantum, as shown in Figure 2. Here is a diagram of the hybrid algorithm, where each colored small grid represents a feature point, and the color category stands for the category of the feature point.
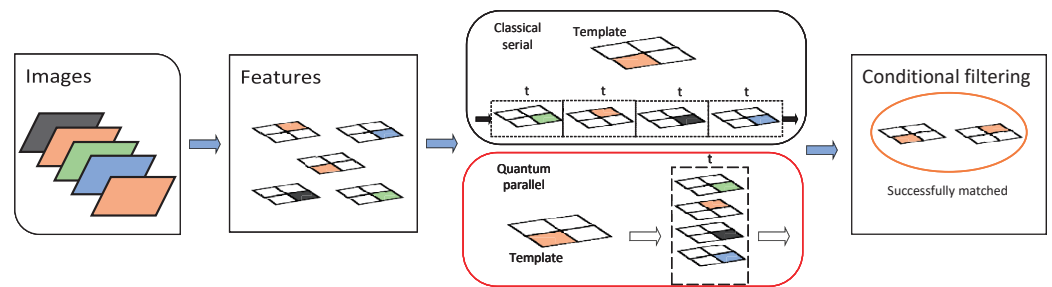
Figure 2. Flow chart of matching algorithm. The part contained in the red box is quantum substitution for classical calculations in black; it is also the structure of the hybrid matching algorithm.

After the descriptors are obtained, the similarity between them needs to be calculated. There are many classical methods of similarity calculation for feature points matching [36,37]. For example, the correlation function is a method to search for a window Y whose correlation function value with window X reaches a maximum value within a certain search range D, and it can be equivalent to the inner product operator and cosine distance:

$$R(X,Y) = \sum_{i=1}^{m}\sum_{j=1}^{n} x_{ij}y_{ij}, \quad Y_{\max} = \max_{Y \in D} R(X,Y) \tag{1}$$

2.2. The Quantum Encoding of Arbitrary Initial State

Many results in quantum information theory require the generation of specific quantum states, or the implementation of encoding a given vector onto qubits for computing, such as the quantum swap-test similarity calculation. However, many states cannot be efficiently realized. There are many methods focusing on the generation of quantum states, for example, Huelga et al. implemented an improved frequency standard experiment that requires the preparation of specific symmetric states on n qubits, where n is a parameter (number of ions) [38]. Kaye et al. proposed an algorithm, efficiently preparing the required symmetric state [39]. For any normalized vector X_i, $X_i = (x_{0i}, x_{1i}, \cdots, x_{n-1i})$. We suppose that $n = 2^k$; otherwise, the vector can be filled with 0. Long et al. proposed an algorithm for preparing arbitrary superposition quantum states with time complexity $O(n(\log n)^2)$, space complexity $O((\log n))$ and no additional auxiliary bits required [40]. The general method of encoding quantum states is shown in Figure 3.
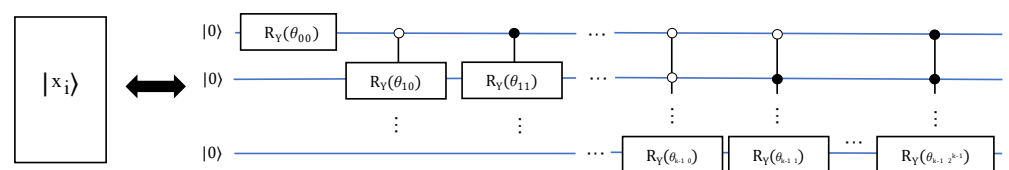

Figure 3. Diagram of encoding a vector to the required quantum state.

Assuming that the probability corresponding to each basis is p_i, $p_i = \frac{|x_{0i}|^2}{\|X_i\|^2}$ the number of required qubit is k. When the first qubit is 0, the probability is $p_0 = \sum_{i=0}^{2^{k-1}-1} p_i = \cos\theta_{00}^2$, which is 1, $p_1 = \sum_{i=2^{k-1}}^{2^k-1} p_i = \sin\theta_{00}^2$. It can just be represented by the quantum gate

$$Ry(2\theta_{00}) = e^{-i\theta_{00}\sigma_y} = \begin{bmatrix} \cos\theta_{00} & -\sin\theta_{00} \\ \sin\theta_{00} & \cos\theta_{00} \end{bmatrix} \tag{2}$$

When the second qubit is 0, the probability is

$$p_{00} = \sum_{i=0}^{2^{k-2}-1} p_i = \cos\theta_{00}^2 \cos\theta_{10}^2,$$

$$p_{01} = \sum_{i=2^{k-2}}^{2^{k-1}-1} p_i = \cos\theta_{00}^2 \sin\theta_{10}^2 \tag{3}$$

When the second qubit is 1, the probability is

$$p_{10} = \sum_{i=2^{k-1}}^{2^{k-1}+2^{k-2}-1} p_i = \sin\theta_{00}^2 \cos\theta_{11}^2,$$

$$p_{11} = \sum_{i=2^{k-1}+2^{k-2}}^{2^{k}-1} p_i = \sin\theta_{00}^2 \sin\theta_{11}^2 \tag{4}$$

Repeating the previous steps according to the rule of dichotomy until all k-bit values are traversed, the circuit in Figure 3 can be obtained. In addition, for a vector set S of m elements, $S = \{X_0, X_1, \cdots, X_{m-1}\}$, where X_i is a n-dimensional vector defined as before. Without loss of generality, we can set m is an integer power of 2, $m = 2^l$, thus an equal probability quantum superposition state $|\phi\rangle$, including labels about these vectors $|X_0\rangle, |X_1\rangle \cdots, |X_{m-1}\rangle$, can be constructed. The number of qubit is $\lceil \log n \rceil + \lceil \log m \rceil$, and the circuit time complexity is $O((mn)(\log_2(m)\log_2(n))^2)$, where the $\lceil \log m \rceil$ qubits are used as the label for state $|\phi\rangle$. The quantum circuit is shown in Figure 4. When the dimensions of the elements in S are inconsistent, the number of bits required is inconsistent. In this case, we take the number of bits required by the vector with the largest dimension as the number of bits for each U_i circuit:

$$|X_i\rangle = \mathcal{A}_1|0\rangle^{\otimes k} = \frac{1}{\sqrt{x_{0i}^2 + x_{1i}^2 + \cdots + x_{n-1i}^2}} \sum_{k=0}^{n-1} |x_{ki}| \, |k\rangle \tag{5}$$

$$|\phi\rangle = \mathcal{A}_2|0\rangle^{\otimes(k+l)} = \frac{1}{\sqrt{m}} \sum_{i=0}^{m-1} |i\rangle |X_i\rangle \tag{6}$$

where $\mathcal{A}_1$ and $\mathcal{A}_2$ represents the unitary operations in Figures 3 and 4 respectively, and U_i stands for the operation to prepare quantum state $|X_i\rangle$.
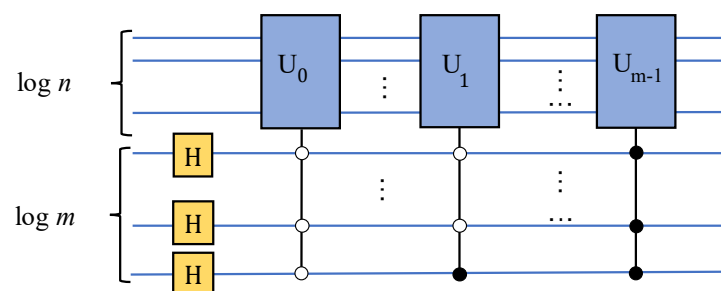

Figure 4. Diagram of constructing a equal probability quantum superposition state.

2.3. Similarity Calculation Based on Swap Test

For comparing the similarity of two states $|\phi\rangle$ and $|\psi\rangle$, where $|\psi\rangle$ is a superposition state consists of p n-dimensional vectors, and the quantum swap-test circuit is available [41,42]:

$$|\psi\rangle = \frac{1}{\sqrt{p}} \sum_{i=0}^{p-1} |i\rangle |Y_i\rangle, \quad |Y_i\rangle = \frac{1}{\sqrt{y_{0i}^2 + y_{1i}^2 + \cdots + y_{n-1i}^2}} \sum_{k=0}^{n-1} |y_{ki}| |k\rangle \tag{7}$$

The result can be obtained by measuring the first register and the labels of state $|\phi\rangle$ and $|\psi\rangle$, as shown in Figure 5. The measurement probability is $P_0 = (1+ |\langle \phi | \psi \rangle|^2)/2$, and $P_1 = 1 - P_0$. Where P_0 represents the probability when the measuring result is $|0\rangle$. Thus, the fidelity-based similarity of vectors ψ and ϕ can be obtained by $\langle \phi | \psi \rangle = \sqrt{(2P_0 - 1)}$.

$$
\begin{aligned}
P_0 &= \frac{1}{4} \|(I + SWAP)(Q|0\rangle|\psi\rangle|\phi\rangle)\|^2 \\
&= \frac{1 + \mathrm{Re}(\langle \psi \otimes \phi | SWAP | \psi \otimes \phi \rangle)}{2} \\
&= \frac{1 + |\langle \phi | \psi \rangle|^2}{2}
\end{aligned}
\tag{8}
$$

and $Q = (H \otimes I^{\otimes n})(C - SWAP)(H \otimes I^{\otimes n})$ represents the operator of the swap test.

$$
\begin{aligned}
U_{swap}|0\rangle|\psi\rangle|\phi\rangle &= \frac{|0\rangle + |1\rangle}{2} |\psi \otimes \phi\rangle + \frac{|0\rangle - |1\rangle}{2} SWAP|\psi \otimes \phi\rangle \\
&= \frac{1}{\sqrt{mp}} \sum_{i}^{m-1} \sum_{j}^{p-1} (|0\rangle|ij\rangle \frac{|X_i \otimes Y_j\rangle + SWAP|X_i \otimes Y_j\rangle}{2} \\
&\quad + |1\rangle|ij\rangle \frac{|X_i \otimes Y_j\rangle - SWAP|X_i \otimes Y_j\rangle}{2})
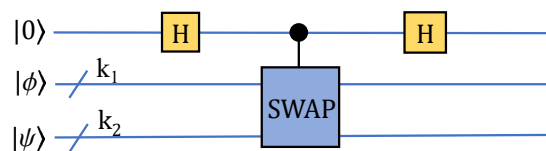\end{aligned}
\tag{9}
$$

Figure 5. Diagram of similarity calculation based on fidelity.

The measurement probability of quantum state $|0\rangle|ij\rangle(|Y_j \otimes X_i\rangle + SWAP|Y_j \otimes X_i\rangle)$ is

$$
\begin{aligned}
P_0^{ij} &= \frac{1}{4mp} \|(I + SWAP)(Q|0\rangle|Y_j\rangle|X_i\rangle)\|^2 \\
&= \frac{1 + \mathrm{Re}(\langle X_i \otimes Y_j | SWAP | X_i \otimes Y_j \rangle)}{2mp} \\
&= \frac{1 + |\langle X_i | Y_j \rangle|^2}{2mp}
\end{aligned}
\tag{10}
$$

when the first register is $|0\rangle$ and we can obtain the probability P_0^{ij} by measuring the labels $|i\rangle$ and $|j\rangle$, i.e., similarity of $\langle X_i | Y_j \rangle = \sqrt{(2mp * P_0^{ij} - 1)}$. For the swap operation, it can be decomposed as three CNOT gates, thus the swap test can be represented as the gate set of Hadamard and CNOT. A simplified equivalent circuit for estimating the expectation value of the swap-test operation is proposed in [43], and the circuit employs one CNOT gate followed by one Hadamard gate, with both qubits being measured, which diagonalizes the swap gate, when it is treated as a Hermitian observable with eigenvalues 1 or -1

(Figure 6). The output distribution that constructs an estimator for tested states $|\phi\rangle$ and $|\psi\rangle$ is $|\langle\psi|\phi\rangle|^2$.
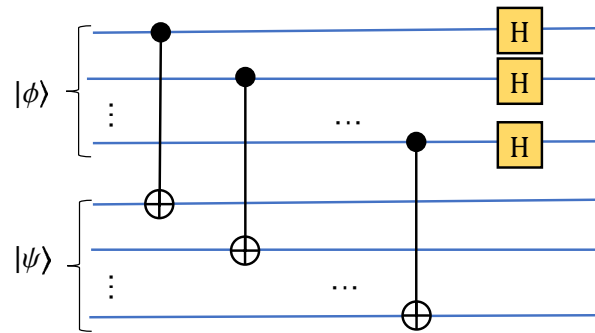

Figure 6. Diagram of simplified swap-test circuit.

For example, when $|\phi\rangle = \alpha_1|0\rangle + \beta_1|1\rangle$, $|\psi\rangle = \alpha_2|0\rangle + \beta_2|1\rangle$, after a CNOT and Hadamard operation, the result is

$$\frac{1}{\sqrt{2}}[(\alpha_1\alpha_2 + \beta_1\beta_2)|00\rangle + (\alpha_1\alpha_2 - \beta_1\beta_2)|10\rangle + (\alpha_1\beta_2 + \beta_1\alpha_2)|01\rangle + (\alpha_1\beta_2 - \beta_1\alpha_2)|11\rangle] \quad (11)$$

The probability of measuring the $|11\rangle$ state is

$$P_{11} = \frac{|\alpha_1\beta_2 - \beta_1\alpha_2|^2}{2} = \frac{(\alpha_1\beta_2 - \beta_1\alpha_2)(\alpha_1\beta_2 - \beta_1\alpha_2)^*}{2} \quad (12)$$

The probability amplitudes in the input qubits obey $|\alpha_1|^2 + |\beta_1|^2 = 1$ and $|\alpha_2|^2 + |\beta_2|^2 = 1$.

$$\begin{aligned} 1 - P_{11} &= \frac{2 - |\alpha_1|^2|\beta_2|^2 - |\beta_1|^2|\alpha_2|^2 + \alpha_1\beta_1^*\alpha_2^*\beta_2 + \alpha_1^*\beta_1\alpha_2\beta_2^*}{2} \\ &= \frac{1 + |\alpha_1|^2|\alpha_2|^2 + |\beta_1|^2|\beta_2|^2 + \alpha_1\beta_1^*\alpha_2^*\beta_2 + \alpha_1^*\beta_1\alpha_2\beta_2^*}{2} \end{aligned} \quad (13)$$

Because of the following relationship

$$\begin{aligned} |\langle\psi|\phi\rangle|^2 &= (\alpha_1\alpha_2^* + \beta_1\beta_2^*)(\alpha_1^*\alpha_2 + \beta_1^*\beta_2) \\ &= |\alpha_1|^2|\alpha_2|^2 + |\beta_1|^2|\beta_2|^2 + \alpha_1\beta_1^*\alpha_2^*\beta_2 + \alpha_1^*\beta_1\alpha_2\beta_2^* \end{aligned} \quad (14)$$

the probability of the swap test is $P = \frac{1 + |\langle\psi|\phi\rangle|^2}{2}$ corresponds to that in Equation (13).

2.4. Extraction of Measurement Results

The results of quantum computing are in the amplitude of the superposition state. To obtain these results, the most direct method is to perform a large number of measurements, but it has no acceleration advantage, so the similarity calculation based on amplitude amplification and amplitude estimation can bring a quadratic acceleration advantage. Let $|\Psi\rangle = \mathcal{A}|0\rangle$ denote the state obtained by applying $\mathcal{A}$ to the initial zero state. The amplification process is realized by repeatedly applying the following unitary operator [44,45] on the state $|\Psi\rangle$:

$$\mathbf{Q} = \mathbf{Q}(\mathcal{A}, \chi) = -\mathcal{A}\mathbf{S}_0\mathcal{A}^{-1}\mathbf{S}_\chi \quad (15)$$

Here, the operator $\mathbf{S}_\chi$ conditionally changes the sign of the amplitudes of the good states,

$$|x\rangle \longmapsto \begin{cases} -|x\rangle \text{ if } \chi(x) = 1 \\ |x\rangle \text{ if } \chi(x) = 0, \end{cases}$$

while the operator $\mathbf{S}_0$ changes the sign of the amplitude if and only if the state is the zero state $|0\rangle$. The operator U_{swap} is well-defined since we assume that $\mathcal{A}$ uses no measurements and, therefore, $\mathcal{A}$ has an inverse. The usefulness of operator U_{swap} stems from its simple action on the subspace $\mathcal{H}_\Psi$ spanned by the vectors $|\Psi_1\rangle$ and $|\Psi_0\rangle$. If we want to obtain the amplitude of $|0\rangle|ij\rangle(|X_i \otimes Y_j\rangle + SWAP|X_i \otimes Y_j\rangle)$, $\sqrt{P_0^{ij}}$, then the amplitude estimation is available, and we can write the result as

$$
\begin{aligned}
|\gamma\rangle = U_{swap}\mathcal{A}_2\mathcal{A}_1|0\rangle = {} & \sqrt{P_0^{ij}}(|0\rangle|ij\rangle(|X_i \otimes Y_j\rangle + SWAP|X_i \otimes Y_j\rangle))\\
& + \sqrt{1 - P_0^{ij}}(\sum_{s=0,s\neq i}^{m-1}\sum_{t=0,t\neq j}^{p-1}|0\rangle|st\rangle(|X_s \otimes Y_t\rangle + SWAP|X_s \otimes Y_t\rangle)\\
& + \sum_{s=0}^{m-1}\sum_{t=0}^{p-1}|1\rangle|st\rangle(|X_s \otimes Y_t\rangle - SWAP|X_s \otimes Y_t\rangle))
\end{aligned}
\tag{16}
$$

If we remark that $\widetilde{\mathcal{A}} = U_{swap}\mathcal{A}_2\mathcal{A}_1$, $\widetilde{\mathbf{Q}} = -\widetilde{\mathcal{A}}\mathbf{S}_0\widetilde{\mathcal{A}}^{-1}\mathbf{S}_\chi$, $\sqrt{P_0^{ij}} = \sin\theta/2$, $\theta \in [0, \pi]$, and we can obtain

$$
\begin{aligned}
|\gamma\rangle = \widetilde{\mathcal{A}}|0\rangle|0\rangle|0\rangle = {} & \sin\theta/2(|0\rangle|ij\rangle(|X_i \otimes Y_j\rangle + SWAP|X_i \otimes Y_j\rangle))\\
& + \cos\theta/2(\sum_{s=0,s\neq i}^{m-1}\sum_{t=0,t\neq j}^{p-1}|0\rangle|st\rangle(|X_s \otimes Y_t\rangle + SWAP|X_s \otimes Y_t\rangle)\\
& + \sum_{s=0}^{m-1}\sum_{t=0}^{p-1}|1\rangle|st\rangle(|X_s \otimes Y_t\rangle - SWAP|X_s \otimes Y_t\rangle))
\end{aligned}
\tag{17}
$$

$\mathbf{S}_\chi|\gamma\rangle = -\sin\theta/2|\gamma_1\rangle + \cos\theta/2|\gamma_0\rangle$, and after k interactions, the result is

$$
\widetilde{\mathbf{Q}}^k|\gamma\rangle = \cos((2k+1)\theta)|\gamma_0\rangle + \sin((2k+1)\theta)|\gamma_1\rangle
\tag{18}
$$

By amplitude estimation (Figure 7), we can obtain an estimation value of $\widetilde{\theta}$. Thus $\sqrt{P_0^{ij}} = \sin\widetilde{\theta}/2$, where k is an integer power of 2 with respect to t. Thus, we can obtain the similarity calculation of each feature vector of the target image and the feature vector of matching images, and it can bring a quantum quadratic acceleration advantage compared with direct measurement.
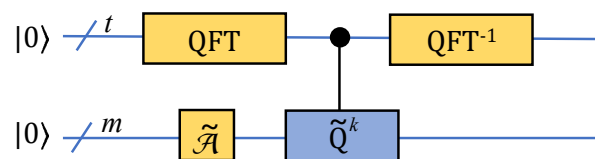

Figure 7. Diagram of amplitude estimation.

3. Results and Discussion

For exhibiting the effectiveness of the hybrid quantum image matching algorithm, we implemented the experiments based on an open NWPU-RESISC45 dataset, a benchmark for remote sensing image scene recognition created by Northwestern Polytechnic University, and another open handwritten character dataset, Mnist. We used the SIFT feature points of these datasets in experiments, and applied classical cosine similarity and the quantum swap test for calculation. An example of image matching is shown in Figure 8. Figure 8a represents the template image; Figure 8b, the references to be matched; and Figure 8c, the result of the feature matching.
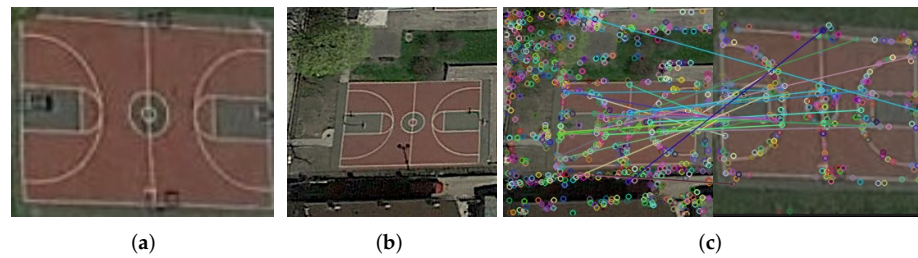
Figure 8. Example of matching. (**a**) This is template image; (**b**) this is the reference to be matched; (**c**) this is result of feature matching.

Experiments are implemented with IBM qasm-simulator, which supports 32-qubit and limited circuit depth. When the number of feature points is small, we apply the circuit in Figure 4; otherwise, the circuit in Figure 3 can be used to prepare the state one by one. The similarity is calculated according to the ratio between the nearest neighbor distance $d_1(R_i, S_i)$, and the next nearest neighbor distance $d_2(R_i, S_j)$ is less than a certain threshold. The experiments use the cosine distance, and we obtain quantum and classical results under different thresholds, respectively:

$$\frac{d_1(R_i, S_i)}{d_2(R_i, S_j)} < threshold \tag{19}$$

There are two evaluation indicators of the matching system effect, recall and precision. The recall represents the proportion of matched feature points among all feature points, which stands for the recall rate of the system. The precision stands for the proportion of correctly matched feature points among all the matched feature points, which represents the matching accuracy of the matching system:

$$Recall = \frac{\mid correct\ matches \mid}{\mid correspondences \mid} \tag{20}$$

$$Precision = \frac{\mid correct\ matches \mid}{\mid correct\ matches \mid + \mid false\ matches \mid} \tag{21}$$

where $\mid \cdot \mid$ represents the number of elements in the set, and *correspondences* stands for the set of all feature points that need to be matched in template images.

Experimental results and analysis: Figure 9 is the result of NWPU-RESISC45 dataset. From Figure 9a–c, the horizontal axis represents the threshold, and the vertical axis corresponds to the values of recall, precision and F1, respectively. In the experiments, the quantum results are obtained by random sampling of the measured results. Due to the influence of the sampling accuracy, the number of misidentified points in the experiments will be more than those of the classicals. Under the same threshold, the recall of quantum swap-test exhibits better performance than that of the classicals. As the threshold increases, the matching accuracy of similarity calculation based on classical cosine similarity and quantum swap-test also increases. That is because the upper bound of the ratio between the nearest neighbor and the next nearest neighbor becomes larger as the threshold increases, which leads to an increase in the number of mismatched points, and thus the recall rate (Figure 9a).
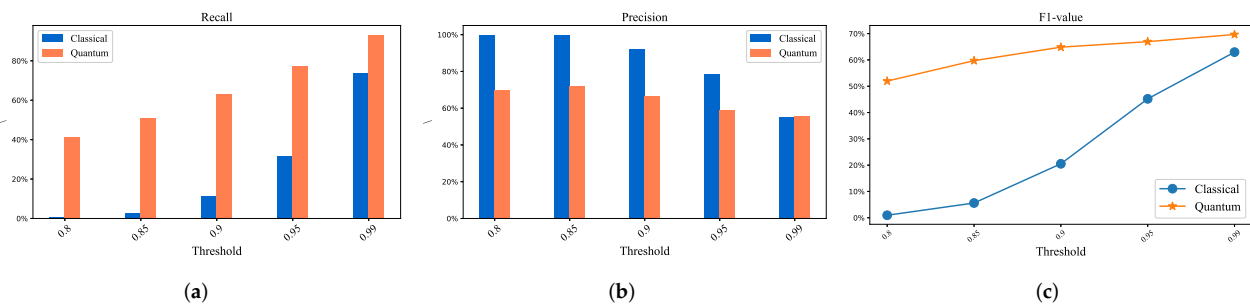
Figure 9. The comparative experimental results of classical and hybrid quantum method based on NWPU-RESISC45 dataset. (**a**) Relationship between threshold and Recall. (**b**) Relationship between threshold and Precision. (**c**) Relationship between threshold and F1-value.

Obviously, the greater the number of misidentified points, the lower the accuracy. Thus, the quantum precision is slightly lower than classical precision, as shown in Figure 9b. Precision and recall are inherently contradictory; therefore, a comprehensive evaluation index F-score is required, which is the average of the weighted sum of precision and recall. When $\alpha = 1$, it is called the F1-score.

$$F = \frac{(\alpha^2 + 1)P * R}{\alpha^2(P + R)} \tag{22}$$

The larger the value of F, the better the matching effect. The trend line in Figure 9c shows that the hybrid algorithm based on the swap test is better than the classical cosine similarity calculation method.

Similar to Figure 9, a–c in Figure 10 stands for recall, precision, and F1, in turn. The results show that the same conclusion can be obtained on the Mnist dataset, which proves that the hybrid algorithm is better than the classical algorithm again.
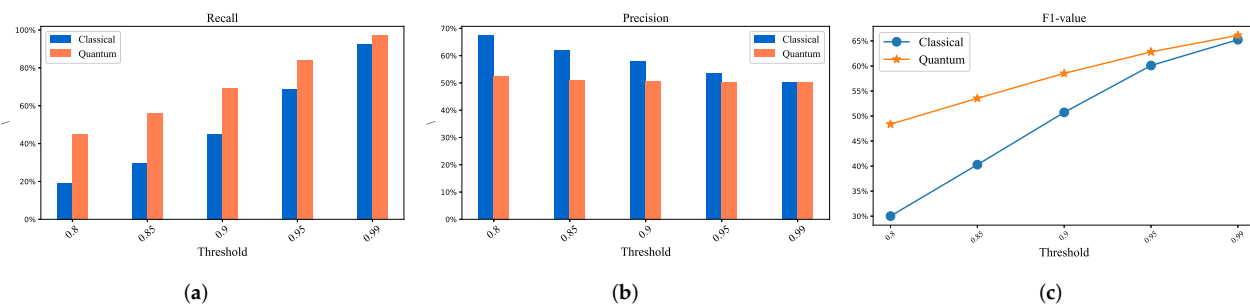

Figure 10. The comparative experimental results of classical and hybrid quantum method based on Mnist dataset. (**a**) Relationship between threshold and Recall. (**b**) Relationship between threshold and Precision. (**c**) Relationship between threshold and F1-value.

4. Conclusions

By conducting a series of experiments based on classical and our hybrid methods, some useful conclusions can be achieved:

(1) The validity and accuracy of the hybrid method based SIFT are proved through experimental verification, and it was also demonstrated that the hybrid algorithm performs better than the classicals.

(2) By the amplitude encoding, the quantum resources consumption is logarithmically less than that of classical resources, making the space complexity exponentially reduced. The computing resources can be reduced from $O(mpn^2)$ to $O(2log(n) + logp + logm)$.

(3) Making use of quantum entanglement and superposition properties, the time complexity of the similarity calculation between two images with s and t feature descriptors is reduced from $O(st)$ to $O(1)$. Additionally, the value of st is usually very large.

(4) The acquisition of quantum computing results requires measurement, and direct measurement does not show the quantum advantage. Thus, it is possible to obtain the similarity by estimating value of θ, $\sin^2 \widetilde{\theta}/2 = P_0^{ij}$, which brings quadratic acceleration advantage $O(\sqrt{mp})$.

This paper theoretically proves the quadratic acceleration advantage of the quantum hybrid algorithm and logarithmic resource consumption advantage over classicals. At the same time, it is also experimentally proved that the comprehensive performance of the hybrid algorithm performs better than that of the classical algorithm, showing that the quantum hybrid algorithm is used in machine learning, especially the great potential in the field of image matching. With the development of quantum computers with large number of bits, high bit connectivity, and long decoherence time, the limitations of massive data calculations in the research fields, such as images processing, are eliminated by quantum computing.

References

1. Jiang, X.; Ma, J.; Xiao, G.; Shao, Z.; Guo, X. A review of multimodal image matching: Methods and applications. *Inf. Fusion* **2021**, *73*, 22–71. [CrossRef]
2. Remondino, F.; Spera, M.G.; Nocerino, E.; Menna, F.; Nex, F. State of the art in high density image matching. *Photogramm. Rec.* **2014**, *29*, 144–166. [CrossRef]
3. Kong, Y.; Yu, T. A deep neural network model using random forest to extract feature representation for gene expression data classification. *Sci. Rep.* **2018**, *8*, 16477. [CrossRef] [PubMed]
4. Li, L.; Liang, J.; Weng, M.; Zhu, H. A multiple-feature reuse network to extract buildings from remote sensing imagery. *Remote Sens.* **2018**, *10*, 1350. [CrossRef]
5. Jayanthi, N.; Indu, S. Comparison of image matching techniques. *Int. J. Latest Trends Eng. Technol.* **2016**, *7*, 396–401.
6. Zitova, B.; Flusser, J. Image registration methods: A survey. *Image Vis. Comput.* **2003**, *21*, 977–1000. [CrossRef]
7. Tuytelaars, T.; Mikolajczyk, K. *Local Invariant Feature Detectors: A Survey*; Foundations and Trends® in Computer Graphics and Vision; Now Publishers Inc.: Shanghai, China, 2008.
8. Timo, O.; Matti, P.; David, H. A comparative study of texture measures with classification based on featured distributions. *Pattern Recognit.* **1996**, *29*, 51–59.
9. Dou, J.; Qin, Q.; Tu, Z. Robust image matching based on the information of SIFT. *Optik* **2018**, *171*, 850–861. [CrossRef]
10. Liu, H.; Yan, S. Common visual pattern discovery via spatially coherent correspondences. In Proceedings of the 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Francisco, CA, USA, 13–18 June 2010; pp. 1609–1616.
11. Li, X.; Hu, Z. Rejecting mismatches by correspondence function. *Int. J. Comput. Vis.* **2010**, *89*, 1–17. [CrossRef]
12. Ma, J.; Zhao, J.; Jiang, J.; Zhou, H.; Guo, X. Locality preserving matching. *Int. J. Comput. Vis.* **2019**, *127*, 512–531. [CrossRef]
13. Bian, J.; Lin, W.-Y.; Matsushita, Y.; Yeung, S.-K.; Nguyen, T.-D.; Cheng, M.-M. GMS: Grid-Based Motion Statistics for Fast, Ultra-Robust Feature Correspondence. *Int. J. Comput. Vis.* **2020**, *128*, 1580–1593. [CrossRef]
14. Ma, J.; Jiang, X.; Fan, A.; Jiang, J.; Yan, J. Image matching from handcrafted to deep features: A survey. *Int. J. Comput. Vis.* **2021**, *129*, 23–79. [CrossRef]

15. Ma, J.; Fan, A.; Jiang, X.; Xiao, G. Feature Matching via Motion-Consistency Driven Probabilistic Graphical Model. *Int. J. Comput. Vis.* **2022**, *130*, 2249–2264. [CrossRef]

16. Zheng, L.; Xiao, G.; Shi, Z.; Wang, S.; Ma, J. MSA-Net: Establishing Reliable Correspondences by Multiscale Attention Network. *IEEE Trans. Image Process.* **2022**, *31*, 4598–4608. [CrossRef]

17. Tarjan, R.E. *Data Structures and Network Algorithms*; SIAM: Philadelphia, PA, USA, 1983.

18. Barnea, D.I.; Silverman, H.F. A class of algorithms for fast digital image registration. *IEEE Trans. Comput.* **1972**, *100*, 179–186. [CrossRef]

19. Biamonte, J.; Wittek, P.; Pancotti, N.; Rebentrost, P.; Wiebe, N.; Lloyd, S. Quantum machine learning. *Nature* **2017**, *549*, 195–202. [CrossRef]

20. Dunjko, V.; Briegel, H.J. Machine learning & artificial intelligence in the quantum domain: A review of recent progress. *Rep. Prog. Phys.* **2018**, *81*, 074001.

21. Rebentrost, P.; Mohseni, M.; Lloyd, S. Quantum support vector machine for big data classification. *Phys. Rev. Lett.* **2014**, *113*, 130503. [CrossRef]

22. Blank, C.; Park, D.; Rhee, J.-K.K.; Petruccione, F. Quantum classifier with tailored quantum kernel. *Phys. Rev. Lett. Quantum Inf.* **2020**, *6*, 41. [CrossRef]

23. Tezuka, H.; Nakaji, K.; Satoh, T.; Yamamoto, N. Grover search revisited: Application to image pattern matching. *Phys. Rev. A* **2022**, *105*, 032440. [CrossRef]

24. Curtis, D.; Meyer, D.A. Towards quantum template matching. *Quantum Commun. Quantum Imaging SPIE* **2004**, *5161*, 134–141.

25. Niroula, P.; Nam, Y. A quantum algorithm for string matching. *NPJ Quantum Inf.* **2021**, *7*, 37. [CrossRef]

26. Yang, Y.-G.; Zhao, Q.-Q.; Sun, S.-J. Novel quantum gray-scale image matching. *Optik* **2015**, *126*, 3340–3343. [CrossRef]

27. Luo, G.; Zhou, R.-G.; Liu, X.; Hu, W.; Luo, J. Fuzzy matching based on gray-scale difference for quantum images. *Int. J. Theor. Phys.* **2015**, *57*, 2447–2460. [CrossRef]

28. Jiang, N.; Dang, Y.; Wang, J. Quantum image matching. *Quantum Inf. Process.* **2016**, *15*, 3543–3572. [CrossRef]

29. Jiang, N.; Dang, Y.; Zhao, N. Quantum image location. *Int. J. Theor. Phys.* **2016**, *55*, 4501–4512. [CrossRef]

30. Dang, Y.; Jiang, N.; Hu, H.; Zhang, W. Analysis and improvement of the quantum image matching. *Quantum Inf. Process.* **2017**, *16*, 269. [CrossRef]

31. Zhou, R.-G.; Liu, X.; Zhu, C.; Wei, L.; Zhang, X.; Ian, H. Similarity analysis between quantum images. *Quantum Inf. Process.* **2018**, *17*, 121. [CrossRef]

32. Montanaro, A. Quantum pattern matching fast on average. *Algorithmica* **2017**, *77*, 16–39. [CrossRef]

33. Ng, P.C.; Henikoff, S. SIFT: Predicting amino acid changes that affect protein function. *Nucleic Acids Res.* **2003**, *31*, 3812–3814. [CrossRef]

34. Ke, Y.; Sukthankar, R. PCA-SIFT: A more distinctive representation for local image descriptors. In Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004, Washington, DC, USA, 27 June–2 July 2004; Volume 2, p. II.

35. Kumar, P.; Henikoff, S.; Ng, P.C. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat. Protoc.* **2009**, *4*, 1073–1081. [CrossRef] [PubMed]

36. Chambon, S.; Crouzil, A. Similarity measures for image matching despite occlusions in stereo vision. *Pattern Recognit.* **2011**, *44*, 2063–2075. [CrossRef]

37. Xuecheng, L. Entropy, distance measure and similarity measure of fuzzy sets and their relations. *Fuzzy Sets Syst.* **1992**, *52*, 305–318. [CrossRef]

38. Huelga, S.F.; Macchiavello, C.; Pellizzari, T.; Ekert, A.K.; Plenio, M.B.; Cirac, J.I. Improvement of Frequency Standards with Quantum Entanglement. *Phys. Rev. Lett.* **1997**, *79*, 3865–3868. [CrossRef]

39. Kaye, P.; Mosca, M. Quantum networks for generating arbitrary quantum states. In Proceedings of the International Conference on Quantum Information, Optical Society of America, PB28, Rochester, NY, USA, 13 June 2001.

40. Long, G.-L.; Sun, Y. Efficient scheme for initializing a quantum register with an arbitrary superposed state. *Phys. Rev. A* **2001**, *64*, 014303. [CrossRef]

41. Zhao, J.; Zhang, Y.-H.; Shao, C.-P.; Wu, Y.-C.; Guo, G.-C.; Guo, G.-P. Building quantum neural networks based on a swap test. *Phys. Rev. A* **2019**, *100*, 012334. [CrossRef]

42. Li, P.; Wang, B. Quantum neural networks model based on swap test and phase estimation. *Neural Netw.* **2020**, *130*, 152–164. [CrossRef]

43. Garcia-Escartin, J.C.; Chamorro-Posada, P. Swap test and Hong-Ou-Mandel effect are equivalent. *Phys. Rev. A* **2013**, *87*, 052330. [CrossRef]

44. Brassard, G.; Hoyer, P. An exact quantum polynomial-time algorithm for Simon's problem. In Proceedings of the Fifth Israeli Symposium on Theory of Computing and Systems, Ramat Gan, Israel, 17–19 June 1997; pp. 12–23.

45. Brassard, G.; Hoyer, P.; Mosca, M.; Tapp, A. Quantum amplitude amplification and estimation. *Contemp. Math.* **2002**, *305*, 53–74.