



A continuous-variable quantum-inspired algorithm for classical image segmentation

Akram Youssry^{1,2} · Ahmed El-Rafei³ · Ri-Gui Zhou⁴

Received: 1 May 2019 / Accepted: 9 October 2019 / Published online: 11 November 2019
© Springer Nature Switzerland AG 2019

Abstract

The probabilistic nature of quantum particles, state space, and the superposition principle are among the important concepts in quantum mechanics. A framework was previously developed by the authors that allowed to take advantage of these quantum aspects in the field of image processing. This was done by modeling each image's pixel by a two-state quantum system which allowed efficient single-object segmentation. However, the extension of the framework to multi-object segmentation would be highly complex and computationally expensive. In this paper, we propose a classical image segmentation algorithm inspired by the continuous-variable quantum theory that overcomes the challenges in extending the framework to multi-object segmentation. By associating each pixel with a quantum harmonic oscillator, the space of coherent states becomes continuous. Thus, each pixel can evolve from an initial state to any of the continuous coherent states under the influence of an external resonant force. The Hamiltonian operator is designed to account for this force and is derived from the features extracted at the pixel. Therefore, the system evolves from an initial ground state to a final coherent state depending on the image features. Finally by calculating the fidelity between the final state and a set of reference states representing the objects in the image, the state with the highest fidelity is selected. The collective states of all pixels produce the final segmentation. The proposed method is tested on a database of synthetic and natural images, and compared with other methods. Average sensitivity and specificity of 97.86% and 99.61% were obtained respectively indicating the high segmentation accuracy of the algorithm.

Keywords Quantum-inspired algorithms · Coherent states · Quantum harmonic oscillator · Signal processing

✉ Akram Youssry
akram.mohamed@student.uts.edu.au;
akram.youssry@eng.asu.edu.eg

Ahmed El-Rafei
ahmed.elrafei@eng.asu.edu.eg

Ri-Gui Zhou
rgzhou@shmtu.edu.cn

- ¹ Centre for Quantum Software and Information, Faculty of Engineering and Information Technology, University of Technology Sydney, Ultimo, NSW 2007, Australia
- ² Electronics and Communication Engineering Department, Faculty of Engineering, Ain Shams University, Cairo, Egypt
- ³ Engineering Physics and Mathematics Department, Faculty of Engineering, Ain Shams University, Cairo, Egypt
- ⁴ College of Information Engineering, Shanghai Maritime University, Shanghai, China

1 Introduction

Quantum image processing has been an active area of research recently. The usual tasks of image processing are performed utilizing the theory of quantum mechanics. This includes image representation (Yan et al. 2016), image matching (Jiang et al. 2016), similarity analysis (Zhou et al. 2018b), interpolation (Zhou et al. 2018a), denoising (Mastriani 2015a), coding (Chapeau-Blondeau and Belin 2016), watermarking (Li et al. 2016), and segmentation (Caraiman and Manta 2015). These attempts are based on representing the pixels of an image as qubits operated on via suitable quantum circuits. However, there are major challenges facing this approach that need further investigation as in Mastriani (2015b). On the other hand, another approach exists in which classical images are processed on classical computers using quantum-inspired models. This includes for example the classical image

segmentation algorithm in Youssry et al. (2015), and its application in segmenting biomedical retinal images (Youssry et al. 2016), quantum K -means (Casper et al. 2012), quantum Gaussian mixture models (Tanaka and Tsuda 2008), and quantum pattern recognition (Sergioli et al. 2016). Some of these methods can be ported to work on a quantum computer such as Youssry et al. (2015, 2016) which opens the path for many future applications.

Conventionally, discrete degrees of freedom of particles (such as the spin of an electron or the polarization of a photon) are used for encoding information in the form of qubits. In general, this approach faces some technological difficulties when it comes to the implementation. For example, we might have to control the polarization of a single photon system or the spin of a single electron which is not easily realizable. As a result, many of the developed ideas and algorithms have not been fully realized yet. Continuous-variable quantum information processing is another approach that depends on using the continuous degrees of freedom of the particle (such as position and momentum) for manipulating the information. This approach provides easier technological implementations, but can be challenging in porting algorithms and protocols from the discrete domain to the continuous domain. Examples of this category of information processing include quantum computation (Adcock et al. 2016), machine learning (Lau et al. 2016), quantum key distribution (Borelli et al. 2016), and identity authentication (Ma et al. 2016).

Image segmentation is an area of image processing that has many applications. It deals with delineating the significant objects in the image and isolating them from the background. Many methods exist to segment images including thresholding, edge detection, supervised and unsupervised machine learning, morphological methods, and deformable models. There exist also quantum-based methods for image segmentation. A short review on these techniques is given in Youssry et al. (2015). Some of the authors previously proposed a general framework that uses the theory of two-state quantum mechanics systems to process images (Youssry et al. 2015). Based on this framework, a general single-object image segmentation was developed and applied to generic images as well as in determining the vessel tree in retinal images. This algorithm showed high efficiency in segmenting images. Although the framework provides the theory for the extension to multi-object segmentation by utilizing discrete multi-state quantum systems, this extension is complex and computationally expensive. The continuous-variable quantum theory provides a solution to this challenge.

In this paper, we propose a new algorithm for image segmentation based on the continuous-variable coherent quantum states that occur in the theory of quantum harmonic oscillators. The work is built upon the framework

presented in Youssry et al. (2015), but extends to the case of multi-object segmentation. The paper starts in Section 2 with a brief theoretical overview essential for introducing the new methodology. Next, the proposed algorithm is presented in Section 3. After that, the materials used for testing as well as the obtained results are shown in Sections 4 and 5. In Section 6, the analysis and the significance of the results are discussed. Finally, the conclusion and the future perspectives are given in Section 7. The Appendices 1 and 2 include some additional proofs given for the sake of completeness.

2 Background

This section starts with a brief overview on the theory of quantum harmonic oscillators, needed for developing the proposed methodology. The details can be found in any standard reference on quantum mechanics or quantum optics such as Griffiths (2005) or Gerry and Knight (2005). Afterwards, a short review on the quantum fidelity measure is given. Finally, the performance measures used for evaluating the segmentation algorithm are discussed.

2.1 Quantum harmonic oscillator

The classical harmonic oscillator is a physical model that describes the motion of a particle under the influence of a restoring force, causing the particle to oscillate about its equilibrium position. A simple example of this model is a simple mass-spring system. The quantum harmonic oscillator (QHO) is the quantum analogue of the classical harmonic oscillator. However, the particle is microscopic and thus follows the rules of quantum mechanics. The QHO model can describe many physical systems such as: phonons (the quanta of lattice vibrations), electromagnetic radiation modes in a cavity, and vibration of diatomic molecules. The Hamiltonian of the QHO described in terms of the position and momentum operators is given by

$$H = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2\hat{x}^2. \quad (1)$$

\hat{p} and \hat{x} are the momentum and position operators, m is the mass of the particle, and ω is the angular frequency of oscillation. The Hamiltonian can then be inserted in Schrödinger's equation to obtain the wavefunctions and the energy levels. The lowest energy level (ground level) of the QHO is called the vacuum state denoted as $|0\rangle$ with energy $E_0 = \frac{1}{2}\hbar\omega$. This is the contrast of the classical picture where the lowest level has zero energy. The n th excited state of the QHO is denoted by $|n\rangle$, and has energy $E_n = \hbar\omega\left(n + \frac{1}{2}\right)$. An important feature to notice is that the number of energy levels is infinite opposed to other finite quantum systems

(such as the intrinsic spin of an electron). Therefore, the quantum states and operators are represented in the form of infinite-dimensional vectors and matrices. The states of a QHO are orthonormal $\langle n|m \rangle = \delta_{nm}$. The eigenstates $|n\rangle$ are also called number states in the field of quantum optics, as they represent a state of the electromagnetic field confined in a cavity with exactly n photons.

Two important operators in the theory of QHO are the creation and annihilation operators. The creation operator \hat{a}^\dagger is defined by its action on a number state $|n\rangle$ as

$$\hat{a}^\dagger |n\rangle = \sqrt{n+1} |n+1\rangle, \tag{2}$$

while the annihilation operator \hat{a} is defined as

$$\hat{a} |n\rangle = \sqrt{n} |n-1\rangle \tag{3}$$

$$\hat{a} |0\rangle = 0. \tag{4}$$

In other words, the creation and annihilation operators increase/decrease the energy of the QHO by one level. These operators are not Hermitian and thus do not represent physical observables, although they are related to the position and momentum of the oscillator. Finally, these operators do not commute,

$$[a, a^\dagger] = 1. \tag{5}$$

Another related operator is the number operator defined as

$$\hat{n} |n\rangle = n |n\rangle, \tag{6}$$

where

$$\hat{n} = \hat{a}^\dagger \hat{a}. \tag{7}$$

This operator is Hermitian and thus can be measured physically.

The aforementioned operators can be used to express the free field Hamiltonian of a QHO (no external forces)

$$H = \hbar\omega \left(\hat{n} + \frac{1}{2} \right) = \hbar\omega \left(\hat{a}^\dagger \hat{a} + \frac{1}{2} \right). \tag{8}$$

This is the matrix (Hilbert space) representation of the Hamiltonian operator, in contrast to the wavefunction (configuration space) representation in Eq. 1. The matrix form turns out to be the more suitable form in this paper.

As known from the postulates of quantum mechanics, a system is generally in a superposition of its basis states until a measurement is performed. An important superposition state of a QHO is the coherent state. It is defined as the eigenstate of the annihilation operator

$$\hat{a} |\alpha\rangle = \alpha |\alpha\rangle. \tag{9}$$

In order to satisfy this equation as well as the normalization condition of the state, it can be shown that the general form of the coherent state defined in terms of number states is

$$|\alpha\rangle = e^{-\frac{1}{2}|\alpha|^2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle. \tag{10}$$

A very important remark is that the eigenvalue α defining the coherent state is a complex number. Thus, there is a continuum set of coherent states rather than a discrete set. The real and imaginary parts of α are related to the position/momentum of the oscillating particle. In quantum optics, the real and imaginary parts are related to the field quadratures, while the magnitude describes the average number of photons in the field. Coherent states can be practically generated. For example, a strong laser field is considered a coherent state. The following are the most important mathematical properties of coherent states:

- They are not orthogonal,

$$\langle \alpha | \beta \rangle = e^{\frac{1}{2}(\alpha^* \beta + \alpha \beta^*)} e^{-\frac{1}{2}|\beta - \alpha|^2}. \tag{11}$$

In the limit of large magnitudes, the states tend to be orthogonal.

- They form an overcomplete basis

$$\int |\alpha\rangle \langle \alpha| d^2\alpha = \pi. \tag{12}$$

If the QHO is coupled to an applied external force $f(t)$, the Hamiltonian in the matrix representation takes the following form

$$H = \hbar\omega \left(\hat{a}^\dagger \hat{a} + \frac{1}{2} \right) + \hbar \left(f(t) \hat{a} + f^*(t) \hat{a}^\dagger \right). \tag{13}$$

If the force is in resonance with the oscillator, i.e.,

$$f(t) = f_0 e^{i\omega t}, \tag{14}$$

and the initial state of the harmonic oscillator is the ground state

$$|\psi(0)\rangle = |0\rangle, \tag{15}$$

then by solving Schrödinger’s equation, the final state can be shown as in Appendix 1 to take the form of coherent state

$$|\psi(t)\rangle = |\alpha\rangle, \alpha = -i e^{-i\omega t} f_0 t. \tag{16}$$

This equation is of great importance in this paper and will be central in the development of the proposed image segmentation algorithm.

2.2 Fidelity distance measure

Commonly in quantum information processing, it is required to define a distance metric between quantum states. There exist many measures in quantum information for this purpose such as the trace distance and fidelity (Nielsen and Chuang 2010). Fidelity is chosen in this paper as it can be easily computed with a closed form in the case of coherent states. But in principle any other distance measure can be used. The fidelity between two quantum states is defined as

$$F = \sqrt{\rho^{1/2} \sigma \rho^{1/2}}, \tag{17}$$

where ρ and σ are the density matrices corresponding to each state. If the two states are pure quantum states $|\psi\rangle$ and $|\phi\rangle$, the definition of fidelity reduces to

$$F = |\langle\phi|\psi\rangle|. \quad (18)$$

Since coherent states are pure states, then we can use this last definition for evaluating the fidelity between two coherent states. This leads to the following form of fidelity

$$F = |\langle\beta|\alpha\rangle| = e^{-\frac{1}{2}|\beta-\alpha|^2}. \quad (19)$$

Fidelity measures the overlap between the two states, and it satisfies the inequality

$$0 \leq F \leq 1. \quad (20)$$

A fidelity of 1 corresponds to a full overlap between the two states (i.e., they are identical), while a fidelity of 0 corresponds to a non-overlapping situation (i.e., orthogonal states).

2.3 Performance measures

The proposed image segmentation algorithm takes a classification approach, where each pixel is classified to belong to either the foreground or the background of one of the objects in the image. Accordingly, the sensitivity and specificity measures are suitable for evaluating the performance of the algorithm quantitatively (Youssry et al. 2015). Sensitivity measures the percentage of pixels of the object's foreground that are correctly classified by the algorithm as foreground. Specificity measures the percentage of the object's background that is correctly classified as background. Ideally, it is favorable to have both sensitivity and specificity of 100%. However, practically this may not be possible. In the case of multi-object segmentation, the algorithm may succeed in segmenting some objects and fails to segment others. Thus, the sensitivity and specificity for each individual object in the image are calculated to evaluate the performance in all cases.

3 Methods

In this section, the proposed methodology is presented. First, an overview is given on the quantum-based framework upon which the proposed image segmentation algorithm is built. This framework has been proposed recently in Youssry et al. (2015). The challenges of extending this framework to the multi-object case are discussed, as well as how the novel algorithm overcomes these challenges. After that, the detailed steps of the developed algorithm for multi-object image segmentation are elaborated.

3.1 Overview of the framework

The proposed algorithm follows the quantum-based framework developed in Youssry et al. (2015). In this framework, an analogy between the signal processing task required to be performed and quantum mechanics is formed. This allows transforming the signal processing problem into a problem that can be solved easily within the well-developed quantum mechanics theory. Afterwards, the obtained quantum solution can be transformed back to the signal processing domain. This idea is used to develop an image segmentation algorithm that was suitable for segmenting single-object images. A classification approach is adopted, where each pixel in the image is classified into one of the two possible classes: background or foreground of the object. In order to accomplish this task, each pixel in the image is associated to a two-level quantum system (qubit). The quantum system starts from an initial state and is evolved to a final state. By measuring the final state, the final outcome representing the class of the pixel is obtained. In order to reach a correct final state, the Hamiltonian operator is designed to be a function in the features of extracted from the pixel. In other words, the feature vector guides the quantum system to reach its correct final state. This requires estimating some parameters so that the features can be combined together in the Hamiltonian. This is done using a supervised learning method. A small window in the image is selected together with its ground truth, and both are used to estimate the parameters targeting the minimization of the error between the resulting segmentation of this window and its ground truth. After this learning phase, the obtained parameters do not change anymore for this image, and they can also be used to segment any other visually similar image.

A straightforward approach to extend this algorithm to the case of multi-object image segmentation is to use multi-level quantum systems (qudit). However, there are four problems in this approach. First, the complexity of computations will increase, as the state vector of an N -level quantum system is represented as an $N \times 1$ vector, and the quantum operators will be represented by $N \times N$ matrices. Since the framework is mainly designed to work on a classical computer, then this can form a bottleneck in the execution in the case of large number of objects. Second, it may be difficult to derive a closed-form solution for Schrödinger's equation in the general case (N -level system) as was proven for the qubit case (2-level system). In this case, the solution must be numerically obtained. When the number of levels increases, this again increases the overall complexity. Third, there is an important issue concerning the controllability of quantum systems. Not every Hamiltonian allows an arbitrary transition between states. Therefore, this issue must be taken into consideration while choosing the Hamiltonian form. Besides increasing

the difficulty of the design process, the result may be a Hamiltonian that does not correspond to an actual physical process. This may prevent the realization of the algorithm on a quantum computer. This opposes the case of the 2-level system where any Hamiltonian can be realized easily. Finally, the number of Hamiltonian parameters (degrees of freedom in the matrix representation) generally increases for larger systems which adds more complexity.

In principle, these challenges can be solved to obtain a generalized model for multi-object image segmentation. However, in this paper a novel model is proposed that does not face those challenges. Additionally, it can be generalized to any number of image objects without an increase in the overall complexity. The basic idea is to map each pixel in the image to a quantum harmonic oscillator system instead of a qudit system. The oscillator is initialized to the ground state. By applying an external resonant force, the oscillator evolves to a final state which will be a particular coherent state. By choosing the Hamiltonian parameters, the final state can be controlled. Therefore, image features are extracted at each pixel and combined together. Next, a training phase is performed to estimate the Hamiltonian parameters. A small window of the image and its ground truth are provided for this step. The training pixels should include representative pixels for all objects. The pixels of each object of the image in the ground truth are assigned to a particular coherent state referred to as the reference state in this paper. For instance, if the image contains $N - 1$ objects plus the background, then we need to define a set of N coherent states to be used as reference states. So, the background is treated as an object as well. Each pixel is associated to its corresponding reference state according to the ground truth segmentation. Consequently, the Hamiltonian will be trained such that it results in the evolution of all the pixels in the training set from the initial state (which is the ground state) to the final state (which should be the corresponding reference coherent state). Once the Hamiltonian is constructed, it is used afterwards without further change. It will be used to evolve the states of pixels in the testing set (the remaining image pixels that are outside the training set). In general, the final state of those test pixels may not coincide exactly with any of reference states. So, in order to determine the class/state of the pixel, the final state is compared with the whole set of reference coherent states representing each object. If the final state of system is close to an object's reference state, the pixel is classified as belonging to the foreground of this object. The quantum fidelity measure is used as a distance measure to quantify the closeness of the final state to any of the reference states.

The quantum harmonic oscillator is an infinite-dimensional quantum system. Working with number states

of the QHO will result in matrices that are of infinite-dimensions. So, it will be impossible to store and process them on a classical computer. However, although the system is infinite-dimensional, it is completely defined by a single complex-valued parameter α . All operations can be done by manipulating this parameter. This simple parameter can be stored and manipulated efficiently on a classical computer. Consequently, the representation of the quantum states as well the required quantum operators will be of a fixed size independent on the number of classes (objects) in the image. This solves the complexity problem in the original framework. The choice of the Hamiltonian generating the coherent states of the QHO solves the second problem as the solution exists in a closed form, as shown in Appendix 1, independent on the number of objects in the image. Additionally, this form guarantees that starting from the ground state, any final coherent state can be reached. Thus, the third challenge related to controllability is resolved. Moreover, the chosen Hamiltonian can be realized easily in the case of implementing on a quantum computer. Finally, as will be shown later, there are only three degrees of freedom in the Hamiltonian representation irrespective to the number of image objects.

3.2 Proposed algorithm

Based on the previous discussion, the algorithm shown in Algorithm 1 is proposed and it consists of five main steps discussed as below.

Algorithm 1 Multi-object image segmentation.

```

1: function SEGMENTIMAGE(Image)
2:   Define a set of reference states for each object as
   well as the background  $\{|\beta_i\rangle\}_{i=0}^{N-1}$ 
3:   for each pixel  $p$  in Image do
4:     Initialize state of  $p$  to  $|0\rangle$ 
5:     Extract feature vector  $\mathbf{x}$  at  $p$ 
6:     Evaluate  $T(\mathbf{x})$  to construct the Hamiltonian
7:     Solve Schrödinger's eq. to obtain final state  $|\alpha\rangle$ 
8:     Calculate all fidelities  $F_i = |\langle\alpha|\beta_i\rangle|$ 
9:     Select the maximum fidelity  $j = \arg \max_i F_i$ 
10:    OutputImage( $p$ )  $\leftarrow j$   $\triangleright$  classify pixel as
    foreground of object  $j$  or background
11:   end for
12:   return OutputImage
13: end function

```

3.2.1 Reference state preparation

The first step is to associate each object in the classical image to a predefined coherent state, that will be referred

to as the reference state. The choice of reference states is arbitrary. In this work, it is assumed that there are $N - 1$ objects to extract as well as the background. Thus, the reference states are chosen to take the form

$$|\beta_k\rangle = \frac{e^{i\frac{2\pi}{N}k}}{2\sin\left(\frac{\pi}{N}\right)}, k = 0, 1, \dots, N - 1 \quad (21)$$

Thus, on the complex plane formed of the real and imaginary components of the coherent state (phase space), the reference states are distributed evenly on a circle with radius $\left(2\sin\left(\frac{\pi}{N}\right)\right)^{-1}$ centered at the origin. Once the reference states are selected, they do not change anymore. The magnitude of these states is scaled to provide enough separation between them. This is related to the uncertainty principle, as experimentally the real and imaginary parts of the state cannot be measured simultaneously. The actual representation of a coherent state in the phase space is a small circle to reflect this uncertainty. In order to prevent the overlap of the reference states, the amplitudes are scaled. However, the scaling factor does not change the result of the algorithm as shown in Appendix 2; therefore, it can be chosen arbitrary. It may also be selected to have sufficiently large value such that the classical behavior dominates. Consequently, normal optical components can be used in this case for experimental realization. It is worth noting that the choice of the reference states affects the training of the Hamiltonian as the goal of this step is to find a set of optimal parameters of the Hamiltonian that results in evolving the initial states to the corresponding reference states as final states.

3.2.2 Pixel state initialization

In order to classify a particular pixel to one of the objects in the image, it is associated to a QHO system. The initial state of the pixel is taken to be the vacuum state.

$$|\psi(0)\rangle = |0\rangle. \quad (22)$$

3.2.3 Feature extraction

In this step, the feature vector is extracted from the pixel and is denoted by \mathbf{x} . Next, the features are combined together into a single feature, denoted by $T = T(\mathbf{x})$. This function can be chosen arbitrarily. In this paper, it is chosen to take the following form

$$T(\mathbf{x}) = p(\mathbf{x})u(p(\mathbf{x})) + 0.01. \quad (23)$$

$p(\bullet)$ is a polynomial function chosen to be of 6th order, (in some cases it was selected to have a Gaussian form depending on the problem) while $u(\bullet)$ is the Heaviside

unit step function which returns 1 if its argument is greater than 0 and returns 0 otherwise. This form assures that T is positive-valued, and thus can represent a time variable. It also guaranties that T^{-1} is not singular at any point. These two requirements are needed as will be shown in the next step of the algorithm. The coefficients of the polynomial are evaluated during the learning phase of the algorithm and are kept fixed afterwards through out the testing.

3.2.4 State evolution

The QHO system of the pixel is then allowed to evolve to its final state governed by the Hamiltonian

$$H = \hbar\omega\left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right) + \hbar f_0\left(\hat{a}e^{i\omega t} + \hat{a}^\dagger\hat{a}e^{-i\omega t}\right). \quad (24)$$

The final state under this form of the Hamiltonian is guaranteed to be the coherent state

$$|\psi(t)\rangle = |\alpha\rangle, \alpha = -ie^{-i\omega t}f_0t. \quad (25)$$

This equation shows that the final state depends on three factors: ω , f_0 , and t . But actually, there are only two degrees of freedom since the state is defined by a complex number that has only real and imaginary parts. The angular frequency is chosen arbitrary in this paper to be unity ($\omega = 1$). The time instant at which we observe the state is chosen to be $t = T(\mathbf{x})$. The force amplitude is chosen to be $f_0 = \frac{1}{2\sin\left(\frac{\pi}{N}\right)T(\mathbf{x})}$. Therefore, the final state of the QHO is

$$|\psi(T)\rangle = |\alpha(T)\rangle, \alpha(T) = \frac{-ie^{-iT}}{2\sin\left(\frac{\pi}{N}\right)}. \quad (26)$$

It is clear now that final state is a function in T which is itself a function of the image-based feature vector \mathbf{x} . In other words, the image-derived features at the pixel control the final state of the QHO.

3.2.5 Measurement

In order to obtain the final outcome for the pixel under consideration, the final state is observed. If it is one of the reference states defined in the first step of the algorithm, then the corresponding class is the final outcome. However, practically this may not happen. In this case, we have to choose one of the reference states and consequently choose the corresponding class for this pixel. An intuitive solution is to choose the reference state that is nearest to the pixel's final state. The fidelity measure in Eq. 19 can be used to obtain the distance between the final state and each of the

reference states. Thus, the pixel is classified to belong to object j if it satisfies that

$$j = \arg \max_i F_i = \arg \max_i |\langle \alpha | \beta_i \rangle|. \quad (27)$$

4 Materials

Following the work in Youssry et al. (2015), the proposed algorithm is tested on two datasets. The first dataset consists of synthetic images of geometric shapes with different types and different number of objects. The ground truth of these images is generated manually. Moreover, noise is applied to some of these images in order to test the performance of the algorithm in the presence of noise. The second dataset consists of natural images and is chosen from the publicly available image segmentation database of Alpert et al. (2007). This database provides images with single and double objects as well as the human segmentation for all images to test the accuracy of segmentation methods. The number of synthetic images are 19 images of which 11 of them are noisy images. Five natural images are included which adds up to a total of 24 images.

5 Results

The presented algorithm is tested on the datasets described in the previous section. For each image in the dataset, features are selected to better represent the objects in the image and used to estimate the Hamiltonian parameters. Many features are used depending on the nature of the image which are mostly simple features such as the gray-level, mean, and median. Nevertheless, in few cases, more complex features like the Morlet transform-based features are utilized. The features used are based on the nature of the object and background in the image. For most of the simple noiseless images (examples are first and third images in Fig. 1 and first image in Fig. 2), the pixels' gray values were sufficient to segment the object. However, in some cases there were no clear class separation between the object and background in the grayscale domain. Thus, other features were incorporated. For example, the second image in Fig. 2 is a textured image, so Morlet wavelet-based features were found to perform better than grayscale features. Moreover, in the presence of noise as in the images in Fig. 3, features from median filters with window size depending on the amount of noise were utilized and provided highly accurate segmentation. The feature selection is a design issue and should depend on the analysis of the images in order to

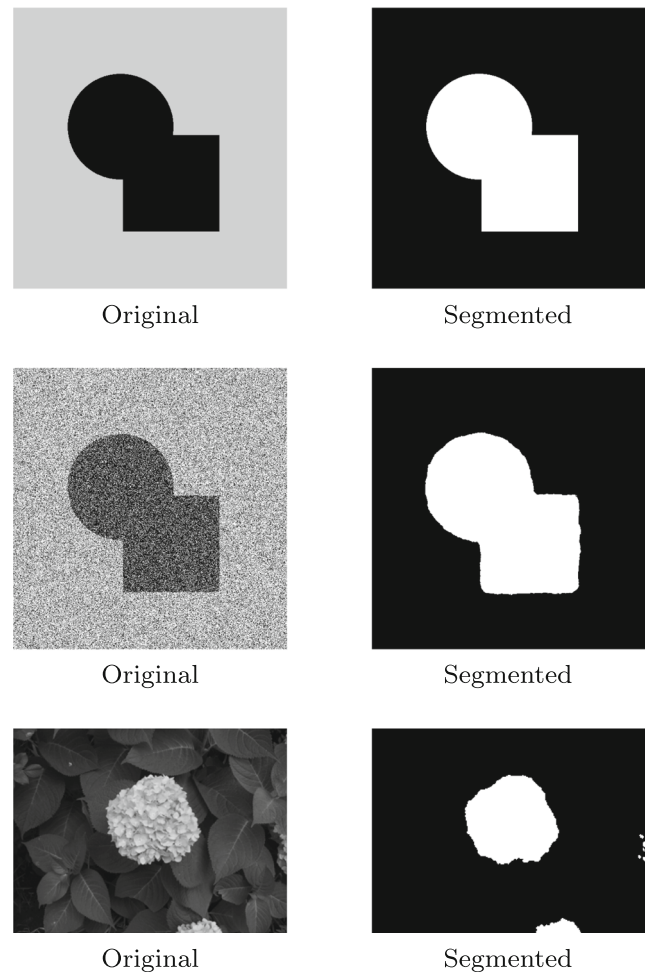


Fig. 1 Proposed algorithm segmentation of single-object images from Youssry et al. (2015)

estimate the Hamiltonian's parameters that will derive the evolution and in turn the segmentation.

The sensitivity and the specificity measures for each class in each image are calculated. First, the algorithm is applied to the images with single object that are used to validate the original framework (Youssry et al. 2015). This system is considered an enhancement to the previously introduced framework. Thus, the purpose of this step is to verify that the system can produce comparable results in case of single object before proceeding to multi-object images. Average sensitivity and specificity of 98.23% and 99.53% were obtained, respectively. This shows that the coherent state-based algorithm performs very efficiently in segmenting images with single object. Samples of segmented objects are shown in Fig. 1. In addition, the results are very similar to the former framework (sensitivity = 98.5% and specificity = 99.7%) which were shown to exceed other existing segmentation methods like active contours and

Fig. 2 Proposed algorithm segmentation of multi-object images

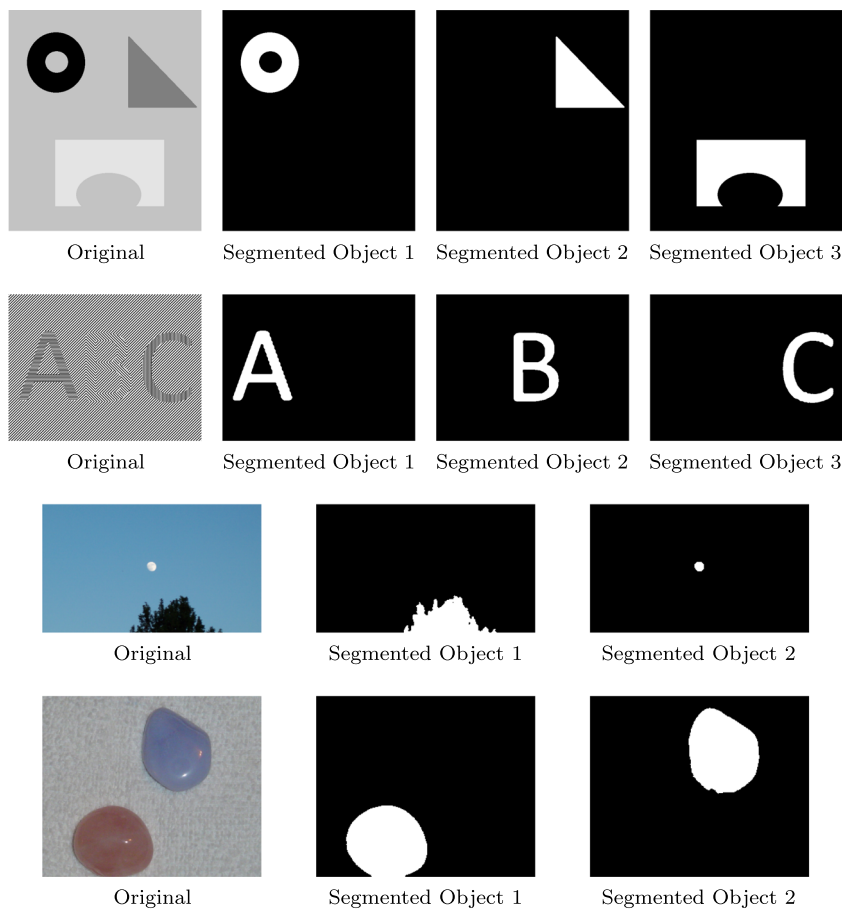
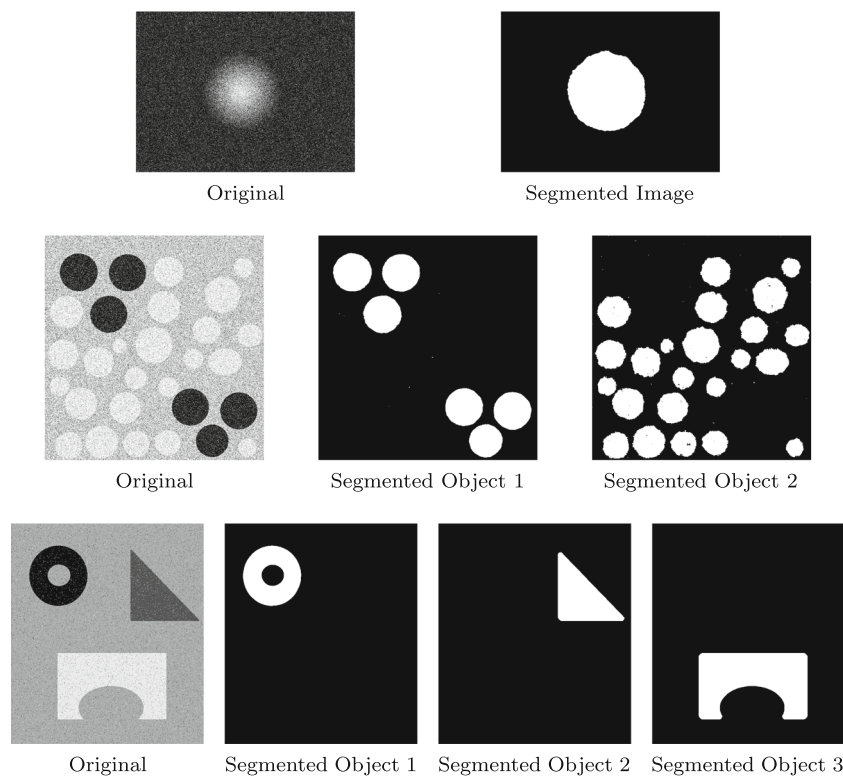


Fig. 3 Proposed algorithm segmentation of noisy images



graph cuts. In this paper, four methods are compared against the proposed algorithm. These segmentation algorithms are *K*-means clustering (Lloyd 1982), Otsu's multithreshold (Otsu 1979), lazy snapping graph cuts (Li et al. 2004), and random forests (Sommer et al. 2011). The results of applying the five algorithms to the entire dataset are summarized in Table 1. Regarding the single-object images, the sensitivities of both the quantum (98.23%) and random forests (98.55%) methods were close and significantly higher than those of the other three methods. The quantum, Otsu, and *K*-means techniques produced similar specificities (around 99.72%) which were 1–2% better than the other two algorithms.

For the multi-objects' dataset including synthetic, noisy, and natural images, the quantum method gave the highest sensitivity over all techniques in comparison with a sensitivity of 97.52%. The sensitivities of the other methods were approximately 4.5–9.2% lower than those of the proposed technique. In terms of specificity, the quantum had a value of 99.61% which is only 0.03% lower than *K*-means and higher than the rest by up to 7.25%. In order to assess the noise performance, Gaussian, salt and pepper, and compression noise are added to some of the synthetic images to create eleven noisy images. Salt and pepper noise was added to the three-objects image in Fig. 3 while other images were modified by the inclusion of Gaussian noise such as the first two images in Fig. 3. The compression noise was added to the single-object image in Fig. 4 at low and medium quality levels, as well as the two-object image in Fig. 4 at low- and high-quality levels. Compression noise was added by compressing then decompressing the image using lossy JPEG scheme at the corresponding quality level. Furthermore, the first image in Fig. 1 was blurred using a multiplicative noise. The test with the noisy images illustrates that the introduced method performed very well in the presence of noise with sensitivity and specificity of 98.30% and 99.30%, respectively. The best noise performing algorithm was the random forests which is slightly higher than the quantum approach by

0.5% and 1.2% for sensitivity and specificity, respectively. Nevertheless, the sensitivities of the quantum method were 3.3–7.2% better than those of the remaining three methods in comparison. Examples of segmented objects in images with multi-objects (Fig. 2) in addition to noisy images (Figs. 3 and 4) are demonstrated for qualitative assessment of the algorithm.

The reported overall average performance measures indicate that the specificities from all methods, except the graph cuts (93.92%), are in close proximity to each other (99.00 to 99.69%) with quantum and *K*-means at the top of the range. However, the superiority of the proposed method over all the other methods under comparison in capturing the target objects for different types of images is evident from the sensitivity results. The quantum technique's average sensitivity is 97.86% compared with the closest value of 94.58% from random forests and the least value of 90.03% from the graph cuts method.

6 Discussion

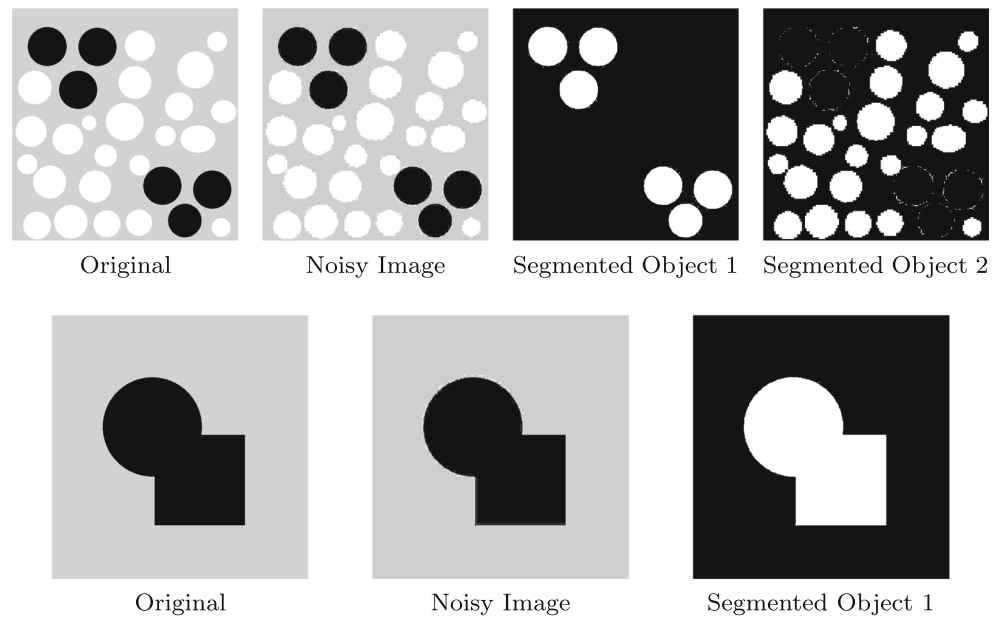
The theory of quantum harmonic oscillators and coherent states provide the bases for the proposed quantum-based image segmentation algorithm. This method relies on treating each pixel in a classical image as QHO initially at the vacuum state. By allowing the system to evolve controlled by features extracted from the image's pixels, the oscillator can reach any of the continuous eigenstates. Principally, this allows for the segmentation of an infinite number of objects. The Hamiltonian parameters are estimated by supervised learning from the image features to lead the evolution to the desired class. The results of applying the system to segment different images indicate that the algorithm can accurately segment multi-objects in many types of images including noisy ones.

The presented method inherits the design flexibility from the original framework (Yousry et al. 2015). So, many design aspects can be adjusted to suit different

Table 1 Comparison of the proposed algorithm with other algorithms on different data subsets, showing the number of images in each subset, and the sensitivity and specificity as percentages

Dataset	Num	Proposed		<i>K</i> -means		Multithreshold		Graph cuts		Random forests	
		Sens	Spec	Sens	Spec	Sens	Spec	Sens	Spec	Sens	Spec
Paper (Yousry et al. 2015)	9	98.23	99.53	92.75	99.76	93.16	99.85	95.53	96.88	98.55	98.75
Multi-objects	12	97.52	99.61	91.12	99.64	89.41	96.57	88.31	92.36	92.78	98.99
Noisy images	11	98.30	99.30	91.71	99.59	91.63	94.98	91.06	99.66	98.84	99.48
All	24	97.86	99.56	91.17	99.69	90.07	97.57	90.03	93.92	94.58	99.00

Fig. 4 Proposed algorithm segmentation of images with compression noise



types of applications such as the form of the Hamiltonian. In this work, the Hamiltonian was selected to lead to a closed-form solution. However for other problems, a more complicated form might be needed which may necessitate obtaining a numerical solution of Schrödinger's equation. The construction of the Hamiltonian was performed using supervised learning which can also be changed to possibly unsupervised learning approach. Additionally, the fidelity metric was adopted in this work as it can be optically implemented as will be discussed later in this section. However, other metrics can be used.

The analysis of the performance summarized in Table 1 shows that the quantum-based method outperforms two of the classical image segmentation techniques in addition to graph cuts and random forests in terms of overall accuracy. All objects in all images were correctly identified by the proposed algorithm. However, *K*-means, graph cuts, and Otsu's method failed to identify objects in some of the images. For example, Otsu's thresholding algorithm missed the second object in the natural image in the third row of Fig. 2 while all three techniques failed to identify the second object in the noisy image in the second row of Fig. 3. Also, it can be seen from Figs. 5 and 6 that

K-means and Otsu's methods tend to undersegment the objects in the natural image of the bird as well as in the gradient noisy image while graph cuts produced greater segmentation inaccuracies. Moreover, Otsu's segmentation of the image in Fig. 7 tends to identify part of object 2 as object 1. The random forests method did not suffer from the aforementioned issues but performed poorly when using complex features in textured images as shown in Fig. 8. The proposed method succeeded in segmenting all objects in the different types of images without missing or undersegmenting any objects. Furthermore, the performance was robust to noise with minimal effect on accuracy. As can be observed from the average sensitivity and specificity for all four types of inserted noise which were only reduced by 1.70% and 0.31% in comparison with the noiseless case. *K*-means is initialized randomly and there is no guarantee that it will produce the correct segmentation. Also, the repeated runs of the algorithm do not yield the same segmentation results. For fair comparison, the tests were done numerously in order to achieve the best segmentation. Also, the features used in the quantum method were used with the *K*-means. The quantum method did not suffer from all the mentioned

Fig. 5 *K*-means, multithreshold, and graph cuts segmentation of a natural image of a bird

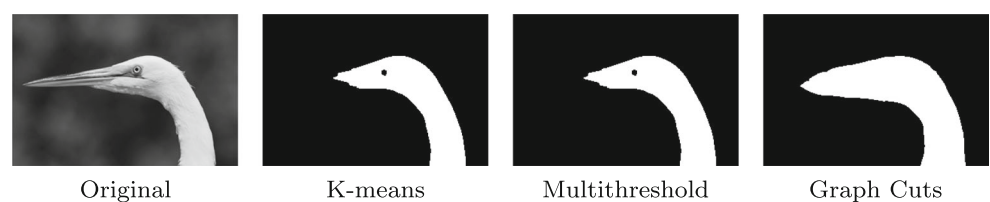
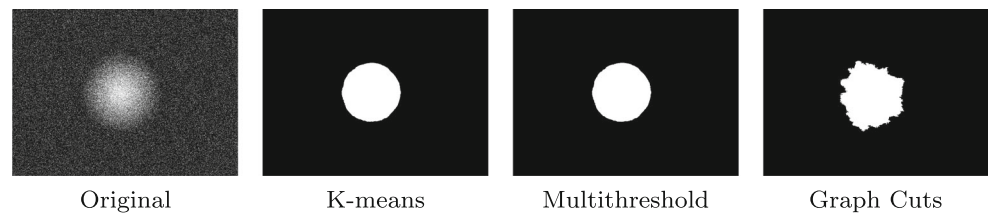


Fig. 6 *K*-means, multithreshold, and graph cuts segmentation of a gradient image with Gaussian noise of variance 0.1



problems. The results do not change by repeated runs of the algorithm. Otsu's method is not suitable for incorporating multiple features (i.e., it works using only one feature). A different approach for image segmentation is the active contours method (Chan and Vese 2001). Although it is very powerful in case of single-object images, it is not suitable for multi-object images. The only way to solve this problem is to customize the location and shape of the initial contour so that it only captures one object and then repeat for other objects. But this requires too much human intervention which is a drawback. In addition to its higher performance, the proposed algorithm does not face these challenges and can automatically do the task with minimal human intervention.

In Sergioli et al. (2016), a framework for pattern recognition has been proposed in which features are mapped to quantum density matrices on the Bloch sphere via a stereographic mapping. This is suitable for a 2D feature vector. For larger number of features, the model can be generalized geometrically by using Bloch spheres of higher dimensions. In this case, higher-dimensional matrices are used. The classification is done by a nearest mean classifier rule based on trace distance. In a broader sense, the work in Sergioli et al. (2016) shares a quantum-based classification approach as the proposed method. Nevertheless, the two methods have many differences. First, the presented work is based on the theory of quantum harmonic oscillators which uses continuous states of infinite-dimensions and is independent of the number of classes as previously discussed. Second, the features are encoded through the Hamiltonian governing the evolution of states. Third, the identification of the final state is performed by first evolving the system then by using the fidelity as a distance metric.

Fourth, the learning is done in a least-square sense for evaluating the Hamiltonian parameters. Finally, despite the ability of the suggested algorithm to be used as a classifier, the focus is on developing a complete for image segmentation technique.

Although the algorithm is designed to work on classical computer, the usage of coherent states opens the path for practical implementation of the algorithm using optical components. Laser beams are described using coherent states and the quantum effects dominate as the number of photons decreases (corresponding to reduction in the beam power). Using a local oscillator (LO) beam, other beams can be generated by splitting the LO beam and then phase-modulating each sub-beam alone. These beams correspond to the fixed reference states $\{|\beta_i\rangle\}$. For each pixel in the image, another beam from the LO is generated and modulated according to the value of the extracted feature at the pixel to produce the final state $|\alpha\rangle$. The beams from the pixel and from the reference states can be combined in a Mach-Zehnder interferometer to estimate the fidelity using the method in Ekert et al. (2002). Then, the state with the highest fidelity is selected to produce the classification of the pixel.

7 Conclusions and future work

This paper proposes an algorithm for segmenting classical images that is formulated from the foundations of quantum mechanics. It can be considered as enhanced extension of the work done in Youssry et al. (2015). In addition to the ability to deploy beneficial aspects from quantum mechanics in image segmentation as the original framework, this

Fig. 7 Multithreshold segmentation of a natural image of marbles

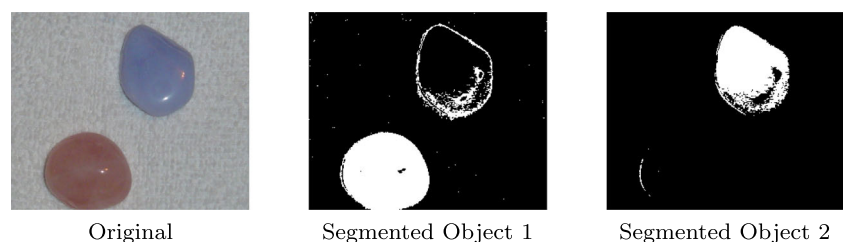
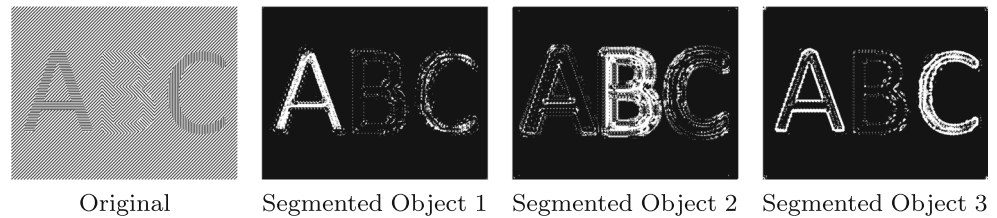


Fig. 8 Random forests segmentation of a textured image



algorithm has a major advantage as it can handle images with multi-objects at no additional computational complexity. This is accomplished by utilizing the theory of quantum harmonic oscillators rather than two-state quantum system. Since the number of coherent states is a continuum, their eigenstates represents a continuous-variable and thus can model any number of objects. The performance of the proposed method demonstrates its high performance in terms of accuracy even when noise is present, while being superior to the original work in Youssry et al. (2015) in terms of complexity. Despite being developed as a classical algorithm, we provided a suggestion on the quantum implementation of the system using the aforementioned optical hardware. The following points should be considered in the future to enhance the system. First, the interaction between neighboring pixels is considered only indirectly in the feature extraction. In order to provide a direct way to incorporate this information and replace the use of complicated image features, the mathematical model of coupled quantum harmonic oscillators could be exploited. Second, the algorithm was presented generally and tested on generic images to show its functionality. It remains to get advantage of its flexibility to efficiently apply it to a particular application. Finally, the system could be physically realized, as described previously, to validate its practically.

Funding information AY is supported by an Australian Government Research Training Program Scholarship. This work is supported by the National Natural Science Foundation of China under Grant No. 61463016, 61763014, National Key R&D Plan under Grant No. SQ2018YFC120002 and “Science and technology innovation action plan” of Shanghai in 2017 under Grant No. 17510740300.

Appendix 1. Proof of Eq. 16

It is required to prove that given the Hamiltonian

$$H = \hbar\omega \left(\hat{a}^\dagger \hat{a} + \frac{1}{2} \right) + \hbar \left(f(t) \hat{a} + f^*(t) \hat{a}^\dagger \right), \quad (28)$$

such that

$$f(t) = f_0 e^{i\omega t}, \quad (29)$$

and the initial state

$$|\psi(0)\rangle = |0\rangle, \quad (30)$$

then the final state obtained after evolution is a coherent state

$$|\alpha\rangle = e^{-\frac{1}{2}|\alpha|^2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle. \quad (31)$$

The Hamiltonian in Eq. 28 can be written in the form

$$H = H_0 + V, \quad (32)$$

where

$$H_0 = \hbar\omega \left(\hat{a}^\dagger \hat{a} + \frac{1}{2} \right) \quad (33)$$

is the free field Hamiltonian which is time independent, while

$$V = \hbar \left(f(t) \hat{a} + f^*(t) \hat{a}^\dagger \right) \quad (34)$$

is the interaction potential which is time dependent. As it is known in the theory of quantum mechanics, there are three pictures (representations) of states and operators:

- The Schrödinger picture in which the states are time dependent, while the operators are time independent.
- The Heisenberg picture in which the states are time independent, while the operators are time dependent
- The Dirac (interaction) picture, where both states and operators are time dependent.

The three representations are equivalent to each other and can be converted easily from one picture to another. The state represented in the interaction picture $|\psi(t)\rangle_I$ is related to Schrödinger’s picture $|\psi(t)\rangle_S$ via the transformation

$$|\psi(t)\rangle_I = U_0^\dagger(t) |\psi(t)\rangle_S, \quad (35)$$

where

$$U_0 = e^{-\frac{i}{\hbar} H_0 t} = e^{-i\omega t \left(\hat{a}^\dagger \hat{a} + \frac{1}{2} \right)} \quad (36)$$

is the free field evolution operator. The interaction potential in the interaction picture $V_I(t)$ is related to the Schrödinger’s picture V_0 via the transformation

$$V_I(t) = U_0^\dagger(t) V_0 U_0(t). \quad (37)$$

By substituting Eqs. 36 in Eq. 37, we get

$$V_I(t) = e^{i\omega t \left(\hat{a}^\dagger \hat{a} + \frac{1}{2} \right)} \hbar \left(f(t) \hat{a} + f^*(t) \hat{a}^\dagger \right) e^{-i\omega t \left(\hat{a}^\dagger \hat{a} + \frac{1}{2} \right)}. \quad (38)$$

This equation is expanded as

$$V_I(t) = \hbar f(t)e^{i\omega t \hat{a}} \hat{a} e^{-i\omega t \hat{a}} + \hbar f^*(t)e^{i\omega t \hat{a}} \hat{a}^\dagger e^{-i\omega t \hat{a}}. \tag{39}$$

The operator part in each of the two terms is in the form $e^{sA} B e^{-sA}$, where $s = i\omega t$, $A = \hat{a}^\dagger \hat{a}$, and $B = \hat{a}$ for the first term and \hat{a}^\dagger for the second term. This allows us to use the Baker-Campbell-Hausdorff (BCH) formula

$$e^{sA} B e^{-sA} = B + [A, B]s + \frac{[A, [A, B]]s^2}{2!} + \frac{[A, [A, [A, B]]]s^3}{3!} + \dots \tag{40}$$

So, we need to calculate these nested commutators. Rewriting the commutator relation in Eq. 5, in the form $\hat{a}\hat{a}^\dagger - \hat{a}^\dagger\hat{a} = 1$, we obtain

$$[\hat{a}^\dagger \hat{a}, \hat{a}] = \hat{a}^\dagger \hat{a} \hat{a} - \hat{a} \hat{a}^\dagger \hat{a} = \hat{a}^\dagger \hat{a} \hat{a} - (1 + \hat{a}^\dagger \hat{a}) \hat{a} = -\hat{a}. \tag{41}$$

$$[\hat{a}^\dagger \hat{a}, [\hat{a}^\dagger \hat{a}, \hat{a}]] = [\hat{a}^\dagger \hat{a}, -\hat{a}] = -[\hat{a}^\dagger \hat{a}, \hat{a}] = \hat{a}. \tag{42}$$

and so on. The nested commutators form an alternating series between \hat{a} and $-\hat{a}$. By substituting the result of Eq. 41 in the BCH formula of Eq. 40, we get

$$e^{i\omega t \hat{a}^\dagger \hat{a}} \hat{a} e^{-i\omega t \hat{a}^\dagger \hat{a}} = \hat{a} - (i\omega t)\hat{a} + \frac{(i\omega t)^2}{2!}\hat{a} - \frac{(i\omega t)^3}{3!}\hat{a} + \dots \tag{43}$$

$$= \hat{a} \left(1 + (-i\omega t) + \frac{(-i\omega t)^2}{2!} + \frac{(-i\omega t)^3}{3!} + \dots \right) \tag{44}$$

$$= \hat{a} e^{-i\omega t}, \tag{45}$$

where the last equation comes from the Taylor expansion of the exponential function. By taking the adjoint we obtain

$$e^{i\omega t \hat{a}^\dagger \hat{a}} \hat{a}^\dagger e^{-i\omega t \hat{a}^\dagger \hat{a}} = \hat{a}^\dagger e^{i\omega t}. \tag{46}$$

Now we can insert the results into Eq. 39 to get

$$V_I(t) = \hbar \left(f(t)\hat{a}e^{-i\omega t} + f^*(t)\hat{a}^\dagger e^{i\omega t} \right). \tag{47}$$

This interaction part does not generally commute in time $[V_I(t_1), V_I(t_2)] \neq 0$, which makes the exact solution difficult to obtain; however, if the external force is resonant with the oscillator so that it takes the form in Eq. 29, and f_0 chosen to be real, we obtain the interaction Hamiltonian in the form

$$V_I(t) = \hbar f_0 (\hat{a} + \hat{a}^\dagger), \tag{48}$$

which clearly commutes in time, so we can obtain a closed-form solution in the interaction picture

$$|\psi(t)\rangle_I = U_I |\psi(0)\rangle_I = e^{-\frac{i}{\hbar} V_I t} |\psi(0)\rangle_I \tag{49}$$

$$= e^{-if_0 t (\hat{a} + \hat{a}^\dagger)} |\psi(0)\rangle_I. \tag{50}$$

We can now return back to Schrödinger’s picture for the initial and final states by using Eq. 35

$$U_0^\dagger(t) |\psi(t)\rangle_S = e^{-if_0 t (\hat{a} + \hat{a}^\dagger)} U_0^\dagger(0) |\psi(0)\rangle_S, \tag{51}$$

which is equivalent to

$$|\psi(t)\rangle_S = U_0(t) e^{-if_0 t (\hat{a} + \hat{a}^\dagger)} U_0^\dagger(0) |\psi(0)\rangle_S. \tag{52}$$

Since $|\psi(0)\rangle_S = |0\rangle$, and $U_0^\dagger(0) = 1$, therefore

$$|\psi(t)\rangle_S = e^{-i\omega t (\hat{a}^\dagger \hat{a} + \frac{1}{2})} e^{-if_0 t (\hat{a} + \hat{a}^\dagger)} |0\rangle. \tag{53}$$

For two operators A and B , if $[A, B] \neq 0$ and $[A, [A, B]] = [B, [A, B]] = 0$, then the following identity holds

$$e^{A+B} = e^A e^B e^{-\frac{1}{2}[A, B]} = e^B e^A e^{\frac{1}{2}[A, B]}. \tag{54}$$

The second exponential clearly satisfies these two conditions with $A = (-if_0 t)\hat{a}$ and $B = (-if_0 t)\hat{a}^\dagger$; consequently the final state becomes

$$|\psi(t)\rangle_S = e^{-i\omega t (\hat{a}^\dagger \hat{a} + \frac{1}{2})} e^{-if_0 t \hat{a}^\dagger} e^{-if_0 t \hat{a}} e^{-\frac{(f_0 t)^2}{2}} |0\rangle. \tag{55}$$

Rearranging the non-operator terms,

$$|\psi(t)\rangle_S = e^{-\frac{i\omega t}{2}} e^{-\frac{(f_0 t)^2}{2}} e^{-i\omega t \hat{a}^\dagger \hat{a}} e^{-if_0 t \hat{a}^\dagger} e^{-if_0 t \hat{a}} |0\rangle. \tag{56}$$

The first term is an overall phase shift that does not change the state, so it can be neglected to get

$$|\psi(t)\rangle_S = e^{-\frac{(f_0 t)^2}{2}} e^{-i\omega t \hat{a}^\dagger \hat{a}} e^{-if_0 t \hat{a}^\dagger} e^{-if_0 t \hat{a}} |0\rangle. \tag{57}$$

The last exponential can be expanded as a Taylor series:

$$|\psi(t)\rangle_S = e^{-\frac{(f_0 t)^2}{2}} e^{-i\omega t \hat{a}^\dagger \hat{a}} e^{-if_0 t \hat{a}^\dagger} \sum_{n=0}^{\infty} \frac{(-if_0 t)^n}{n!} \hat{a}^n |0\rangle. \tag{58}$$

All the terms in the series vanish due to definition of annihilation operator except for the first term ($n = 0$) which is equal to 1 so

$$|\psi(t)\rangle_S = e^{-\frac{(f_0 t)^2}{2}} e^{-i\omega t \hat{a}^\dagger \hat{a}} e^{-if_0 t \hat{a}^\dagger} |0\rangle. \tag{59}$$

Again expanding the last exponential as a Taylor series we get

$$|\psi(t)\rangle_S = e^{-\frac{(f_0 t)^2}{2}} e^{-i\omega t \hat{a}^\dagger \hat{a}} \sum_{n=0}^{\infty} \frac{(-if_0 t)^n}{n!} (\hat{a}^\dagger)^n |0\rangle. \tag{60}$$

By the repeated application of the creation operator on the vacuum state and using Eq. 2,

$$|\psi(t)\rangle_S = e^{-\frac{(f_0 t)^2}{2}} e^{-i\omega t \hat{a}^\dagger \hat{a}} \sum_{n=0}^{\infty} \frac{(-if_0 t)^n}{n!} \sqrt{n!} |n\rangle. \tag{61}$$

Due to linearity, we can operate on the series term by term to get

$$|\psi(t)\rangle_S = e^{-\frac{(f_0 t)^2}{2}} \sum_{n=0}^{\infty} \frac{(-if_0 t)^n}{n!} \sqrt{n!} e^{-i\omega t \hat{n}} |n\rangle. \tag{62}$$

Given that $|n\rangle$ are eigenstates of the number operator $\hat{n} = \hat{a}^\dagger \hat{a}$, with eigenvalues n , the expression becomes

$$|\psi(t)\rangle_S = e^{-\frac{(f_0 t)^2}{2}} \sum_{n=0}^{\infty} \frac{(-if_0 t)^n}{n!} \sqrt{n!} e^{-i\omega t n} |n\rangle. \tag{63}$$

Simplifying this expression, we get

$$|\psi(t)\rangle_S = e^{-\frac{(f_0 t)^2}{2}} \sum_{n=0}^{\infty} \frac{(-if_0 t e^{-i\omega t})^n}{\sqrt{n!}} |n\rangle. \tag{64}$$

Now let

$$\alpha = -i e^{-i\omega t} f_0 t \rightarrow |\alpha|^2 = (f_0 t)^2, \tag{65}$$

then the final state becomes

$$|\psi(t)\rangle_S = e^{-\frac{1}{2}|\alpha|^2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle. \tag{66}$$

This is exactly the definition of the coherent state $|\alpha\rangle$, which is required to prove.

Appendix 2. Proof of invariance of scaling factor in the reference states

Given that the fidelity between a final state $|\alpha\rangle$ and two reference states $|\beta_1\rangle$ and $|\beta_2\rangle$ with equal magnitudes and different phases (i.e lying on the same circle in the complex plane) satisfy that

$$F_1 > F_2. \tag{67}$$

It is required to prove that scaling the reference states do not change that relation. Evaluating the fidelity yields

$$\langle \alpha | \beta_1 \rangle > \langle \alpha | \beta_2 \rangle, \tag{68}$$

or,

$$e^{-\frac{1}{2}|\alpha - \beta_1|^2} > e^{-\frac{1}{2}|\alpha - \beta_2|^2}. \tag{69}$$

This is equivalent to

$$|\alpha - \beta_1|^2 < |\alpha - \beta_2|^2 \tag{70}$$

$$(\alpha - \beta_1)(\alpha - \beta_1)^* < (\alpha - \beta_2)(\alpha - \beta_2)^* \tag{71}$$

$$-\alpha\beta_1^* - \beta_1\alpha^* + |\beta_1|^2 < -\alpha\beta_2^* - \beta_2\alpha^* + |\beta_2|^2. \tag{72}$$

Since $|\beta_1|^2 = |\beta_2|^2$, then

$$-\alpha\beta_1^* - \beta_1\alpha^* < -\alpha\beta_2^* - \beta_2\alpha^* \tag{73}$$

$$\frac{\alpha}{\alpha^*} < \frac{\beta_1 - \beta_2}{\beta_1^* - \beta_2^*}. \tag{74}$$

If scaling is done, such that $\beta_1 \rightarrow \gamma\beta_1$ and $\beta_2 \rightarrow \gamma\beta_2$ for positive γ , then the above equation does not change. Consequently, decisions based on fidelity do not change. This can be generalized to any number of reference states.

References

Adcock MR, Høyer P, Sanders BC (2016) Quantum computation with coherent spin states and the close hadamard problem. *Quantum Inf Process* 15(4):1361–1386

Alpert S, Galun M, Basri R, Brandt A (2007) Image segmentation by probabilistic bottom-up aggregation and cue integration. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*

Borelli L, Aguiar L, Roversi J, Vidiella-Barranco A (2016) Quantum key distribution using continuous-variable non-gaussian states. *Quantum Inf Process* 15(2):893–904

Caraiman S, Manta VI (2015) Image segmentation on a quantum computer. *Quantum Inf Process* 14(5):1693–1715

Casper E, Hung CC, Jung E, Yang M (2012) A quantum-modeled k-means clustering algorithm for multi-band image segmentation. In: *Proceedings of the 2012 ACM research in applied computation symposium*. ACM, pp 158–163

Chan TF, Vese LA (2001) Active contours without edges. *Trans Img Proc* 10(2):266–277

Chapeau-Blondeau F, Belin E (2016) Quantum image coding with a reference-frame-independent scheme. *Quantum Inf Process*: 1–16

Ekert AK, Alves CM, Oi DK, Horodecki M, Horodecki P, Kwak LC (2002) Direct estimations of linear and nonlinear functionals of a quantum state. *Physical Rev Lett* 88(21):217, 901

Gerry C, Knight P (2005) *Introductory quantum optics*. Cambridge University Press, Cambridge

Griffiths DJ (2005) *Introduction to quantum mechanics*. Pearson Education India

Jiang N, Dang Y, Wang J (2016) Quantum image matching. *Quantum Inf Process*: 1–30

Lau HK, Pooser R, Siopsis G, Weedbrook C (2016) Quantum machine learning over infinite dimensions. [arXiv:160306222](https://arxiv.org/abs/160306222)

Li Y, Sun J, Tang CK, Shum HY (2004) Lazy snapping. *ACM Trans Graphics (ToG)* 23(3):303–308

Li P, Xiao H, Li B (2016) Quantum representation and watermark strategy for color images based on the controlled rotation of qubits. *Quantum Inf Process*: 1–26

Lloyd S (1982) Least squares quantization in pcm. *IEEE Trans Inf Theory* 28(2):129–137

Ma H, Huang P, Bao W, Zeng G (2016) Continuous-variable quantum identity authentication based on quantum teleportation. *Quantum Inf Process*: 1–16

Mastriani M (2015a) Quantum boolean image denoising. *Quantum Inf Process* 14(5):1647–1673

Mastriani M (2015b) Quantum image processing? [arXiv:151202942](https://arxiv.org/abs/151202942)

Nielsen MA, Chuang IL (2010) *Quantum computation and quantum information*. Cambridge University Press, Cambridge

Otsu N (1979) A threshold selection method from gray-level histograms. *IEEE Trans Systems, Man, Cybern* 9(1):62–66

Sergio G, Santucci E, Didaci L, Miskczak JA, Giuntini R (2016) Pattern recognition on the quantum bloch sphere. [arXiv:160300173](https://arxiv.org/abs/160300173)

Sommer C, Straehle C, Koethe U, Hamprecht FA (2011) Ilastik: interactive learning and segmentation toolkit. In: *2011 IEEE International symposium on biomedical imaging: from nano to macro*. IEEE, pp 230–233

- Tanaka K, Tsuda K (2008) A quantum-statistical-mechanical extension of gaussian mixture model. In: *Journal of physics: conference series*, vol 95. IOP Publishing, p 012023
- Yan F, Iliyasu AM, Venegas-Andraca SE (2016) A survey of quantum image representations. *Quantum Inf Process* 15(1):1–35
- Yousry A, El-Rafei A, Elramly S (2015) A quantum mechanics-based framework for image processing and its application to image segmentation. *Quantum Inf Process* 14(10):3613–3638
- Yousry A, El-Rafei A, Elramly S (2016) A quantum mechanics-based algorithm for vessel segmentation in retinal images. *Quantum Inf Process* 15(6):2303–2323. <https://doi.org/10.1007/s11128-016-1292-1>
- Zhou R, Hu W, Luo G, Liu X, Fan P (2018a) Quantum realization of the nearest neighbor value interpolation method for ineqr. *Quantum Inf Process* 17(7):166
- Zhou RG, Liu X, Zhu C, Wei L, Zhang X, Ian H (2018b) Similarity analysis between quantum images. *Quantum Inf Process* 17(6):121

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.