QUANTUM MECHANICS IN COMPUTER VISION: AUTOMATIC OBJECT EXTRACTION

Çağlar Aytekin[1], Serkan Kiranyaz[2] and Moncef Gabbouj[2]

[1]*Middle East Technical University,*
Electrical and Electronics Engineering Department
06531, Balgat, Ankara, Turkey

[2]*Tampere University of Technology,*
Signal Processing Department
33101, Tampere, Finland

ABSTRACT

An automatic object extraction method is proposed exploiting the rich mathematical structure of quantum mechanics. First, a novel segmentation method based on the solutions of Schrödinger's equation is proposed. This powerful segmentation method allows us to model complex objects and inherent structures of edge, shape, and texture information along with the grey-level intensity uniformity, all in a single equation. Due to the large amount of segments extracted with the proposed method, the selection of the object segment is performed by maximizing a regularization energy function based on a recently proposed sub-segment analysis indicating the object boundaries. The results of the proposed automatic object extraction method exhibit such a promising accuracy that pushes the frontier in this field to the borders of the input-driven processing only – without the use of "object knowledge" aided by long-term human memory and intelligence.

Index Terms**—** Object extraction, image segmentation, Schrödinger's equation, and quantum mechanics.

1. INTRODUCTION

Object extraction in an image is a challenging problem since segmenting an object needs a combination of various visual cues such as texture information, distribution of the pixel intensities, object boundaries and dissimilarity from other objects/regions. Furthermore, the visual scenery may bear more than one object, or partially occluded objects. Most of the object extraction techniques presented in the literature relies on a supervised approach and/or adopts a class-specific approach, hence requires *a priori* information such as shape templates, manually selected ground-truth information or object part models, see e.g. [1]-[6].A fully unsupervised method which is not only restricted to a specified dataset is required for general applications such as object recognition, content-based image retrieval and object matching. In order to achieve this goal, a robust and powerful segmentation algorithm is needed. However, it is a known fact that the image segmentation is an ill-posed problem [7], since one has to first define what a meaningful region is. There are many segmentation methods, each of which proposes a particular solution of this problem, such as region-based [8]-[12], edge-based [13]-[16] and texture detection [17]-[19] algorithms. All of these algorithms suffer from various problems, such as seed selection, the region merging criteria, the presence of the smooth edges, missing or irregular textures, etc. These deficiencies result in unreliable and often inaccurate object extraction. To remedy this, ideally, a robust segmentation method combining the edge and intensity information along with the ability to detect and use the texture information, dissimilarity from the background and/or to model a certain intensity distribution is thus required as an initial step.

In this study, first a novel segmentation method based on the principals of the quantum mechanics (QM) is proposed. With an extensive analysis, we shall show that it can overcome the aforementioned problems of the classical segmentation algorithms. QM inspired algorithms have been studied and proved useful in various areas [20]-[23]. However, to the best of our knowledge, this is the first attempt to model the image segmentation problem in this domain. The proposed method can generate a massive number of segmentation alternatives and some of them correspond to the object or salient object parts. In order to select appropriate candidates corresponding to the object of interest, we form a regularization function, based on two terms: 1) the object region and, 2) the object boundary both of which are valid assumptions for rigid objects.

The rest of the paper is organized as follows: in Section 2, a brief introduction to QM is given in order to understand the concept behind the proposed method. In Section 3, the proposed segmentation method based on QM is explained and the advantages of the method over typical segmentation algorithms are analyzed. In section 4, the object extraction method proposed in this study is explained in detail and the performance of the proposed method is analyzed in section 5. Finally, Section 6 concludes the study and suggests topics for future research.

2. PRELIMINARIES

In order to get an understanding of the QM, which constitutes the foundation of the proposed method; in this section we shall summarize the four postulates which formalize the rules of the QM concept. The details of the QM can be obtained from [26].

An *observable* is a physical quantity, such as energy, momentum or mass of a particle. To every observable, a corresponding operator exists such that this operation on the observable will yield the values of observables as follows:

$$\hat{A}\boldsymbol{\psi} = a\boldsymbol{\psi} \tag{1}$$

where $\hat{A}$ is an operator and $\boldsymbol{\psi}$ is the eigenfunction (or wave function) of $\hat{A}$ corresponding to the eigenvalue a. One of the important operators in QM is the energy operator, i.e. the Hamiltonian $\widehat{H}$, which is defined as:

$$\widehat{H} = -\frac{\hbar^2}{2m}\nabla^2 + \boldsymbol{V}(\boldsymbol{r}) \tag{2}$$

where ħ is the Planck's constant, $\boldsymbol{r}$ is the 3D position vector and $V(\boldsymbol{r})$ is a potential field on a particle with a mass of m. The eigenvalue corresponding to the Hamiltonian is the energy of the particle E. Combining Eq. 1 and Eq. 2 gives the *time-independent Schrödinger's equation* which helps us to find the possible energy values and corresponding wave functions of a particle of mass m under a potential of $V(\boldsymbol{r})$:

$$-\frac{\hbar^2}{2m}\nabla^2\boldsymbol{\psi} = -\boldsymbol{V}(\boldsymbol{r})\boldsymbol{\psi} + E\boldsymbol{\psi} \tag{3}$$

The set of wave functions ψ are orthogonal to each other and each wave function describes a different state of the particle. The absolute square of a wave function defines the probability of a particle to exist in a specific location for that state, hence the following holds:

$$\int \boldsymbol{\psi}^* \boldsymbol{\psi} d\boldsymbol{r}=1 \tag{4}$$

3. THE PROPOSED SEGMENTATION METHOD

In this work, we consider the absolute square of a wave function $\boldsymbol{\psi}$ as a labeling vector for a segment, since this information is directly related to the probability of occurrence of a particle in space where 0 and 1 correspond to background and foreground respectively. Note that each pixel in an image is considered as a potential field, $V(x,y)$. Hence, each wave- (eigen-) function, $\psi(x,y)$, represents a meaningful region (a certain set of pixels), which corresponds to an imaginary quantum particle with a mass *m* being at a certain energy level with a certain probability of occurrence, basically satisfying Eq. (3). Therefore, solving the Schrödinger's equation will provide us with wave-functions representing meaningful segments (in a sense of having the same quantum-mechanical properties) in the image. In order to understand why these segments are meaningful (i.e., that has the potential to represent the object region); we now perform an extensive mathematical analysis. First let us multiply both sides of Eq. 3 with $\boldsymbol{\psi}^T$:

$$\frac{-h^2}{2m}\boldsymbol{\psi}^T \nabla^2 \boldsymbol{\psi} = \boldsymbol{\psi}^T (E - \boldsymbol{V})\boldsymbol{\psi} \tag{5}$$

Now, rewrite the Laplacian by the finite difference approach as in [27]:

$$\nabla^2 \boldsymbol{\psi}|_i = \left(\sum_{n \in N_i} \boldsymbol{\psi}(n) \right) - |N_i| * \boldsymbol{\psi}(i) \tag{6}$$

where N_i is the set of neighbours with the index i, and $|N_i|$ is the cardinality of, i.e the number of elements in N_i. Combining Eq. 5 and Eq. 6, we obtain:

$$\frac{h^2}{2m}\sum_{i=1}^{M}\left(\boldsymbol{\psi}(i)^2 * |N_i| - \sum_{n \in N_i} \boldsymbol{\psi}(i) * \boldsymbol{\psi}(n)\right) = \sum_{i=1}^{M} \boldsymbol{\psi}(i)^2 * \left(E - \boldsymbol{V}(i)\right) \tag{7}$$

where *M* is the number of elements in $\boldsymbol{\psi}$. The left-hand side of this equation is a measure of how similar the labels in a neighborhood are, i.e. a measure of spatial coherence. Now consider that we have given the potential $\boldsymbol{V}$ as the gray-level pixel intensities in a picture as input. Then, the right-hand side is a measure of how close the pixel values in a segment are to a constant value E, hence to each other. Hence, solution of the Schrödinger's equation provides segments whose spatial coherence is related with their pixel similarity. Note that m (the mass of the particle) is a parameter of regularization between the spatial coherence and intensity similarity within a segment. A large mass selection will favor the pixel similarity against spatial coherence.

Therefore, in the proposed segmentation algorithm, we put gray-level pixel intensities in the potential function $\boldsymbol{V}$. Furthermore, by approximating the Laplacian as in Eq. 6, the wave functions can be numerically calculated by solving the eigenvectors of Hamiltonian operator in matrix form which is defined as follows:

$$H(i,j) = \begin{cases} V(i) + |N_i| * \dfrac{h^2}{2m}, & i = j \\ -\dfrac{h^2}{2m}, & j \epsilon N_i \\ 0, & e.w \end{cases} \tag{8}$$

Since the size of *H* matrix is equal to $|V| \times |V|$, where $|V|$ is the cardinality of V, and in this case V is the gray level pixel intensities of an image, performing a decomposition to calculate the eigenvectors of this matrix will be very costly. Instead we calculate the minimum and maximum eigenvalues with the power iteration method [28], and then find the eigenvectors with eigenvalues closest to a number of values regularly selected between the minimum and maximum eigenvalues, with the inverse power iteration method [29]. In this work, we empirically select this number as 300. Note that the mass of the quantum particle is the only parameter in the 2D time-independent Schrodinger's equation. The effect of the mass in this application is crucial in the sense that it has a close relation to the *scale* in the proposed segmentation algorithm, which will not be shown herein due to the space limitations. In order to adopt a multi-scale approach, we perform this for six different mass selections corresponding to six $\hbar^2/2m$ values uniformly assigned in an empirical range of [0.2, 1.2], ending up with 1800 eigenvector in total. Then, we threshold the absolute square of the eigenvectors $\boldsymbol{\psi}(i)^2$ to obtain the segments, where the threshold is empirically set to 10^{-6}. Before going into details of the proposed object extraction method, we show how this segmentation method is robust to noise and how can it extract textures by analyzing a unique property of quantum mechanics: *Tunneling*.

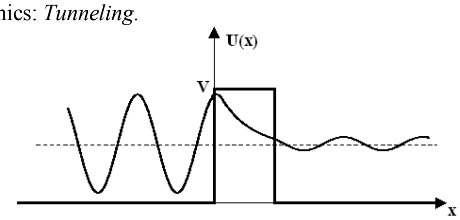

Figure 1: Tunneling of a particle through a 1-D barrier of potential V in a potential distribution U(x).

In quantum mechanics, the tunneling effect allows a particle to be present in a potential *higher* than its energy. The main idea is that the particle can *tunnel* through a potential barrier higher than its energy and reappear on the other side of the barrier (from left to right) with a small probability of occurrence (see Fig. 1).The reverse of this phenomenon also holds. The wave functions with higher eigenvalues than the potential of the barrier are focused on the barrier, which is logical if one considers the orthogonality constraint on wave functions. These wave functions may also occur in regions with low potential then their energy, but slowly vanish similar to the above case. This unique property provides a significant robustness against variations and disturbances such as intensity and color changes, shadows, and noise on objects (see Fig. 2). Furthermore, with tunneling through repeating parts of textures, this algorithm can also extract textures as segments. However, due to tunneling effect, object boundaries may not be very accurate. Due to space limitations, an extensive analysis of tunneling effect is omitted in this paper.

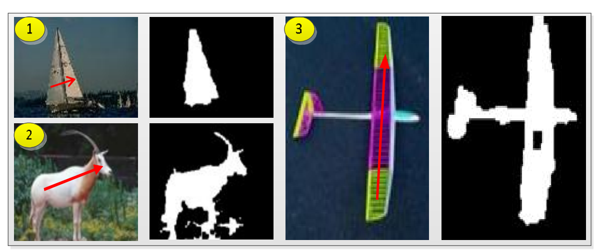

Figure 2: Object segmentation examples with tunneling effect. Red arrows indicate the regions where tunneling occurs in-between.

4. PROPOSED OBJECT EXTRACTION METHOD

There are a massive number of wave functions each of which indicate a potential segment, and are extracted from the Schrödinger's equation. Some of these correspond to the object(s) in the image, whereas the others correspond to other meaningful regions such as the (part of) background or specific object parts with a certain structure. Hence, one should consider a number of constraints to narrow down the selection to one or few proper segments for a specific application, which is object extraction in this work. In order to accurately select the one or few segment(s) corresponding to the object, we rely on the basic rigid object assumption, *objects have boundaries* and output of a multi-scale edge detection (referred as sub-segments) algorithm proposed in [24] is exploited and segments with largest encapsulation of edges are searched. However, an area constraint should also be applied, in order to avoid the bias towards large regions. To accomplish this, we form the following regularization function and search for ψ^* that maximizes:

$$\psi^* = \arg\max_{\psi}(w \times I(\psi, SS) - (1-w) \times A) \qquad (9)$$

where, $I(\psi, SS)$ is the ratio of the number of pixels of the sub-segments (edge parts), *SS*, encapsulated by ψ to their total number, and A is the area of the segment represented by ψ unit-normalized by the total image area. The parameter w is the weight in the interval [0, 1] which controls the trade-off between the sub-segment encapsulation, $I(\psi, SS)$ and the area of the wave function. By increasing w, one will favor the encapsulation of the sub-segments and vice versa for the minimization of the area.

The optimal segment(s) that maximizes the objective function in Eq. 9 with two weight settings: w=0.5 and w=0.7, are then selected from 1800 segments extracted by proposed segmentation method. To demonstrate the accuracy of the proposed object extraction method the best 5 segments with the highest objective function values for each weight setting are then taken into account. As a result, we aim to show that the object(s) can be accurately extracted within a total number of 2 (weights) x 5 (segments) = 10 segments.

5. EXPERIMENTAL RESULTS AND DISCUSSIONS

The experiments are conducted over an extensive set of images from several datasets and personal image collections. Due to space limitations, we can only put few results in this paper and the readers are referred to [25] for an extensive collection of images and object extraction results.

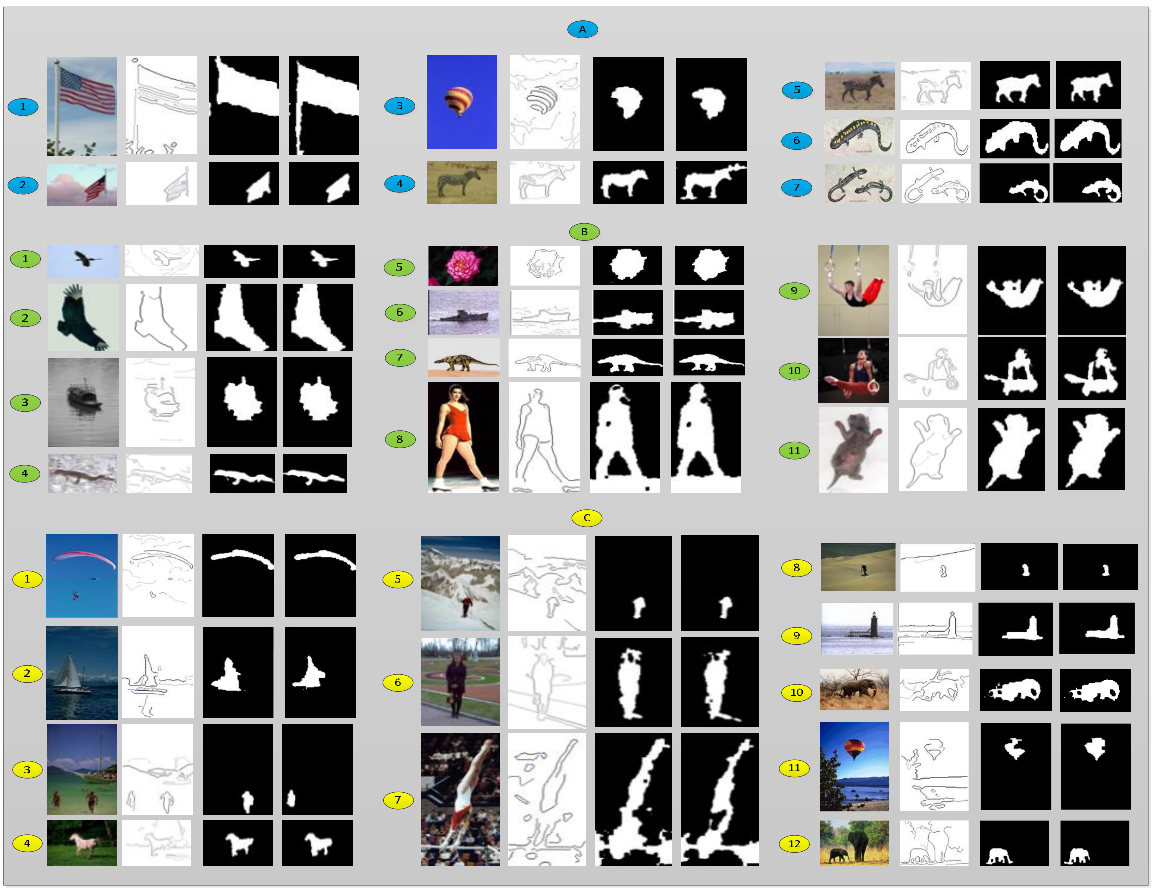

Figure 3: *(column-wise)* Original images, 20 most relevant sub-segments, and two object extraction results. *(row-wise)* (A) 7 images with object(s) containing strong texture information or significant intensity/color variations. (B) 11 images with complex object(s) and a uniform background. (C) 12 images with complicated sub-segments.

First, we analyze the performance of our algorithm, i.e. ability to extract the object in the image within the best 10 segments. Figure 3 contains images having objects with (A) texture, (B) complex structure, color varieties, (C) complicated edge information. As explained above, due to tunneling effect, textures and complex objects with color variations, with the presence of noise and/or shadows on them can be extracted with the cost of a certain inaccuracy on the object boundaries that are often negligible. One can see such effects nearly in all the examples in Figure 3. Furthermore, by the segment selection method explained in Section 4, we achieve the object extraction manually selected within the best 10 segments.

In order to compare our method with existing algorithms, one should select the competitors as automatic, unsupervised, and class independent object extraction algorithms or some well-known segmentation methods for a fair comparison. For that, we compare our method against the object extraction algorithm in

[24] along with the GBIS [3] and KMCC [11] segmentation methods. As observed from Figure 4, the competing methods can achieve a reasonable segmentation/object extraction performance only under certain circumstances, i.e. when some certain assumptions and conditions are met for the image/object. Therefore, it is fairly expected that an image, which is straightforward to segment by one method can be infeasible with another e.g. note examples in (1-3, 5, 7, 9, 10, 13-16). Furthermore, in some of the examples such as (4, 6, 11), none of the competitors are able to extract the object accurately (over/under segmentation occurs) whereas the proposed method achieves an elegant performance in all cases.

Finally we perform an extensive evaluation over a total number of 644 images collected from benchmark VOC datasets [30] all of which usually include images with a salient object and a cluttered background. For each image, we evaluate precision (P) and recall (R) performance measures of the segment maximizing the F1-score [31] among the 10 best segments automatically selected with our algorithm. We also repeat the same process for all 1800 segments, so that we can measure the maximum performance one can achieve with the proposed segmentation algorithm. The first and second order statistics (mean, μ and variance, σ^2) for the P and R performance measures are listed in Table 1. It is evident from the average Recall rate that the proposed segmentation technique can extract the object segment in a high accuracy (>75%) even from images with complex scenery/background. However, comparably lower μP indicates that the extracted object segment can have a certain amount of non-object parts. Although, tunneling is a reason of low precision, the cues used for object segment selection (edge encapsulation and area restriction) are also not sufficient and there is an imminent need for some other "objectness" cues in order to select the right segment(s) extracted by the proposed method as object(s) for better performance.

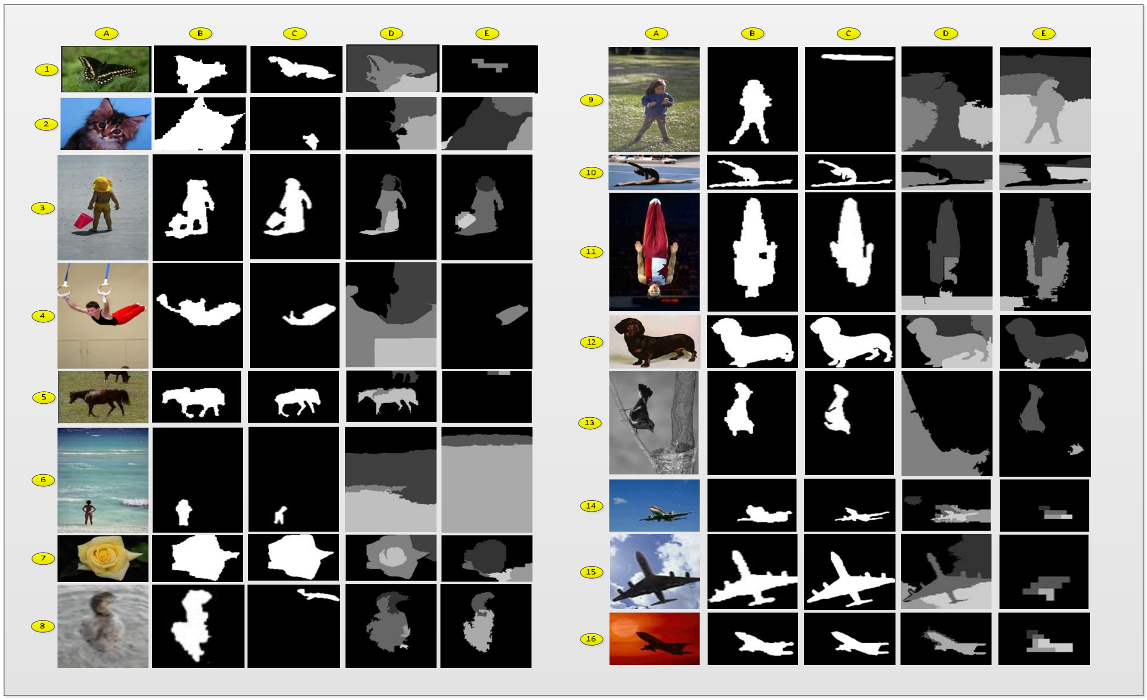

Figure 4: (A) Original images and object extraction/segmentation results from: (B) the proposed method (C) the method in [24], (D) GBIS [3] (E) KMCC [11].

Table 1: P and R statistics (μ and σ^2) for the segment maximizing the F1 score among the best 10 segments (first row), and among all segments (second row).

Segments	μP	μR	σ^2P	σ^2R
Best 10	0.4214	0.7919	0.0582	0.0510
All	0.5847	0.7540	0.0527	0.0388

6. CONCLUSIONS

In this paper, we have demonstrated that the segmentation problem can be efficiently and accurately modeled by the quantum mechanics thanks to its rich mathematical structure and unique properties. Particularly, it provides the basis and resources needed for overcoming the severe problems and limitations of classical segmentation methods. In fact the proposed method is capable of providing a massive number of segmentation alternatives, among which the proper segmentation scheme (that is the model or criteria that one seeks to perform required segmentation) can be selected according to the problem at hand. In this study, the focus is particularly drawn on automatic object extraction, which is based on the proposed segmentation method. In order to select the true segment(s) that encapsulate(s) the object(s) of interest, we then adopted the sub-segment analysis proposed in a recent work, which provides the reliable information and the visual cue needed to select the proper segment(s) among a massive number of candidates. An extended set of experiments approve the accuracy and superiority of the proposed object extraction method especially on complex objects and scenes where recent state-of-the-art segmentation or automatic, unsupervised and class independent object extraction methods fail. This is due to the fact that the proposed method achieves an "all-in-one" solution and thus negates the need for designing/tuning the method with respect to the object/image properties.

The proposed method may fail to extract the entire object if the sub-segments are too noisy or unreliable, indicating a significant complexity for both the object and the background. Such cases sometimes constitute the limit of the automatic and unsupervised methods.

REFERENCES

[1] T. Cour and J. Shi, "Recognizing Objects by Piecing Together the Segmentation Puzzle," Proc. IEEE Conf. Computer Vision and Pattern Recognition, pp. 1-8, 2007.

[2] Z. Tu, X. Chen, A.L. Yuille, and S.C. Zhu, "Image Parsing: Unifying Segmentation, Detection, and Recognition," Proc. Toward Category-Level Object Recognition, pp. 545-576, 2006.

[3] H.E.A.E. Munim and A.A. Farag, "A Shape-Based Segmentation Approach: An Improved Technique Using Level Sets," Proc. 10th IEEE Int'l Conf. Computer Vision, pp. 930-935, 2005.

[4] L. Zhao and L.S. Davis, "Closely Coupled Object Detection and Segmentation," Proc. 10th IEEE Int'l Conf. Computer Vision, pp. 454- 461, 2005.

[5] T. Liu, Z. Yuan, J. Sun, J. Wang, N. Zheng, X. Tang and H-Y. Shum, "Learning to Detect a Salient Object", IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 33, no. 2, pp. 353-367, Feb. 2011.

[6] G. Liu, Z. Lin, Y. Yu and X. Tang, "Unsupervised Object Segmentation with a Hybrid Graph Model", IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 32, no. 5, pp. 910-924, May. 2010.

[7] J.L. Marroquin, S. Mitter, T. Poggio, "Probalistic solution of three ill-posed problems in computational vision," J. Am. Stat. Assoc. 82 (397) pp. 76–89, 1987.

[8] S. A. Hijjatoleslami and J. Kitter, "Region growing: A new approach", IEEE Trans. On Image Proc., vol. 7, pp. 1079-1084, 1998.

[9] I. Kompatsiaris and M. G. Strintzis, "Content-based Representation of Colour Image Sequences", IEEE Int. Conf. on Acoustics, Speech and Signal Proc. 2001 (ICASSP), Salt Lake City, USA, May 2001.

[10] E. Tuncel and L. Onural, "Utilization of the recursive shortest spanning tree algorithm for video-object segmentation by 2-D affine motion modeling", IEEE Trans. On Circuits and Systems for Video Tech., vol. 10, no. 5, pp. 776-781, Aug. 2000.

[11] V. Mezaris, I. Kompatsiaris and M. G. Strintzis, "Still Image Segmentation Tools for Object-based Multimedia Applications", Int. Journal of Pattern Recognition and Artificial Intelligence, vol. 18, no. 4, pp. 701-725, June 2004.

[12] Y. Deng, and B. S. Manjunath, "Unsupervised segmentation of color-texture regions in images and video", IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI '01), Vol. 23, No. 8, pp. 800-810, Aug. 2001.

[13] J. P. Fan, D. K. Y. Yau, A. K. Elmagarmid, W. G. Aref, "Automatic image segmentation by integrating color-edge extraction and seeded region growing", IEEE Trans. on Image Processing, Vol. 10, No. 10, pages 1454-1466, Oct. 2001.

[14] P. F. Felzenszwalb and D. P. Huttenlocher, "Efficient Graph-Based Image Segmentation", Int. Journal of Computer Vision, Vol. 59, No. 2, Sep. 2004.

[15] Q. Gao, D. Qing, S. Lu, "Fuzzy Classification of Generic Edge Features", ICPR 2000, vol. 3, pp. 668-671, 2000.

[16] S. Mahamud, L. R. Williams, K. K. Thornber, and K. Xu, "Segmentation of multiple salient closed contours from real images", IEEE Trans. on Pattern Analysis and Machine Intelligence, 25(4):433-444, 2003.

[17] J. Portilla and E. P. Simoncelli, "A Parametric Texture Model Based on Joint Statistics of Complex Wavelet Coefficients", International Journal of Computer Vision 40(1) , pp. 49-71, 2000.

[18] C. Palm and T. M. Lehmann, "Classification of Color Textures by Gabor Filtering", Machine Graphics and Vision, vol.11, no. 2/3, pp. 195-219 2002.

[19] M. Varma and A. Zisserman, "Texture Classification: Are Filter Banks Necessary?", Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, vol 2., pp. 601-698,2003.

[20] C. C. Tseng and T. M. Hwang, "Quantum Digital Image Processing Algorithms", 16th IPPR Conference on Computer Vision, Graphics and Image Processing, 2003.

[21] Y. C. Eldar, "Quantum Signal Processing", Ph.D. dissertation, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, 2001.

[22] D. Horn and A. Gottlieb, "Algorithm for data clustering in pattern recognition problems based on quantum mechanics", Phys. Rev. Lett. Vol 88, No 1, pp. 1-4, 2002.

[23] N. Nasios, A. G. Bors, "Kernel-based classification using quantum mechanics," Journal of Pattern Recognition, Elsevier, vol. 40, no. 3, pp. 875-889, Mar. 2007.

[24] S. Kiranyaz, M. Ferreira, M. Gabbouj, "Automatic Object Extraction over Multi-Scale Edge Field for Multimedia Retrieval", in IEEE Transactions on Image Processing, pp. 3759-3772, December, 2006.

[25] Object Extraction Results, [Online]: http://muvis.cs.tut.fi/quantum_suppl.htm

[26] R.L. Liboff, "The Postulates of Quantum Mechanics. Operators, Eigenfunctions, and Eigenvalues," in *Introductory Quantum Mechanics,* 4th ed. San Francisco: Addison Wesley, 2003, ch 3, pp. 68-84.

[27] K.W. Morton and D.F. Mayers, *Numerical Solution of Partial Differential Equations, An Introduction*, 2nd ed. Cambridge, UK: Cambridge University Press, 2005.

[28] R. von Mises and H. Pollaczek-Geiringer, "Praktische Verfahren der Gleichungsauflösung", ZAMM - Zeitschrift für Angewandte Mathematik und Mechanik, vol 9, pp. 152-164, 1929.

[29] Ernst Pohlhausen, "Berechnung der Eigenschwingungen statisch-bestimmter Fachwerke", ZAMM - Zeitschrift für Angewandte Mathematik und Mechanik vol. 1, pp. 28-42, 1921.

[30] http://pascallin.ecs.soton.ac.uk/challenges/VOC/

[31] C.J van Rijsbergen, *Information Retrieval,* 2nd ed. London, UK: Butterworths, 1979.