Quantum Motion Segmentation

Federica Arrigoni [1] Willi Menapace [2] Marcel Seelbach Benkner [3] Elisa Ricci [2,4] Vladislav Golyanik [5]

[1] Politecnico di Milano [2] University of Trento [3] University of Siegen [4] Bruno Kessler Foundation [5] MPI for Informatics

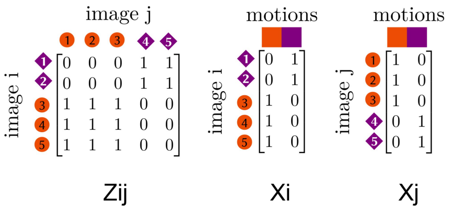
Overview and Contributions

- Task of motion segmentation (MS) is to *classify points in multiple images into different motions*

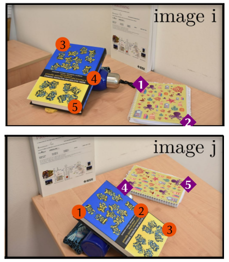
- Applications of motion segmentation: multi-body SfM [45], motion estimation [51]
- MS is an NP-Hard problem
- **Contributions** of this paper:
    - QuMoSeg: A new MS approach employing AQC
    - A new Q-MSEG dataset for MS
- QuMoSeg reaches SotA accuracy on a wide range of problems (QA and SA versions)

Motion Segmentation

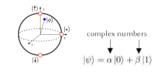
- **Goal**: identify independent motions in multiple images
- **Assumption:** two-frame matches as input, which are used to recover two-frame segmentations Zij via multi-model fitting [40]
- **Task**: recover unknown absolute segmentations X1, …, Xn such that they are consistent with the known two-frame segmentations Zij, namely: $Z_{ij} = X_i X_j^\mathsf{T}$



$$Z_{ij} = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \end{bmatrix} \quad X_i = \begin{bmatrix} 0 & 1 \\ 0 & 1 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \end{bmatrix} \quad X_j = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}$$



Adiabatic Quantum Computing



AQC leverages quantum fluctuations

Main parameter: transverse magnetic field

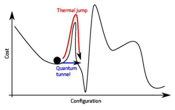


Image: Steger *et al.*, SPIE, 2015.

Initial state:

$$|\psi(t=0)\rangle = \bigotimes_{i=1}^{n} \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$$

Transition (simplified):

$$H(t) = \left(1 - \frac{t}{\tau}\right) H_I + \frac{t}{\tau} H_P$$

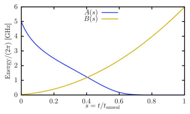


Image: Willsch *et al.*, Computer Physics Communications, 2022.

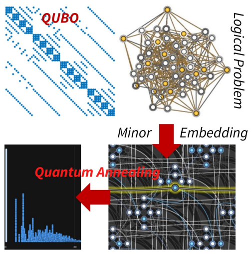


Image: [10]

- It optimizes objectives over **binary variables** and obtains globally-optimal or low-energy solutions with high probabilities, leveraging quantum mechanics.
- The objective must be expressed as a **QUBO** (quadratic unconstrained binary optimization)

$$\min_{\mathbf{y} \in \mathcal{B}^k} \mathbf{y}^\mathsf{T} Q \mathbf{y} + \mathbf{s}^\mathsf{T} \mathbf{y}$$

- Linear constraints are treated as soft ones

$$\min_{\mathbf{y} \in \mathcal{B}^k} \mathbf{y}^\mathsf{T} Q \mathbf{y} + \mathbf{s}^\mathsf{T} \mathbf{y} + \sum_i \lambda_i ||A_i \mathbf{y} - \mathbf{b_i}||^2$$

Our Approach

Motion segmentation via synchronization (**combinatorial objective**)

$$\min_{X_1,\dots X_n} \sum_{(i,j)\in\mathcal{E}} ||Z_{ij} - X_i X_j^\mathsf{T}||_F^2, \qquad X = \begin{bmatrix} X_1 \\ X_2 \\ \dots \\ X_n \end{bmatrix}, \quad Z = \begin{bmatrix} 0 & Z_{12} & \dots & Z_{1n} \\ Z_{21} & 0 & \dots & Z_{2n} \\ \dots & & & \dots \\ Z_{n1} & Z_{n2} & \dots & 0 \end{bmatrix}$$

$$\text{s.t. } \operatorname{vec}(X_i) \in \mathcal{B}^{p_i}, \quad X_i \mathbf{1}_d = \mathbf{1}_{pi} \quad \forall i = 1, \dots, n$$

binary variables

Each point belongs to one motion only

QuMoSeg-v1 (no additional assumptions, dense matrix)

$$\max_{X_1,\dots X_n} \sum_{(i,j)\in\mathcal{E}} \operatorname{trace}(X_i^\mathsf{T}(2Z_{ij} - \mathbf{1}_{p_i \times p_j})X_j)$$
$$s.t. \ \operatorname{vec}(X_i) \in \mathcal{B}^{p_i}, \quad X_i \mathbf{1}_d = \mathbf{1}_{pi} \quad \forall i = 1, \dots, n$$

$\Leftrightarrow$

$$\max_X \operatorname{vec}(X)^\mathsf{T}(I_{d\times d} \otimes (2Z - \mathbf{1}_{p\times p})) \operatorname{vec}(X)$$
$$s.t. \ \operatorname{vec}(X) \in \mathcal{B}^{dp}, \quad X\mathbf{1}_d = \mathbf{1}_p.$$

we know the number of points per motion in every image

QuMoSeg-v2 (additional assumptions, sparse matrix)

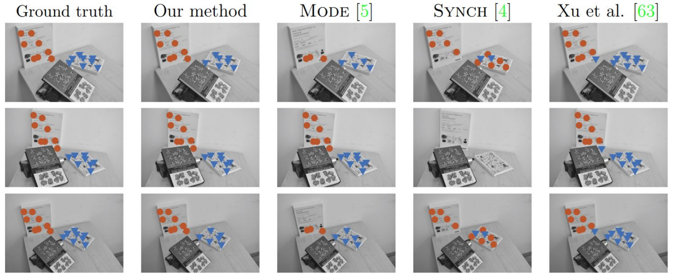
$$\max_{X_1,\dots X_n} \sum_{(i,j)\in\mathcal{E}} \operatorname{trace}(X_i^\mathsf{T} Z_{ij} X_j),$$
$$s.t. \ \operatorname{vec}(X_i) \in \mathcal{B}^{p_i}, \quad X_i \mathbf{1}_d = \mathbf{1}_{pi}, \quad \mathbf{1}_{p_i}^\mathsf{T} X_i = \mathbf{m}_i^\mathsf{T} \quad \forall i = 1, \dots, n.$$

$\Leftrightarrow$

$$\max_X \operatorname{vec}(X)^\mathsf{T}(I_{d\times d} \otimes Z) \operatorname{vec}(X),$$
$$s.t. \ \operatorname{vec}(X) \in \mathcal{B}^{dp}, \quad X\mathbf{1}_d = \mathbf{1}_p, \quad KX = M$$

additional constraints

Experimental Results



$\mu =$	0.97	0.93	0.93	0.89

Qualitative results on Q-MSEG dataset with 96 qubits. Each color (symbol) represents a distinct planar motion.

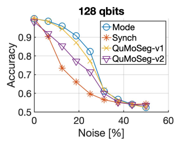
Q-MSEG

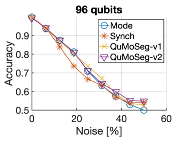
# Qubits:	96	102	120	126	128	136	160	168	180	190	200	216	220	243
Xu et al. [63]	0.89	0.89	0.94	0.75	0.96	0.97	0.86	0.86	0.97	0.88	0.96	0.77	0.83	0.74
Mode [5]	0.93	0.93	0.96	0.93	**0.97**	0.97	**0.98**	0.99	**0.98**	**0.99**	**0.99**	0.93	**1**	**0.94**
Synch [4]	0.93	0.94	0.95	0.95	0.84	0.92	0.97	**1**	0.89	0.95	0.90	**0.94**	0.99	0.92
QuMoSeg-v1	0.97	0.97	0.97	0.96	0.95	**0.98**	**0.98**	0.99	**0.98**	**0.99**	**0.99**	0.64	–	–
QuMoSeg-v2	0.96	0.97	0.95	0.94	0.89	0.89	0.88	0.85	0.74	0.75	0.79	0.59	0.75	0.58
QuMoSeg-v1, SA	0.97	0.97	0.97	0.96	0.95	**0.98**	**0.98**	**1**	**0.98**	**0.99**	**0.99**	0.68	0.98	0.72
QuMoSeg-v2, SA	**0.98**	**0.99**	**0.99**	**1**	0.96	**0.98**	**0.98**	**1**	0.94	0.97	**0.99**	0.80	**1**	0.59

Hopkins

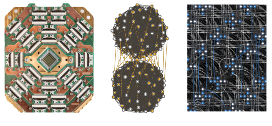
# Qubits:	120	126	132	138	144	156	162	168	174	180	186	192	198	204	210	216	222	228	234	240
Xu et al. [63]	0.80	0.78	0.81	0.79	0.83	0.81	0.84	0.81	0.85	0.89	0.88	0.94	0.96	0.96	0.97	**1**	0.98	**1**	**0.99**	**1**
Mode [5]	0.89	0.91	0.90	0.93	0.92	0.94	0.95	0.95	0.96	0.95	0.97	0.98	0.98	0.98	0.98	0.99	0.98	0.98	**0.99**	0.99
Synch [4]	0.87	0.93	0.95	0.96	**0.99**	0.96	**0.99**	**0.99**	0.96	0.99	**1**	**1**	0.99	**1**	**1**	0.70	0.97	**1**	**0.99**	0.65
QuMoSeg-v1	0.92	0.89	0.93	0.93	0.93	0.93	0.95	0.94	0.96	0.95	0.96	0.97	0.96	-	-	-	-	-	-	-
QuMoSeg-v2	0.91	0.92	0.91	0.92	0.94	0.89	0.91	0.89	0.90	0.88	0.88	0.89	0.88	0.89	0.88	-	-	-	-	-
QuMoSeg-v1, SA	0.93	0.90	0.92	0.94	0.93	0.94	0.95	0.96	0.96	0.96	0.98	0.98	0.98	0.98	0.99	0.99	0.98	0.98	**0.99**	0.99
QuMoSeg-v2, SA	**0.96**	**0.97**	**0.98**	**0.98**	**0.99**	**0.99**	**0.99**	0.97	**0.99**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**0.99**	**1**	**0.99**	**1**

Increasing Noise Levels





Minor Embeddings

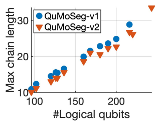


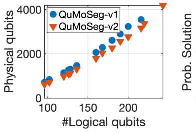
D-Wave Adv4.1

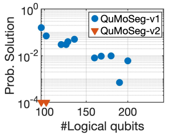
Logical Problem (96 qubits)

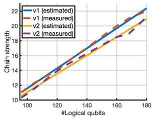
Minor Embedding (803 qubits; excerpt)

Execution on D-Wave Quantum Annealer









References

[4] Arrigoni and Pajdla. Motion segmentation via synchronization. ICCVW, 2019.
[5] Arrigoni and Pajdla. Robust motion segmentation from pairwise matches. ICCV, 2019.
[10] Birdal and Golyanik *et al.* Quantum permutation synchronization. CVPR, 2021.
[40] Magri and Fusiello. Multiple structure recovery via robust preference analysis. Image and Vision Computing (2017)
[45] Ozden *et al.* Multibody structure-from-motion in practice. TPAMI, 2010.
[51] Sabzevari and Scaramuzza. Multi-body motion estimation from monocular vehicle-mounted cameras. T-RO, 2016.
[60] Tron and Vidal. A benchmark for the comparison of 3-d motion segmentation algorithms. CVPR, 2007.
[63] Xu *et al.* 3d rigid motion segmentation with mixed and unknown number of models. TPAMI, 2019.