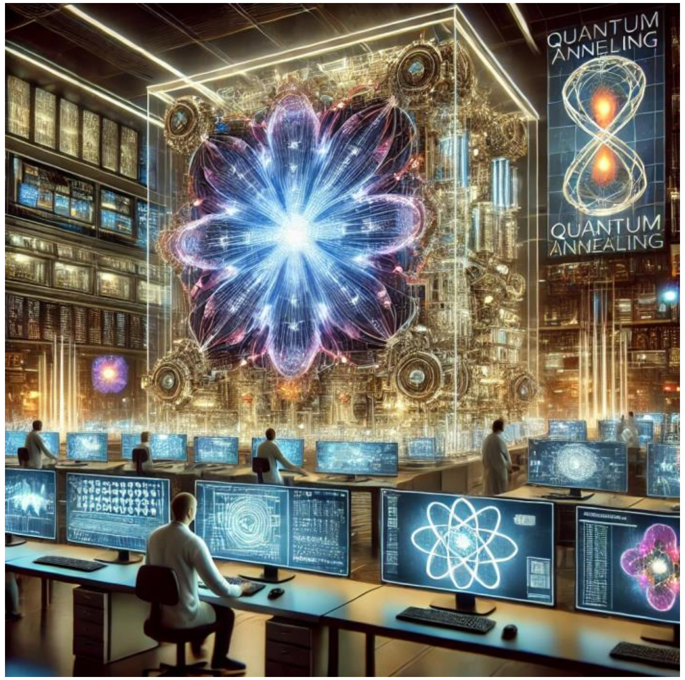
Quantum Annealing in Computer Vision

Shahrokh Heidari

Supervisors:

Patrice Delmas, Michael Dinneen, Cristian Calude

Table of Contents



Credit: DALL·E text-to-image model
(prompt: quantum annealing in computer vision)

Motivations

Deep learning:

Cons

- Increased model size
- Hardware requirements
- No optimality proof
- Learning phase requirements

Pros

- High accuracy
- Feature learning
- Scalability
- Versatility

Quantum computing :

Cons

- Technical Challenges
- Scalability Issues
- Quantum Error
- Accessibility
- Interdisciplinary Knowledge

Pros

- Optimality
- Exponential Speedup
- Limited space requirements
- Enhanced Optimization

Motivations

A framework to standardize and adapt computer vision problems for quantum computing

- Such a framework facilitates the future integration and application of quantum algorithms in the field of computer vision.
- Selecting a computer vision problem that is challenging to solve and still extensively researched.
  - **Stereo Matching**
- Quantum computing: Gate-based or Annealing?
  - Gate-based: uses quantum gates to perform operations on qubits, similar to classical logic gates but with quantum properties.
  - Annealing: a method for finding/estimating the global minimum of an objective function
- Selecting a quantum computation approach that aligns well with the nature of computer vision and machine learning problems → **optimization**

Quantum Annealing

Stereo Vision

- **Stereo-vision system**
  - A system to mimic human vision to interpret the 3D world.
  - Each camera in a stereo-vision system records an image that, while fundamentally similar, features a certain degree of displacement (disparity)
- **Stereo Matching problem**
  - The most complex part of the system
- **Motivations**
  - Stereo Matching is still considered an open problem
  - Incorporating QA with Stereo Matching
  - Proof of concept for when quantum computers can handle large images

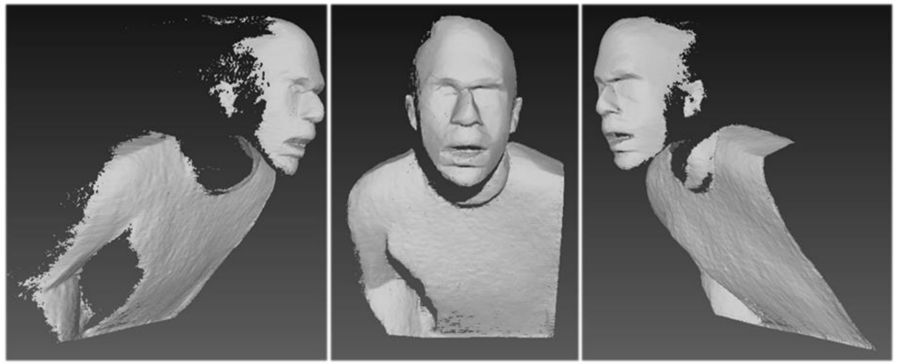


Fig. 1: 3D reconstruction example from a stereo-vision system

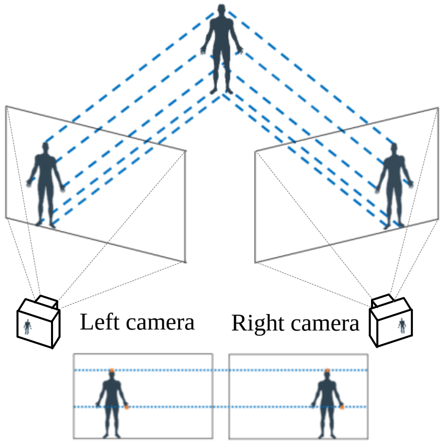


Fig. 2: The outline of a stereo vision system

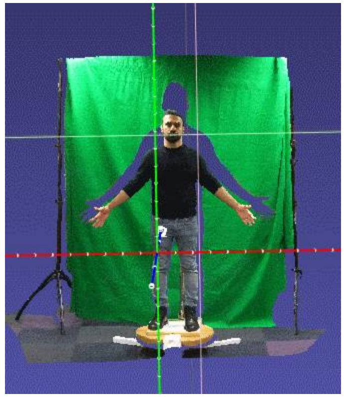


Fig. 3: UoA-IVSLab Stereo Workstation

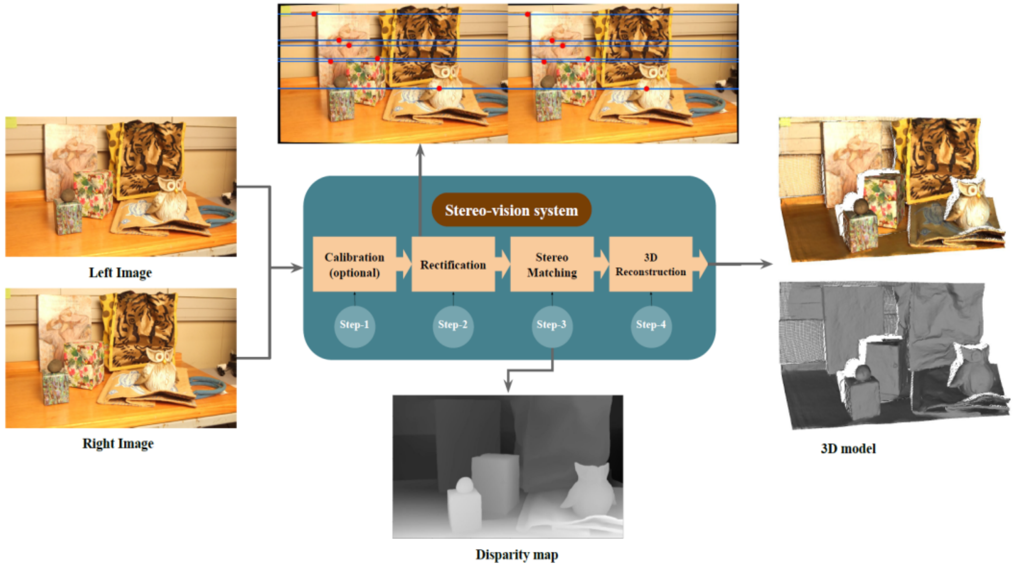


Fig. 4: The end-to-end pipeline from a pair of stereo images to the final 3D reconstruction

What is stereo matching?

- In a stereo-vision system, a 3D point in the scene is projected into the left and right camera planes
- The process of identifying corresponding 2D projections in stereo images is known as **Stereo Matching**.

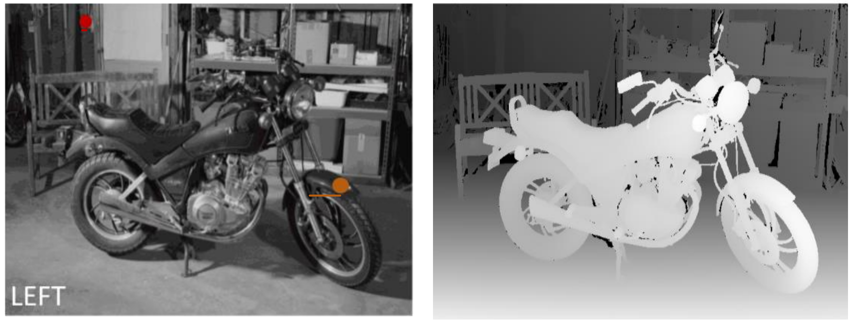


Fig. 5: A pair of stereo images with the corresponding ground truth disparity map

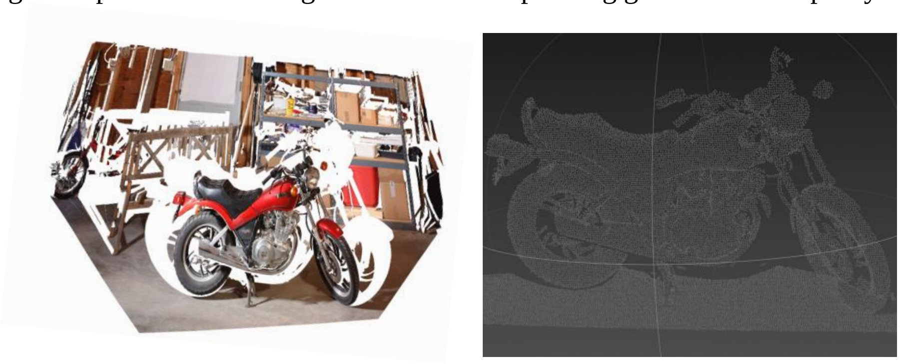


Fig. 6: 3D reconstruction example from a stereo-vision system

Stereo Matching is intrinsically solved using a labeling problem.

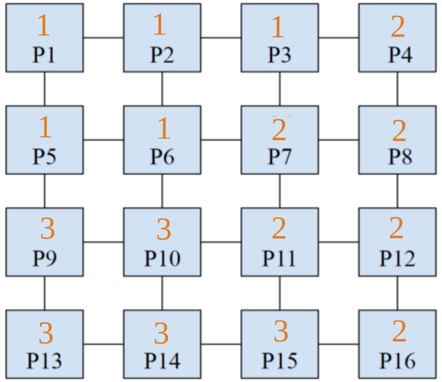


Fig. 7: Stereo Matching as a labeling problem, where a set of pixels is labeled by a set of possible disparities

A 4×4 stereo image labelled by a set of possible disparities as $\{1,2,3\}$

Stereo Matching energy function usually has two terms

$$E(\mathbf{w}) = E_{data}(\mathbf{w}) + \lambda E_{spatial}(\mathbf{w})$$

$\mathbf{w}$ is an $n \times m$ matrix of integer variables with $\mathbf{w} \in \{d_{min}, \dots, d_{max}\}^{n \times m}$

where $n \times m$ is the size of stereo images, d_{min} and d_{max} are the minimum and maximum possible disparity values

E_{data} computes the cost of allocating a disparity value to the given pixel

$E_{spatial}$ is a constraint to the disparities allocated to the neighboring pixels

$\lambda \in \mathbb{R}$ is the penalty factor for the second term

Depending on how $E_{spatial}$ is defined the complexity of the corresponding minimization changes

Quantum Annealing

QA refers to the procedure of identifying a state of a quantum dynamical system that has the lowest energy in accordance with the time-dependent Hamiltonian

QA aims to address **optimization** problems

Quantum Unconstraint Binary Optimization (**QUBO**)

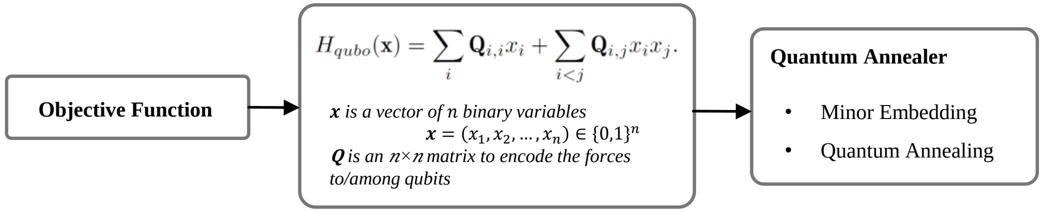


Fig. 8: Quantum model preparation for Quantum Annealing

D-Wave Advantage QPU

A Pegasus working graph with 5000 qubits and 35000 couplers

Main contribution:

Modelling different Stereo Matching objective functions as QUBOs for Quantum Annealing optimization

"**Annealing**" in metallurgy refers to a process of slowly cooling a material to reach a low-energy state, minimizing defects and improving its properties.

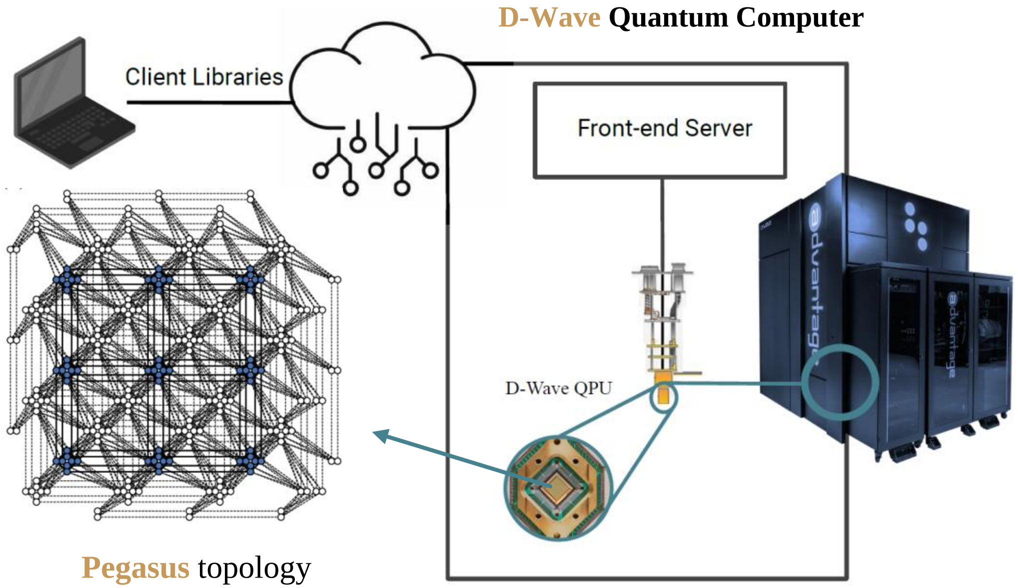


Fig. 9: The outline of a D-Wave quantum computer

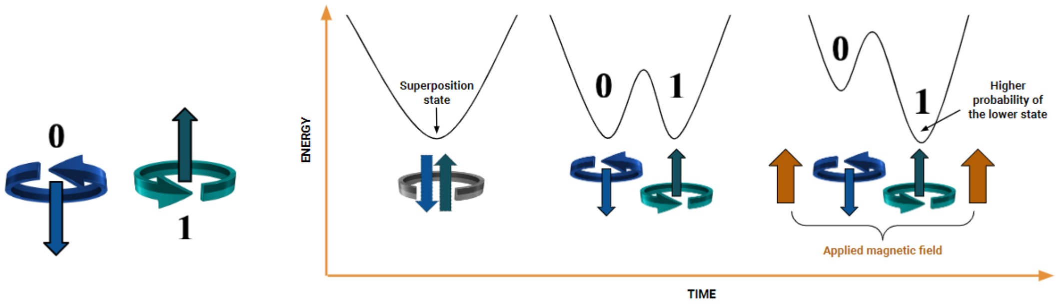


Fig. 10: Energy diagram of the D-Wave Quantum Annealing for a single qubit

Proposed Quantum Models for Stereo Matching

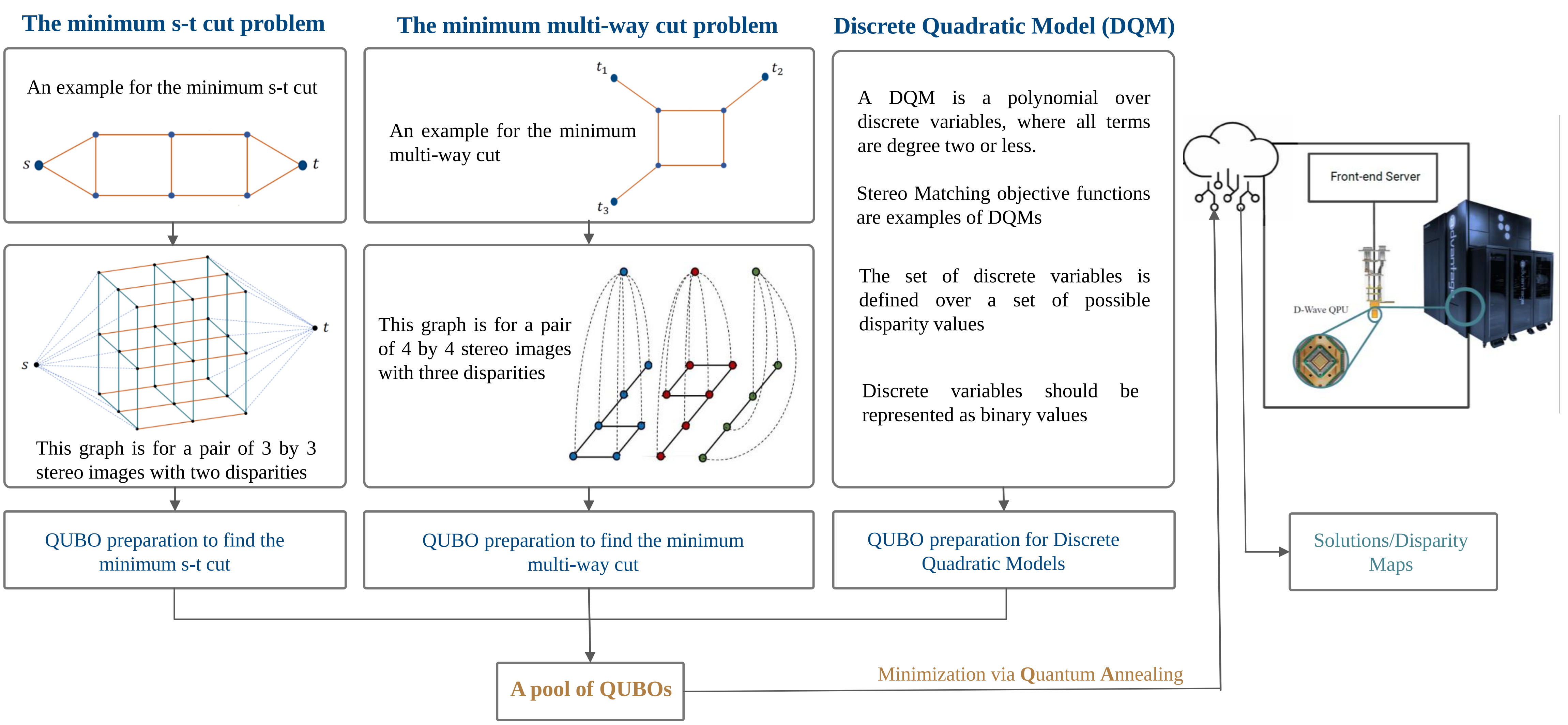


Fig. 11: The outline of our proposed quantum models

QUBO Variable/Preparation Complexities

Given $(\boldsymbol{n} \times \boldsymbol{m})$ as the size of a pair of stereo images with $\boldsymbol{k}$ disparities

Tab. 1: The qubit complexity for the quantum models

QUBO	DESCRIPTION	#VARIABLE NUMBER
QUBO-cruz*	The minimum s-t cut	$7nmk + 9nm - 2nk - 2mk - 2n - 2m + 2$
QUBO-st	The minimum s-t cut	$nmk + nm + 2$
QUBO-mw	The minimum multi-way cut	$nmk + k^2$
QUBO-dqm	Discrete quadratic model	nmk

* The only Quantum Stereo Matching method found in the literature

Tab. 2: The QUBO preparation complexity for the quantum models

QUBO	Problem type	QUBO Preparation Complexity
QUBO-cruz*	Polynomial time	$O(nmk)$
QUBO-st	Polynomial time	$O(nmk)$
QUBO-mw	NP-hard	$O(nmk^3)$
QUBO-dqm	NP-hard	$O(nmk^2)$

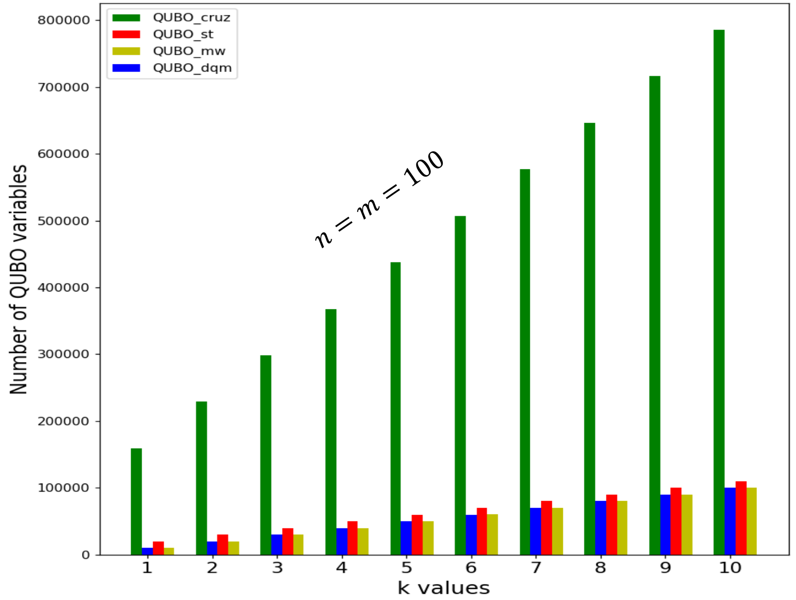


Fig. 12: Number of QUBO variables for each model, given a pair of 100×100 stereo images

- The lack of available qubits on current quantum hardware is a significant limitation.
- The fewer variables in the QUBO, the better
- As QUBO variables are mapped to the physical qubits on the hardware.

Hybrid quantum-classical segment-based Stereo Matching

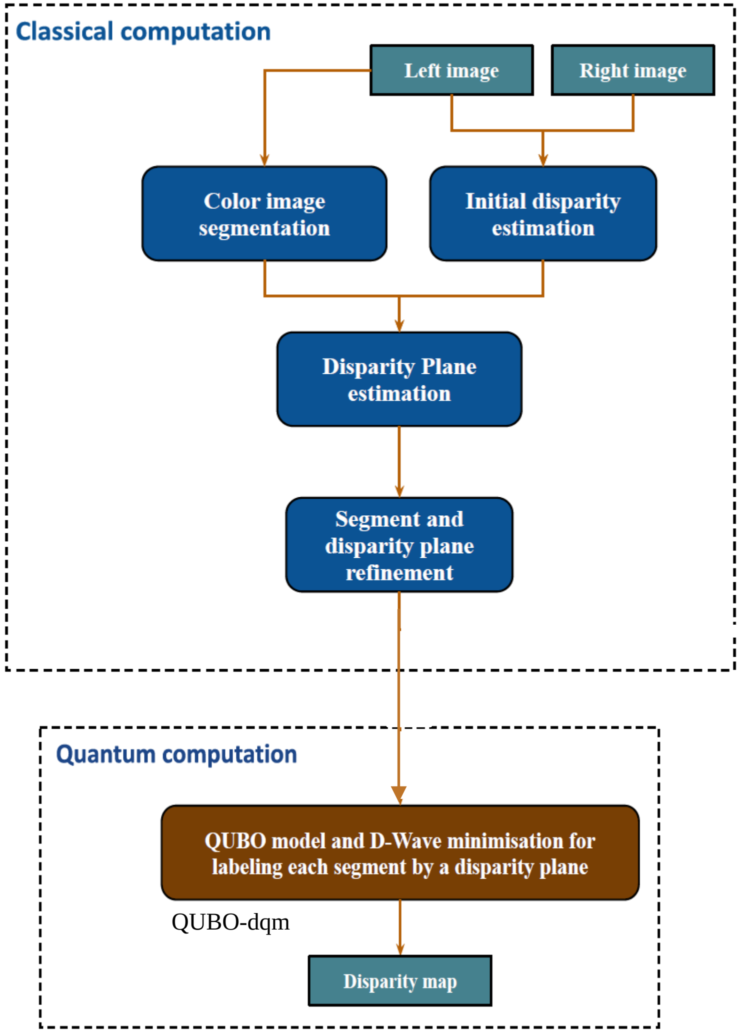


Fig. 13: The outline of the hybrid quantum-classical segment-based Stereo Matching method

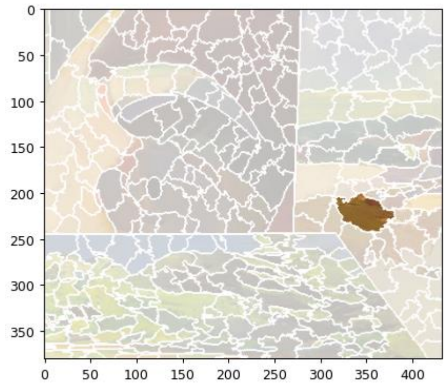


Fig. 14: Color image segmentation (using Quickshift algorithm)

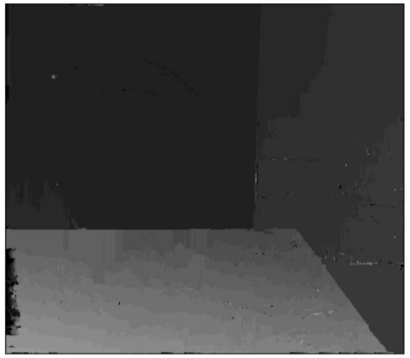


Fig. 15: Initial disparity estimation (using a cross-based local method)

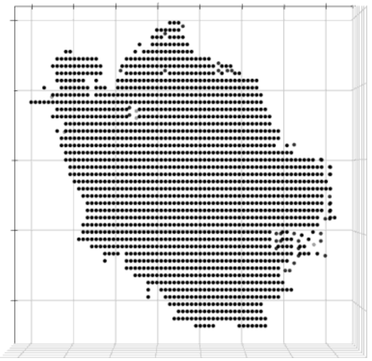


Fig. 16: Initial disparity estimation for the segment

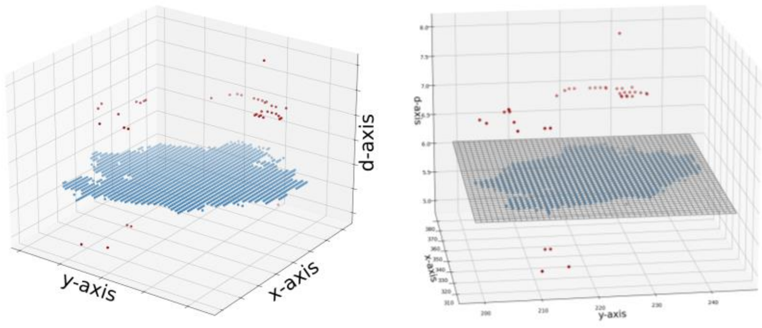


Fig. 17: Plane fitting to the corresponding initial disparity belonging to the selected segment

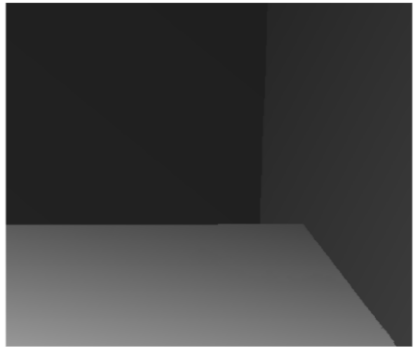


Fig 18. Final disparity map

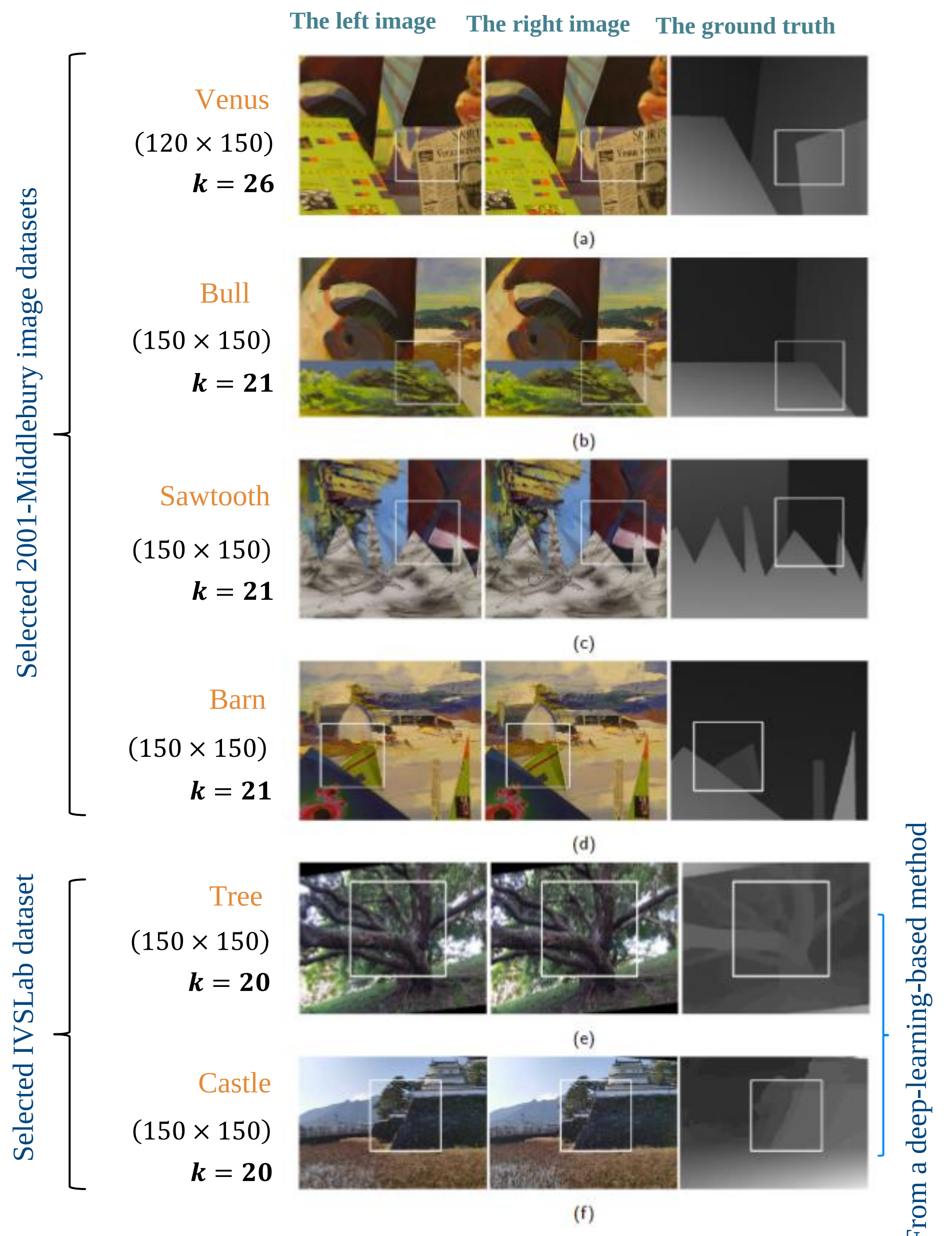


Fig. 19: Selected stereo images for our evaluation/comparison

- D-Wave Advantage QPU provides 5000 Qubits
- D-Wave hybrid Q/C solvers can handle problems with a significantly higher number of variables than those directly solvable by a D-Wave QPU
- As a proof of concept, we utilize a D-Wave hybrid solver to minimize the proposed quantum Stereo Matching models
- Comparing the results with that of four selected state of the art CV minimization methods

Tab. 3: Benchmark Stereo Matching methods

Computation	Methods	Description
Quantum Annealing	QUBO-st	The minimum s-t cut
	QUBO-mw	The minimum multi-way cut
	QUBO-dqm	Discrete quadratic model
Classical Computation	Swap	Swap-move graph-cut minimization
	Expansion	Expansion-move graph-cut minimization
	BP-M	The max product version of Loopy Belief Propagation
	TRW-S	The improved Tree-Reweighted Message-passing

Experimental Results

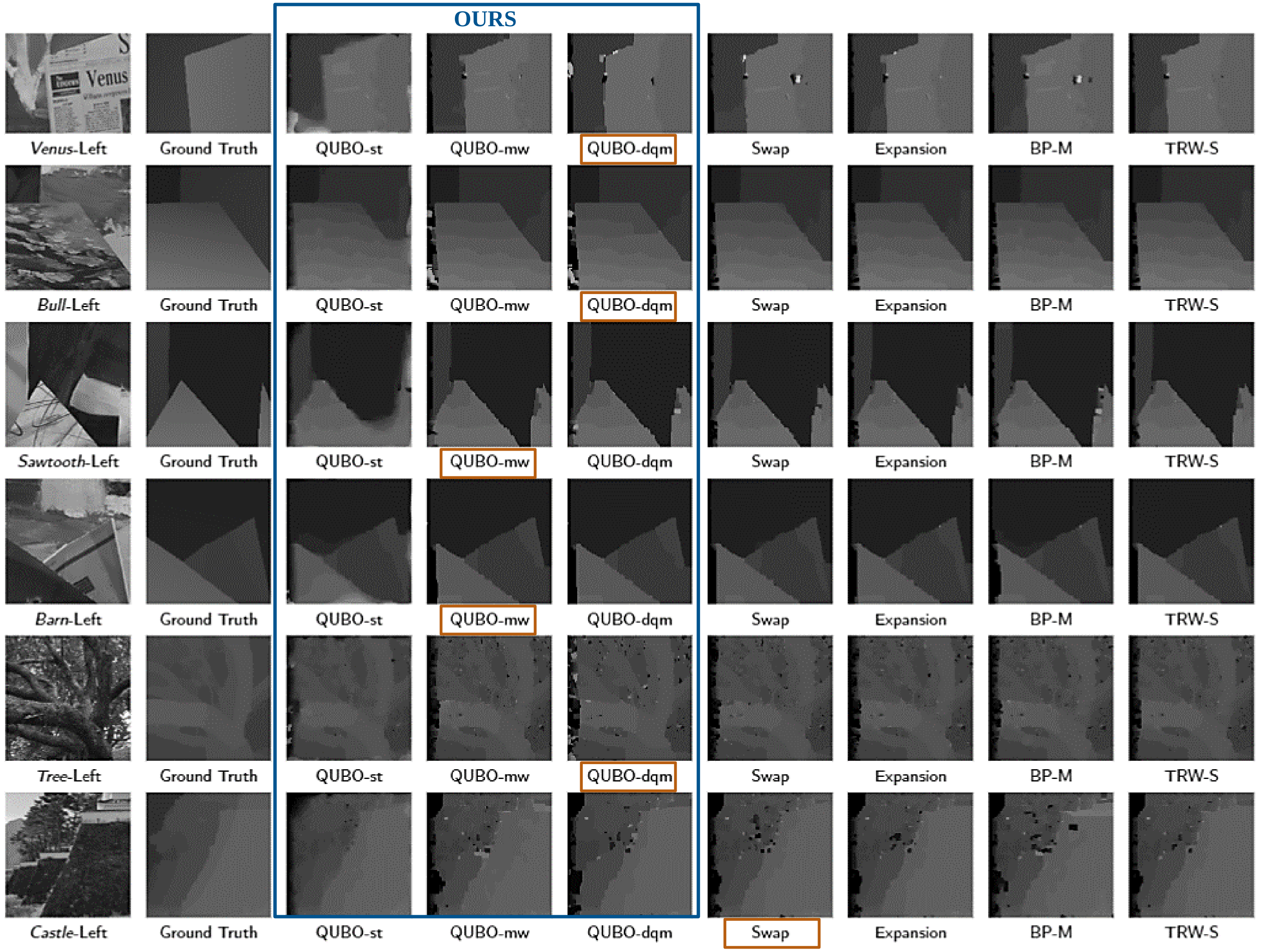


Fig. 20: Computed disparity maps from our stereo datasets based on the selected benchmark Stereo Matching methods

Since QUBO-dqm is a direct equivalent to the defined Stereo Matching energy function, we can compare the energies of the solutions with that of classical minimization methods

QUBO-st and QUBO-mw are based on graph-cuts and their returned energies are expected to differ

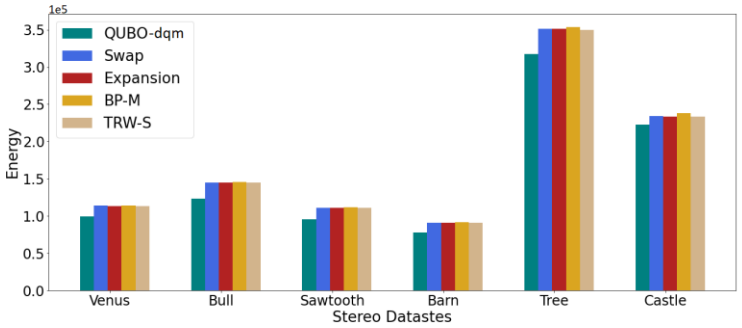
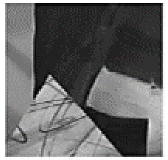
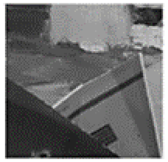


Fig. 21: A comparison between energies obtained by QUBO-dqm and the benchmarking classical minimization methods

- QUBO-dqm demonstrated a capacity to obtain solutions of lower energy in comparison to the iterative classical minimization methods for each provided stereo dataset.
- This observation underscores the effectiveness of QUBOs when solved by D-Wave hybrid solvers

Numerical Evaluation

Tab. 4: Numerical evaluation for benchmarking algorithms on the prepared stereo datasets. The best results are shown in gray

	*	Method	*rms*	*bad-0.5 (%)*	*bad-1.0 (%)*	*bad-2.0 (%)*	*accuracy (%)*
	*Venus*	QUBO-st	2.67	50.55	17.70	15.33	49.45
		QUBO-mw	1.87	45.16	8.92	8.33	54.84
		QUBO-dqm	2.25	40.44	10.45	10.01	59.56
		Swap	2.09	47.57	10.23	9.48	52.43
		Expansion	1.94	43.81	9.76	9.29	56.19
		BP-M	1.96	47.09	9.30	8.75	52.91
		TRW-S	1.92	44.07	9.54	8.96	55.93
	*Bull*	QUBO-st	2.22	38.73	11.42	9.81	61.27
		QUBO-mw	2.19	36.16	6.92	6.73	63.84
		QUBO-dqm	2.33	36.07	7.08	6.75	63.93
		Swap	2.38	37.43	7.40	6.64	62.57
		Expansion	2.38	37.28	7.36	6.77	62.72
		BP-M	2.39	37.42	7.25	6.60	62.58
		TRW-S	2.38	37.06	7.32	6.66	62.94
	*Sawtooth*	QUBO-st	2.92	33.22	20.63	17.78	66.78
		QUBO-mw	2.16	22.09	10.26	9.96	77.91
		QUBO-dqm	2.27	22.54	10.26	9.83	77.46
		Swap	2.44	22.76	10.27	9.84	77.24
		Expansion	2.41	22.85	10.44	10.12	77.15
		BP-M	2.41	23.59	10.30	9.90	76.41
		TRW-S	2.36	22.67	10.36	10.00	77.33
	*Barn*	QUBO-st	2.19	32.32	13.53	11.80	67.68
		QUBO-mw	2.29	14.20	7.53	7.53	85.8
		QUBO-dqm	2.27	14.37	7.41	7.41	85.63
		Swap	2.23	16.11	7.51	7.39	83.89
		Expansion	2.21	16.09	7.55	7.52	83.91
		BP-M	2.38	20.21	8.25	8.18	79.79
		TRW-S	2.23	15.54	7.33	7.29	84.46
	*Tree*	QUBO-st	2.62	33.82	15.36	11.66	66.18
		QUBO-mw	3.16	33.00	13.92	10.59	67.0
		QUBO-dqm	2.99	24.99	13.22	10.47	75.01
		Swap	3.32	33.85	14.39	10.85	66.15
		Expansion	3.27	31.73	13.23	10.61	68.27
		BP-M	3.16	33.45	13.91	10.45	66.55
		TRW-S	3.12	31.81	13.22	10.05	68.19
	*Castle*	QUBO-st	2.13	36.54	16.15	11.73	63.46
		QUBO-mw	2.82	41.46	17.69	13.98	58.54
		QUBO-dqm	2.74	34.52	17.62	14.46	65.48
		Swap	2.83	32.85	17.25	14.28	67.15
		Expansion	2.76	33.68	16.94	14.12	66.32
		BP-M	2.99	41.12	21.21	17.92	58.88
		TRW-S	2.66	33.36	16.62	13.12	66.64

As have proof of concept and demonstrate the feasibility of our proposed quantum models, we have employed the D-Wave hybrid solver to minimize them and compared their performance with the classical benchmark algorithms.

The results show that the quantum models performed competitively with the classical ones in some cases, and in most cases, they outperformed them.

Let $\mathcal{D}$ be the disparity map, and $\mathcal{T}$ be the ground truth disparity map defined by $n \times m$ matrices of integers

- **Root-mean-squared error**

$$rms = \sqrt{\frac{1}{nm}\sum_{i=1}^{n}\sum_{j=1}^{m}\left(\mathcal{D}_{i,j} - \mathcal{T}_{i,j}\right)^2}$$

- **Percentage of bad (missed) matching pixels**

$$bad\text{-}\beta = \left(\frac{1}{nm}\sum_{i=1}^{n}\sum_{j=1}^{m}\left(|\mathcal{D}_{i,j} - \mathcal{T}_{i,j}| > \beta\right)\right) \times 100$$

- **Percentage of exact disparities**

$$accuracy = \left(\frac{1}{nm}\sum_{i=1}^{n}\sum_{j=1}^{m}\delta\left(\mathcal{D}_{i,j}, \mathcal{T}_{i,j}\right)\right) \times 100$$

Hybrid quantum-classical segment-based Stereo Matching

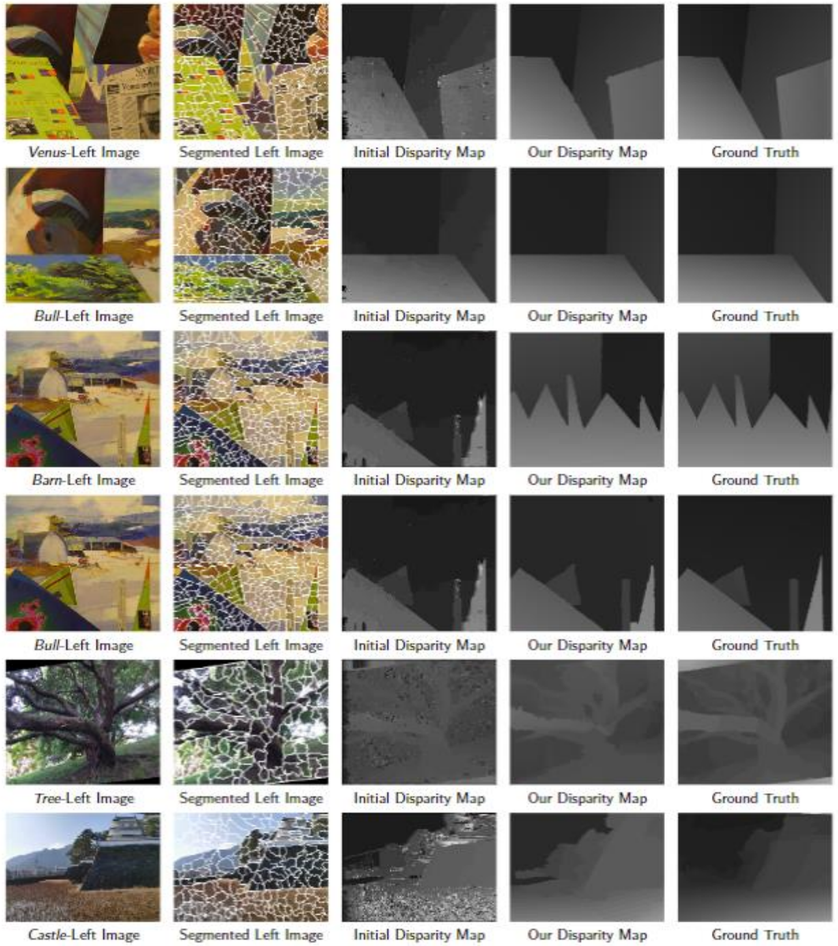


Fig. 22: Computed disparity maps from our stereo datasets based on the hybrid quantum-classical segment-based Stereo Matching.

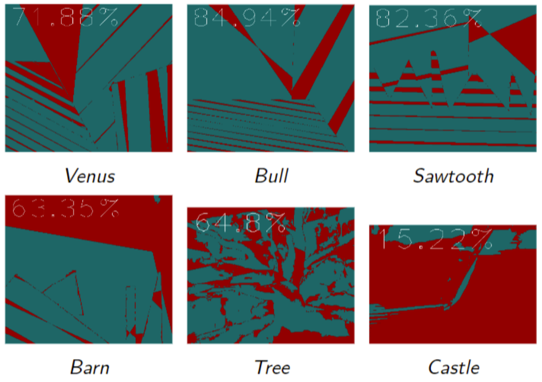


Fig. 23: The percentage of the exact disparity values computed by the hybrid model.

- **Green** areas show the accuracy percentage shown in Tab. 5. **Red** areas belong to the inaccurate disparities

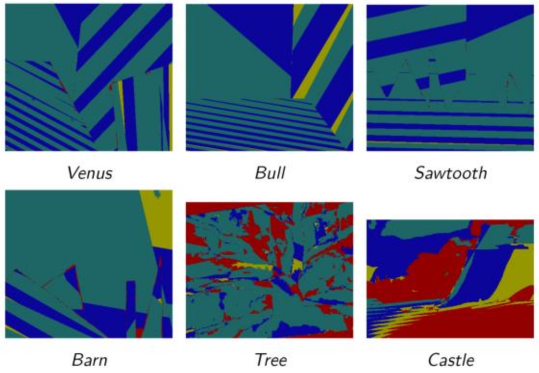


Fig. 24: Disparity-variation representation

- Let $\varphi_{i,j} = \left|\mathcal{D}_{i,j} - \mathcal{T}_{i,j}\right|$. In the disparity-variation representation, the color of the pixels follows the below conditions

■: if $\varphi_{ij} \le 0.5$ ■: if $0.5 < \varphi_{ij} \le 1.0$

■: if $1.0 < \varphi_{ij} \le 2.0$ ■: if $\varphi_{ij} > 2.0$.

Tab. 5: Numerical evaluation for the Hybrid quantum-classical segment-based Stereo Matching

Dataset	rms	*bad*-0.5	*bad*-1.0	*bad*-2.0	*accuracy*
Venus	0.61	28.12	0.33	0.30	71.88
Bull	0.40	0.00	0.00	0.00	84.94
Sawtooth	0.54	17.64	0.26	0.25	82.36
Barn	1.19	36.65	7.23	1.18	63.35
Tree	2.50	35.20	10.26	5.64	64.80
Castle	3.34	84.78	47.62	29.58	15.22

Limitations and Challenges of Quantum Stereo Matching

Classical challenges

- Classically defining an accurate energy function that reflects the scene structure is challenging.
- This energy function calculates how well a particular disparity map aligns with the input stereo images.
- Main challenges in defining a Stereo Matching energy function
    - Occlusions, photometric variations, textureless regions, depth discontinuities, and fine tuning the parameter settings.

Left image　Ground truth　QUBO-st　QUBO-mw

Fig. 25: An example of handling depth discontinuities based on defined Stereo Matching energy function

Quantum (D-Wave) Limitations

D-Wave Physical Qubits:

D-Wave QPUs are characterized by a finite number of qubits, which poses a fundamental constraint on the complexity of problems they can effectively solve.

QUBO Preparation

The given objective should be converted to a standard model

D-Wave Minor Embedding:

Minor Embedding finds a way to represent every vertex of the QUBO graph using one or multiple vertices of the D-Wave graph, which is computationally expensive

D-Wave Bits of Precision:

D-Wave QPUs are not high-precision digital computers, and they are analog and noisy

D-Wave QPUs use a precision of 4 to 5 bits for the floating point coefficients

Main contributions

Quantum Stereo Matching based on the minimum *s-t* cuts

- **IVCNZ-2021 Conference, New Zealand**

Quantum Stereo Matching based on the minimum multi-way cuts

- **International Journal of Unconventional Computing**

Quantum Stereo Matching based on discrete quadratic models

- **Future Generation Computer Systems**

Hybrid quantum-classical segment-based Stereo Matching

- **ACIVS-2023 Conference, Japan**

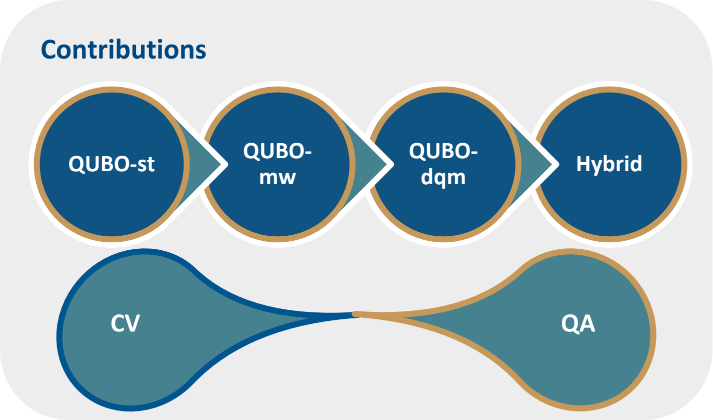


Fig. 26: The outline of our contributions

The future of QA in Computer Vision

Quantum annealing holds significant potential for enhancing computer vision by providing faster and more efficient solutions to complex optimization problems.

However, realizing this potential will require overcoming current hardware limitations

As quantum hardware improves, with more qubits and better error rates, the applicability of quantum annealing in computer vision will expand.

Given the significant search space for objective functions in computer vision problems involving images, the near future of quantum computer vision is likely to rely on hybrid methods rather than solely on quantum annealing.

UNIVERSITY OF AUCKLAND
Waipapa Taumata Rau
NEW ZEALAND

SCIENCE

Thank you for your time and attention

Special thanks to my supervisors

Shahrokh Heidari