

Appendix of ACMMM 2024 Submission #4489

Task-Oriented Multi-Bitstream Optimization for Image Compression and Transmission via Optimal Transport

A PROOF OF EQ. (12)

We assume that there is a change in network capacity between two consecutive transmission periods, e.g., from ρ_1 to ρ_2 . The matrix $K_{\rho_2}^u$ associated with the link capacities within the period ρ_2 can then be denoted by $K_{\rho_2}^u = K_{\rho_1}^u + \Delta K^u$. Meanwhile, we can leverage the variable $\psi_i^{\rho_1}$ to speed up the iteration process of $\psi_i^{\rho_2}$. We then show the initialization detail of $\psi_i^{\rho_2}$. Assuming that $\psi_\tau^{\rho_2}$ is approximately equal to $\psi_\tau^{\rho_1}$, we can obtain $\psi_{\tau-1}^{\rho_2}$ and $\psi_{\tau-2}^{\rho_2}$ via the following derivation:

$$\psi_\tau^{\rho_2} \approx \psi_\tau^{\rho_1} \quad (13a)$$

$$\psi_{\tau-1}^{\rho_2} = (\psi_\tau^{\rho_2} \odot K_{\rho_2}^u) \text{diag}(u_\tau) K^{*T} \quad (13b)$$

$$\begin{aligned} &= (\psi_\tau^{\rho_2} \odot (K_{\rho_1}^u + \Delta K^u)) \text{diag}(u_\tau) K^{*T} \\ &\approx (\psi_\tau^{\rho_1} \odot (K_{\rho_1}^u + \Delta K^u)) \text{diag}(u_\tau) K^{*T} \\ &= (\psi_\tau^{\rho_1} \odot K_{\rho_1}^u) \text{diag}(u_\tau) K^{*T} + (\psi_\tau^{\rho_1} \odot \Delta K^u) \text{diag}(u_\tau) K^{*T} \\ &= \psi_{\tau-1}^{\rho_1} + (\psi_\tau^{\rho_1} \odot \Delta K^u) \text{diag}(u_\tau) K^{*T} \\ &= \psi_{\tau-1}^{\rho_1} + f(\psi_\tau^{\rho_1}, \Delta K^u, u_\tau), \\ \psi_{\tau-2}^{\rho_2} &= (\psi_{\tau-1}^{\rho_2} \odot K_{\rho_2}^u) \text{diag}(u_{\tau-1}) K^{*T} \quad (13c) \\ &= (\psi_{\tau-1}^{\rho_2} \odot (K_{\rho_1}^u + \Delta K^u)) \text{diag}(u_{\tau-1}) K^{*T} \\ &\approx (\psi_{\tau-1}^{\rho_2} \odot K_{\rho_1}^u) \text{diag}(u_{\tau-1}) K^{*T} + (\psi_{\tau-1}^{\rho_2} \odot \Delta K^u) \text{diag}(u_{\tau-1}) K^{*T} \\ &\approx (\psi_{\tau-1}^{\rho_1} \odot K_{\rho_1}^u) \text{diag}(u_{\tau-1}) K^{*T} + (((\psi_\tau^{\rho_1} \odot \Delta K^u) \text{diag}(u_\tau) K^{*T}) \odot K_{\rho_1}^u) \text{diag}(u_{\tau-1}) K^{*T} \\ &\quad + (\psi_{\tau-1}^{\rho_1} \odot \Delta K^u) \text{diag}(u_{\tau-1}) K^{*T} + (((\psi_\tau^{\rho_1} \odot \Delta K^u) \text{diag}(u_\tau) K^{*T}) \odot \Delta K^u) \text{diag}(u_{\tau-1}) K^{*T} \\ &= \psi_{\tau-2}^{\rho_1} + (\psi_{\tau-1}^{\rho_1} \odot \Delta K^u) \text{diag}(u_{\tau-1}) K^{*T} + (((\psi_\tau^{\rho_1} \odot \Delta K^u) \text{diag}(u_\tau) K^{*T}) \odot (K_{\rho_1}^u + \Delta K^u)) \text{diag}(u_{\tau-1}) K^{*T}. \end{aligned}$$

Notice that the values of the elements of $K_{\rho_1}^u$, ΔK^u and u_i all fall within the range of $[-1, 1]$, K^* is a binary matrix (with values of either 0 or 1). Furthermore, the elements of third item of $\psi_{\tau-2}^{\rho_2}$ is significantly smaller than these of the first two items, and therefore can be disregarded in the calculation. The rest iteration process of $\psi_i^{\rho_2}$ for $i = \tau - 3, \dots, 0$ can be obtained in the same manner as follows:

$$\psi_i^{\rho_2} = \psi_i^{\rho_1} + (\psi_{i+1}^{\rho_1} \odot \Delta K^u) \text{diag}(u_{i+1}) K^{*T}. \quad (14)$$

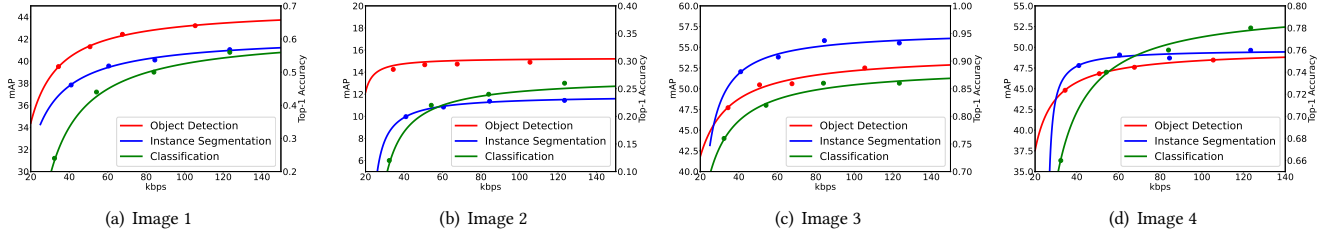
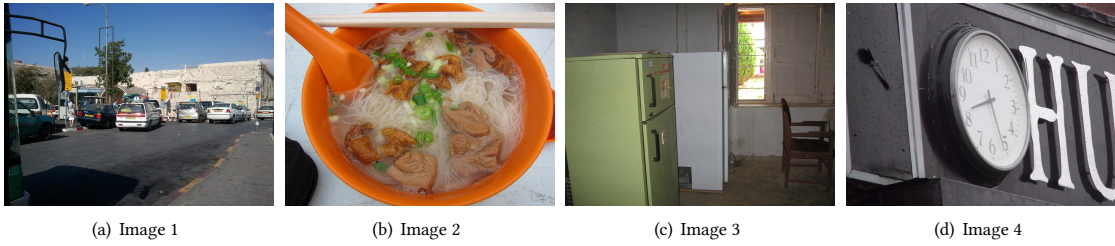
We assume that the capacity of link \bar{e}_i changes between period ρ_1 and ρ_2 . Correspondingly, the non-zero elements of ΔK^u solely exist in the \bar{e}_i -th column. For each link with altered capacity, we denote the \bar{e}_i -th column of $\psi_{i+1}^{\rho_1}$ and ΔK^u as $[\psi_{i+1}^{\rho_1}]_{\bar{e}_i}$ and $[\Delta K^u]_{\bar{e}_i}$, respectively. Similarly, we designate the \bar{e}_i -th element of vector u_{i+1} and the \bar{e}_i -th row of K^T as $[u_{i+1}]_{\bar{e}_i}$ and $[K^T]_{\bar{e}_i}$. Then, the matrix product $(\psi_{i+1}^{\rho_1} \odot \Delta K^u) \text{diag}(u_{i+1})$ can be expressed as a matrix derived from the Hadamard product of vectors $[u_{i+1}]_{\bar{e}_i} ([\psi_{i+1}^{\rho_1}]_{\bar{e}_i} \odot [\Delta K^u]_{\bar{e}_i})$ and $[K^T]_{\bar{e}_i}$. In this case, we replace the complex matrix multiplication with vector products, thereby reducing computational complexity from $O(n^2)$ to $O(n)$.

B ABLATION STUDY OF TRANSMISSION PERIOD Γ

In the aforementioned experiments, we set the values of transmission period Γ as $\{7, 7, 8, 37\}$ according to the size of each network. Here, we further investigate the impact of Γ on the performances of *Tombo*. We fix the duration of each interval in the time-expanded network as $\Delta\tau = 1$ millisecond and change Γ over the four networks. The results of total objectives, transmission costs, surrogate task values and average bitrates are presented in Table 5. As the results shown, when the value of Γ increases, more bits can be transmitted, resulting in smaller task losses at the receiving side. However, this also leads to higher transmission costs. Consequently, the overall objective may either increase, as observed in Janetbackbone and Carnet, or remain basically unchanged, as seen in B4 and UsCarrier. In Carnet, with the increase in transmission period, there is a reduction in task loss due to the enhanced task performance facilitated by higher bitrates. However, in the other three networks, this improvement is tiny. Moreover, as the value of Γ increases, the increase in τ results in an expansion in the problem size, thereby prolonging the running time.

Table 5: Ablation results w.r.t. transmission period Γ , with a fixed $\Delta\tau = 1$ ms, and varied τ and Γ .

Topology	Γ (ms)	Objective			Violation \downarrow	Avg. bitrate (bpp) \downarrow	Running time (s) \downarrow
		Total Obj. \downarrow	Trans. cost \downarrow	Task-surr. \downarrow			
B4	7	86.16	46.69	39.47	2.27e-13	0.1191	3.00
	8	85.24	46.74	38.50	0	0.1209	3.37
	9	85.26	46.76	38.49	0	0.1209	3.29
	10	85.26	46.76	38.49	0	0.1209	3.10
	14	85.26	46.76	38.49	0	0.1209	3.42
Janetbackbone	7	298.49	128.71	169.78	2.88e-3	0.1284	11.19
	8	307.19	138.13	169.05	3.23e-3	0.1288	11.41
	9	313.93	145.43	168.49	3.47e-3	0.1291	11.34
	10	319.66	151.46	168.20	3.59e-3	0.1292	11.57
	14	331.44	163.68	167.76	3.76e-3	0.1295	12.03
Carnet	8	593.63	181.82	411.81	0.083	0.1187	131.09
	9	613.19	225.36	387.83	0.084	0.1249	131.80
	10	636.68	252.81	383.87	0.084	0.1265	138.04
	11	672.45	291.74	380.71	0.084	0.1280	138.31
	16	848.65	480.89	367.75	0.085	0.1350	138.10
UsCarrier	15	312.42	162.51	149.91	0.46	0.1246	56.10
	20	311.84	169.39	142.45	9.09e-13	0.1277	57.84
	25	311.85	169.39	142.45	9.09e-13	0.1277	61.56
	37	311.85	169.39	142.45	1.36e-12	0.1277	42.71
	50	311.85	169.39	142.45	9.09e-13	0.1277	81.98

**Figure 7: Fitted curves across three machine vision tasks. Images 1-4 are randomly sampled from dataset COCO2017.****Figure 8: The visual contents of the Images 1-4.**

C DETAIL OF THE FITTED FUNCTIONS ACROSS THREE IMAGE TASKS

We provide the fitted values for the coefficients $\{a_{l,\theta}, b_{l,\theta}, c_{l,\theta}\}$ in the bitrate-task performance fitting function $\Phi_{l,\theta}(r^l) = -a_{l,\theta}/(r^l - b_{l,\theta}) + c_{l,\theta}$ in Table 6, and illustrate the fitted curves for three machine vision tasks and different images in Fig. 7. The images are shown in Fig. 8. Note that the performance of the same task varies greatly among different images at the same bitrate, as demonstrated in Figs. 7(a) and 7(c). For instance, at a low bitrate of 40kbps, the mean average precision for the instance segmentation task in Fig. 7(c) can exceed 50, whereas for

Table 6: Illustration of the fitted values for the coefficients $\{a_{l,\theta}, b_{l,\theta}, c_{l,\theta}\}$ in bitrate-task performance function $\Phi_{l,\theta}(r^l)$.

Image Task θ	Object Detection				Instance Segmentation				Classification			
Image l	1	2	3	4	1	2	3	4	1	2	3	4
$a_{l,\theta}$	1567.1	131.3	1939.0	1164.6	1576.3	394.7	1246.3	299.3	84.61	20.31	29.73	23.91
$b_{l,\theta}$	49.5	157.8	43.2	103.9	53.8	204.0	159.9	249.4	104.66	193.33	97.49	151.41
$c_{l,\theta}$	44.8	15.3	54.2	49.7	42.3	11.9	57.0	49.7	0.62	0.27	0.89	0.80

Table 7: Illustration of the fitting metric R^2 for each fitted curve.

Image Task	R^2			
	Image 1	Image 2	Image 3	Image 4
Object Detection	0.9969	0.9872	0.9544	0.9996
Instance Segmentation	0.9884	0.9877	0.9363	0.9478
Classification	0.9957	0.9967	0.9233	0.9565

Fig. 7(a), with a higher bitrate of $140kbps$, the mean average precision still falls below 50. This presents a challenge for rate allocation among different tasks. Moreover, across three machine vision tasks, such as object detection and instance segmentation illustrated in Fig. 7(d), the same bitrate yields different task performance. Hence, an identical bitrate allocation strategy for various images and machine vision tasks is not suitable. Furthermore, we provide the fitting metric R^2 for each curve in Table 7, demonstrating that the derived functions accurately match the data points collected from various image tasks.