1. Supplementary Material

In this supplementary material, we present several figures to demonstrate the effectiveness and robustness of our method. In cel animation production, different coloring scenarios exist, including close-up, medium shot, and long shot. Each of these scenes features different characteristics of the characters and varying degrees of motion. For example, long shots typically involve large-scale movements. Therefore, in the supplementary section, we show the coloring results of our method in three distinct scenes, as shown in Figures 1, 2, and 3. For these experiments, we chose a sequence length of 10, using k=0 and k=10 as reference keyframes for the coloring process. The experimental results show that our method exhibits robust performance in all three scenarios.

Furthermore, to explore the gain of varying the number of reference keyframes, we conducted an experiment with k = 0, 5 and 10 as keyframes for coloring, as shown in Figure 4. The results do not indicate significant differences compared to the configuration k = 0, 10. We hypothesize that the number of reference frames may have a diminishing return effect, where too many references could complicate the coloring process without providing substantial benefit.

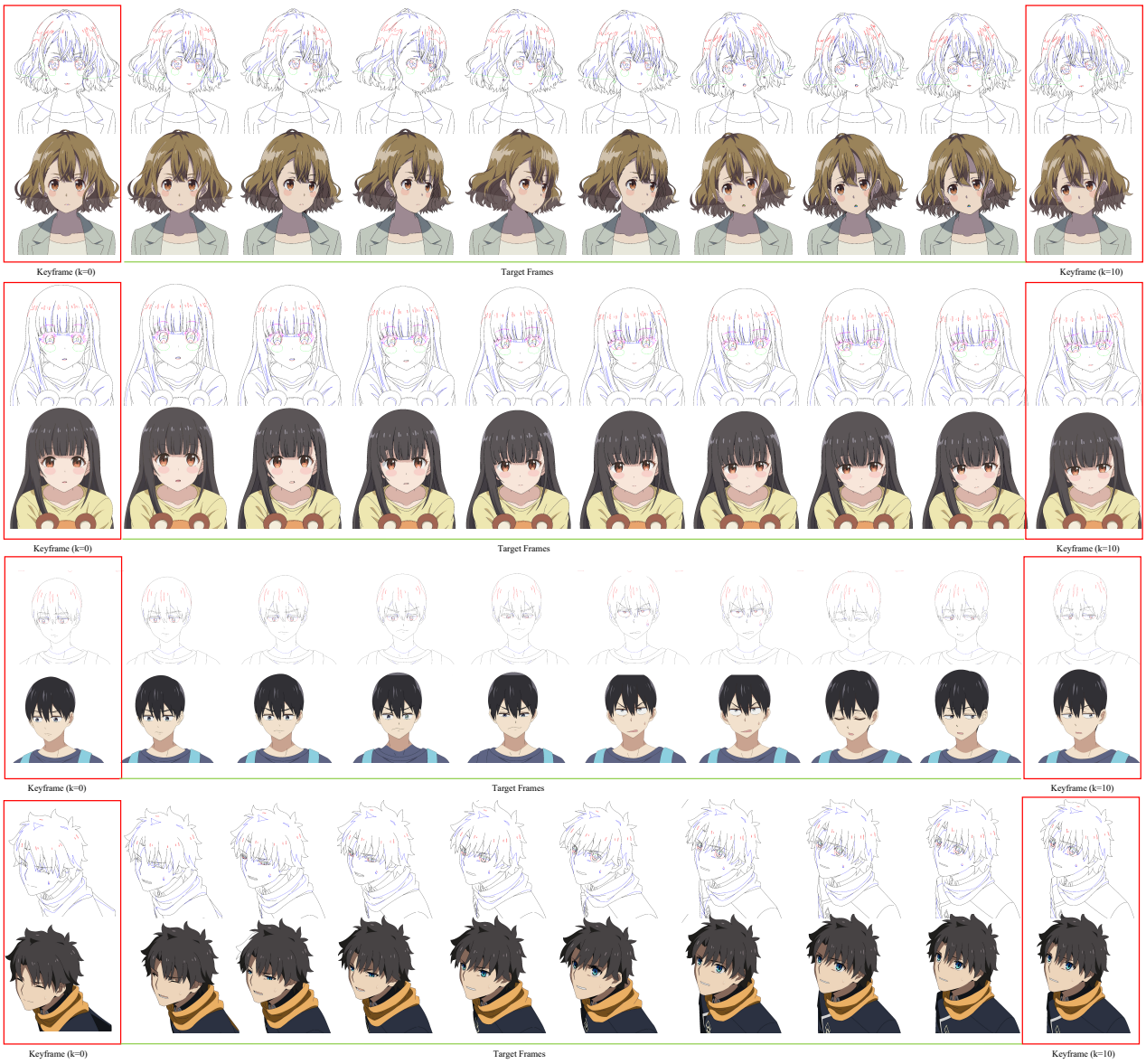

Figure 1. This figure presents the coloring results of our method on a close-up shot (typically including the shoulders to the head). We use a sequence length of 10 for the carts, with the first and last frames as reference keyframes (highlighted in red).

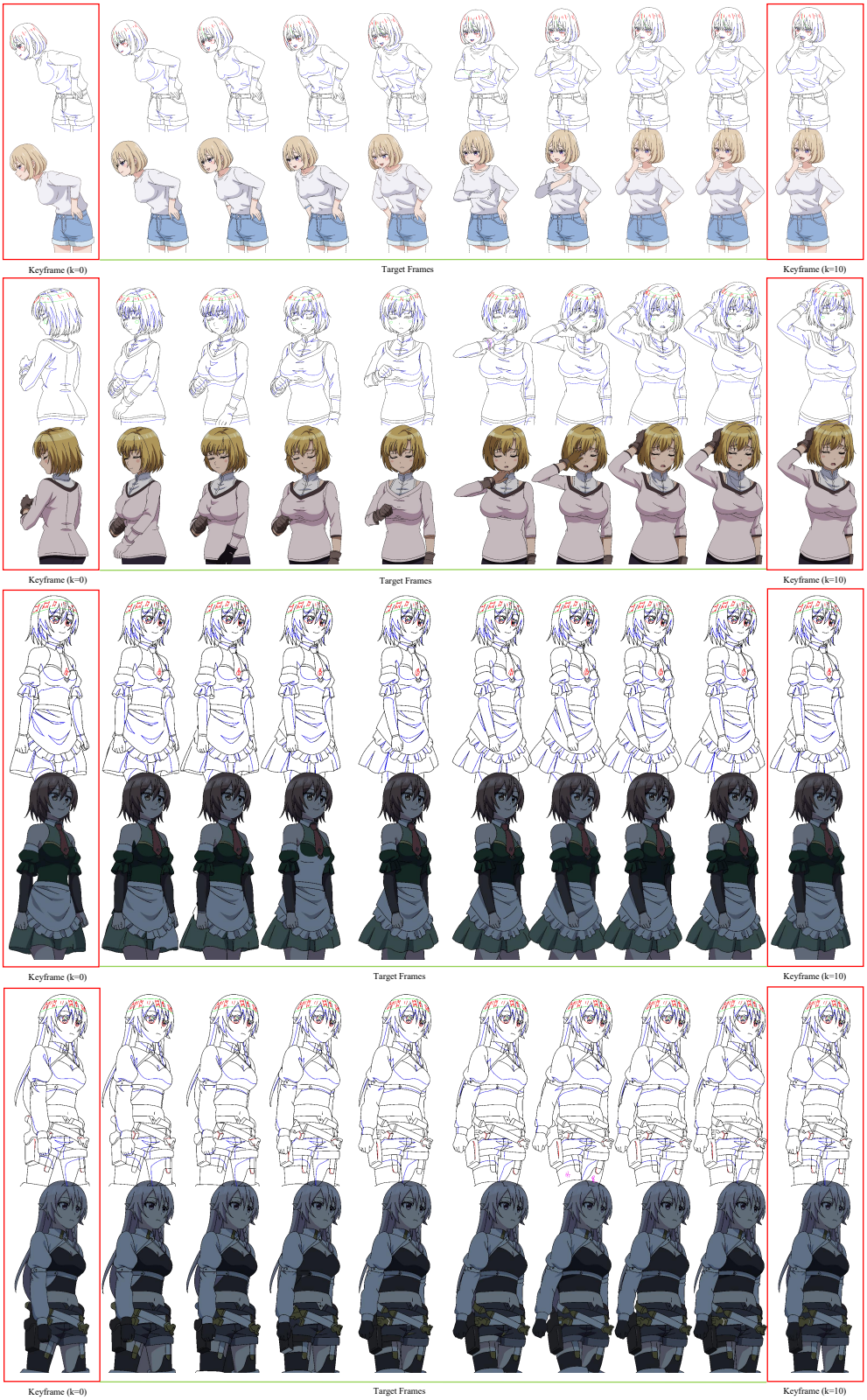

Figure 2. This figure presents the coloring results of our method on a medium shot (typically from the knees to the head). We use a sequence length of 10 for the carts, with the first and last frames as reference keyframes (highlighted in red).

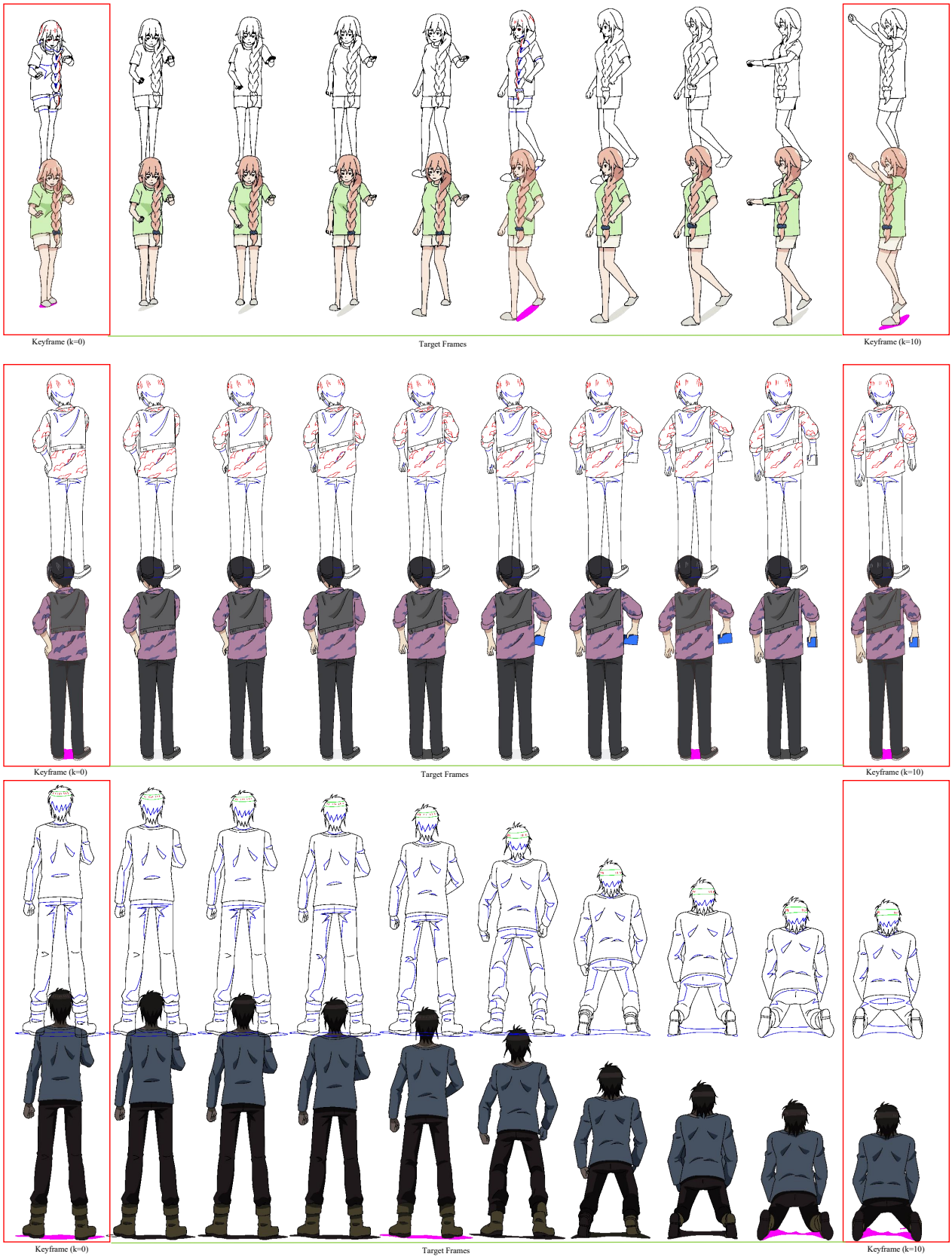

Figure 3. This figure presents the coloring results of our method on a long shot (typically showing the full body of the character). We use a sequence length of 10 for the carts, with the first and last frames as reference keyframes (highlighted in red).

Figure 4. This figure demonstrates the coloring results of our method using three keyframes for reference. We use keyframes at k=0, 5, and 10 to color the other parts. The three sets of results display the outcomes for long shot, medium shot, and close-up shot, respectively.