Slide 1: Title

Slide 2: Problem and Applications

- The k-means algorithm is a method for assigning points to a fixed number (k) of clusters
- It has applications in machine learning and machine vision
- In machine vision, we often want to segment images for object detection or background detection
- K-means is fast and effective, but you must know the number of clusters k
- Here is an example of image segmentation using k-means: the initial image is segmented into clusters, which are displayed as different shades of gray on the grayscale image.

Slide 3: K-means Algorithm

- The first step is initialization: select k initial cluster centers. In the most basic form of k-means, the k centers are randomly selected from the particles.
- The second step is iteration:
    - Assign each particle of the n particles to a center using a distance measure such as Euclidean distance
    - Re-compute the k cluster centers using the mean position of points assigned to each cluster
    - Terminate when an iteration does not change any particle assignments
- There is potential to optimize the initialization step by using a more sophisticated algorithm for assigning the k initial centers
- There is potential to parallelize parts of the iteration, such as the cluster assignments and center re-computation

Slide 4: Implementation

- We implemented a serial version of k-means image segmentation using C++ and the OpenCV open source computer vision library.
- To parallelize our implementation, we used Intel's Threading Building Blocks open source parallel library.
- We also optimized our initial cluster selection using the k-means++ algorithm, where each particle is given a probability of being selected as one of the k cluster centers proportional to its distance from the cluster centers selected so far.
- Using TBB, we converted the for loop that assigned points to clusters to a parallel_for construct, and the for loop that re-computed the cluster centers using a parallel_reduce construct

Slide 5: Results 1

Slide 6: Results 2

Slide 7: Next Steps