Similarity search in image databases relies on comparing the query with a set of images based on features like shape, colour, texture, and spatial locations. As the size of database grows, query processing strategies were proposed to increase performance by reducing the number of distance calculations. Existing strategies include brute-force sequential matching and two-step prune (or filter) and refine processing, with the initial "prune" step based on a high-dimensional spatial index followed by a "refine" step performing expensive computation. Further, many of these strategies assume underlying metric similarity models where lower bounding distance functions exist for pruning.

In this talk, existing strategies are briefly discussed. Then, recent research on approximate similarity search in a Japanese Kamon Image Database is highlighted. The choice of shapes as features is deliberate because kamons are in black and white, and their meanings are conveyed by shapes. Further, a novel three-step "prune-filter-refine" strategy targeting models with non-metric distance functions is introduced. Compared to the two-step approach, this strategy achieves a further reduction in number of distance calculations needed but with close to no change in precision.