AM 207 final project proposal

April 4, 2016

1 2D tiling

Background

This problem is motivated by applications in laser machining developed by the Harvard Microrobotics Lab. In the field of meso-scale fabrication it is advantageous to construct 3D structures out of 2D laminate materials. Composite materials such as carbon fiber are often anisotropic which makes traditional manufacturing infeasible. Consequently, we cut out 2D features and patterns from planar templates using diode pulsed laser. An illustration of this manufacturing process is shown in figure 1. The material templates have fixed size to facilitate the lamination process, which implies that multiple manufactured parts must be tilted onto a common template. Usually this tiling process is done by a human, which can be time consuming and inefficient. Here we aim to improve tiling efficiency and speed by using stochastic optimization techniques dicussed in AM207.

Problem formulation

This optimization problem can be formulated as:

$$\begin{gathered} argmax \sum_i z_i A_i \\ \text{s.t } S_i \cap S_j = \{\} \forall i, j \\ S_i \subset B \forall i \end{gathered}$$

Here we want to find the best indicator vector z_i and centroid position (x_i, y_i) for each cut feature such that we maximize the sum feature area. The constraint is that no two region can overlap and that each feature is entirely contained within the template. This is an integer programming problem and it is NP hard. Here we aim to find an approximate solution using stochastic methods. Here $z_i = \{0, 1\}$ which is the selection vector. $A_i \in R$ is the area of feature i and S_i is the set covered by feature i. Finally, B denotes the set covered by the bounding box.

This problem is derived from a real world application and is a generalization of the 1D knapsack problem we've done in homework 3. Specifically, we need to

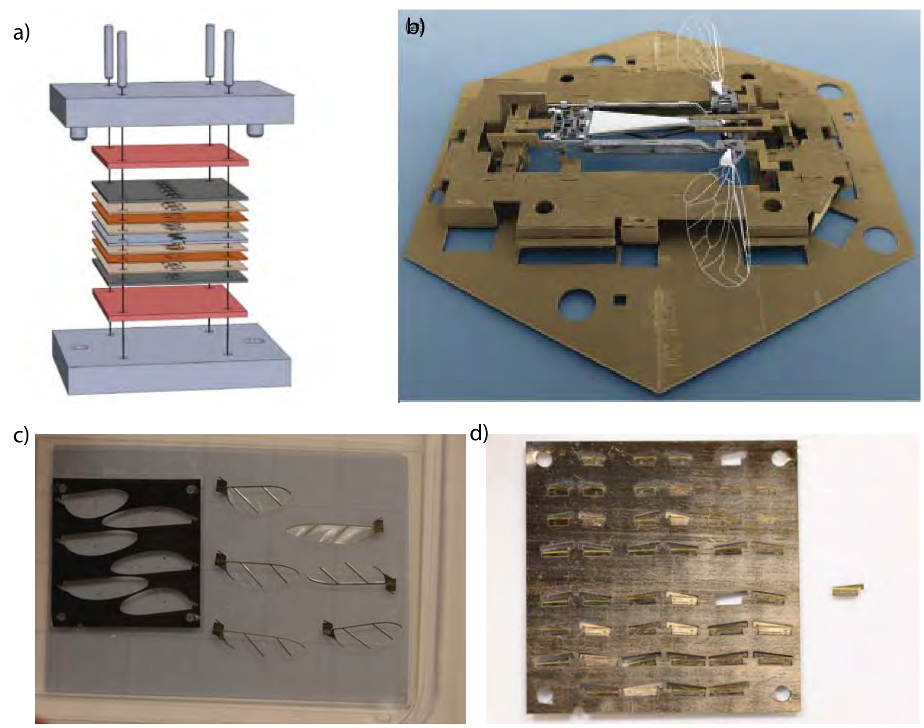

Figure 1: Illustration of laser machining process. a) The lamination process. b) A micro-robot fabricated from the process. c) Example of manually tiled wings. d) Example of manually tiled wing hinges.

choose a set of items to maximize an objective function subject to constraints. However, this is more difficult because it involves "latent variables" that specify the placement of items.

Proposed algorithm

There are a number of algorithms that we can implement to address this problem. For a base line comparison, we can implement a deterministic algorithm called strip packing. This is a greedy algorithm that can be used for base line comparison

- Strip packing

There are a class of strip packing algorithms discussed in http://cgi.csc.liv.ac.uk/~epa/surveyhtml.html. These methods have complexity of O(nlogn), which gives a greedy approximation of solution. We plan to implement several deterministic algorithms for base line comparison.

- Simulated annealing / stochastic gradient descent

Since we define an objective function, we can implement simulated annealing and use the function value as criteria for rejecting or accepting the proposal. We expect the implementation to be a two-step process. In one step, we swap old and new items out of the bag. In the next step, we optimize the arrangement of items based on the current set of items. In the second step, we can define a attractive potential function and converge to the minima by doing stochastic gradient descent. This iterative method may converge to a local solution.

- Genetic algorithm

Genetic algorithm is another stochastic optimization method we've discussed in lecture. Here we can start from a number of arrangement and configuration. Next, we combine the best scoring sets and produce descending generations. We repeat this process to converge to an approximate solution. We will compare the effectiveness of the simulated annealing and genetic algorithm.

2 Social network

Background

Here we can model a social network using a graph $G(V, E)$, where V is a set of nodes that denote the individuals consist of the network and E is a set of connections that denote the relationship between the individuals. Suppose we have a limited number of rewards that we want to distribute to a selected group such that we want to influence the most number of people. Here we can set up this problem in a probabilistic manner, which means we can split a reward between several people. Consequently, we aim to maximize the expected number of people under influence given a constraint on the number of samples we have. This model is usually used to study marketing problems.

Mathematically, the optimization problem can be formulated as:

$$argmax \sum_{i=1}^{n} 1 - \Pi_{j:(j,i)\in E}(1 - x_j)$$
$$\text{s.t. } \sum_{j=1}^{n} x_j \leq K, x_j \in [0, 1]$$

This problem is a non-convex optimization but can be approximated using a convex upper bond. Usually an approximate solution is obtained through interior point methods. Here we can use stochastic methods discussed in lectures to tackle this problem.

Proposed algorithm

- Simulated annealing

Since we define an objective function, we can implement simulated annealing and use the function value as criteria for rejecting or accepting the proposal. In each iteration we can modify the distribution of x_j under the constraint K. We can later compare the algorithm effectiveness to deterministic ones.

- Genetic algorithm

We can start with many different distributions and then merge the promising ones. This algorithm is usually good for objective functions with many local extrema. We can compare the effectiveness of this algorithm to others.

3 Image classification

Background

Here we can identify features from an image and classify images using the expectation maximization methods discussed in lectures. The expectation step involves identifying image features and the maximization step involves classifying the images accordingly to the identified features. We can compare this method to other machine learning algorithms such as othogonal matching pursuit.

Proposed algorithm

- Expectation maximization