CPS296.2 Geometric Optimization **10 April, 2007**

Lecture 24: Iterative Closest Point and Earth Mover's Distance

Guest Lecturer: Jeff Phillips *Scribe: Nihshanka Debroy*

In this lecture, we discuss the Iterative Closest Point Algorithm (ICP) and the Earth Mover's Distance. ICP is used to compute a matching that minimizes the root mean squared distance between two point-sets. The Earth Mover's Distance provides a measure of the dissimilarity between two multi-dimensional distributions.

24.1 Iterative Closest Point (ICP)

ICP is an iterative algorithm for matching point-sets [1]. Consider 2 point-sets A, $B \subseteq \mathbb{R}^d$ where $|A| = n$ and $|B| = m$. We are interested in a one-to-one matching function $\mu : A \to B$ that minimizes the root mean squared distance (RMSD) between A and B. Mathematically, we want to minimize the following:

$$RMSD(A, B, \mu) = \sqrt{\frac{1}{n} \sum_{a \in A} ||a - \mu(a)||^2} \tag{24.1}$$

Incorporating rotation and translation into the matching, we want to find:

$$\min_{\mu: A \to B, t \in \mathbb{R}^d, R \in SO(d)} \sum_{a \in A} ||Ra - t - \mu(a)||^2 \tag{24.2}$$

where R is the rotation matrix, t is the translation vector and $SO(d)$ is the set of special orthogonal matrices in d dimensions.

24.1.1 ICP Algorithm

The ICP algorithm seeks to minimize the RMSD, by alternating between a matching step and a transformation step. In the matching step, given a certain rotation and translation, the optimal matching is calculated by minimizing the RMSD. In the transformation step, given a matching, the optimal rotation and translation are computed. This alternating process terminates when the matching remains unchanged in sucessive iterations. The following is an algorithmic description of ICP:

ICP(A,B)

1. Initialize $R = I$ (the identity matrix), $t = 0$.
2. *Matching Step*: Given R and t, compute optimal μ by finding $\min_\mu RMSD(A, B, \mu)$.

3. *Transformation Step*: Given μ, compute optimal R and t by finding $\min_{R,t} RMSD(RA - t, B, \mu)$.
4. Go to step 2 unless μ is unchanged.

The matching and transformation steps are described in more detail below.

Matching Step

$\forall a \in A$, find closest $b \in B$ in the following manner:

- Construct Voronoi diagram on B.
- $\forall a \in A$, do point-location in $Vor(B)$.

If $d = 2$, we take $O(m \log m)$ pre-processing time and $O(\log m)$ time per point in A (where $m = |B|$). In practice, a k-d tree is used to find the nearest-neighbor quickly. The matching step is usually the slowest part of the algorithm.

Transformation Step

Assume $\mu(a_i) = b_i$, where $a_i \in A$ and $b_i \in B$. We want to find the matrix R and vector t that minimize:

$$\sum_{i=1}^{n} ||Ra_i - t - b_i||^2 \tag{24.3}$$

We can define the centroid of A as

$$\bar{a} = \frac{1}{|A|} \sum_i a_i \tag{24.4}$$

Similarly for B,

$$\bar{b} = \frac{1}{|B|} \sum_i b_i \tag{24.5}$$

Defining the centroids as above will help simplify the minimization problem. Let us now define the following:

$$a_i\prime = a_i - \bar{a} \tag{24.6}$$

$$b_i\prime = b_i - \bar{b} \tag{24.7}$$

$$\Rightarrow a_i = a_i\prime + \bar{a}, \quad b_i = b_i\prime + \bar{b} \tag{24.8}$$

Hence,

$$\sum_{i=1}^{n} ||Ra_i - t - b_i||^2 \tag{24.9}$$

$$= \sum_{i=1}^{n} ||R(a_i\prime + \bar{a}) - t - (b_i\prime + \bar{b})||^2 \tag{24.10}$$

$$= \sum_{i=1}^{n} ||Ra_i\prime - b_i\prime + (R\bar{a} - \bar{b} - t)||^2 \tag{24.11}$$

Now if,

$$t = R\bar{a} - \bar{b} \tag{24.12}$$

$$\Rightarrow \sum_{i=1}^{n} ||Ra_i - t - b_i||^2 = \sum_{i=1}^{n} ||Ra_i\prime - b_i\prime||^2 \tag{24.13}$$

Equation (24.11) can be expanded, using $tr(.)$ to represent the matrix trace operation and the superscript ^{T} to represent a matrix transpose. Hence,

$$\sum_{i=1}^{n} ||Ra_i\prime - b_i\prime||^2 \tag{24.14}$$

$$= RR^T \sum_{i=1}^{n} ||a_i\prime||^2 - 2tr(R\sum_{i=1}^{n} a_i\prime b_i\prime^T) + \sum_{i=1}^{n} ||b_i\prime||^2 \tag{24.15}$$

$$= \sum_{i=1}^{n} ||a_i\prime||^2 - 2tr(R\sum_{i=1}^{n} a_i\prime b_i\prime^T) + \sum_{i=1}^{n} ||b_i\prime||^2 \tag{24.16}$$

since $RR^T = I$. Let $N = \sum_{i=1}^{n} a_i\prime b_i\prime^T$. Hence, the problem of minimizing equation (24.16) reduces to that of maximing the following:

$$tr(RN) \tag{24.17}$$

Taking the singular value decomposition (SVD) of matrix N,

$$N = U\Sigma V^T \tag{24.18}$$

where U and V are orthogonal matrices and $\Sigma = \text{diag}(\sigma_1, \sigma_2, ..., \sigma_d)$ such that $\sigma_1 \geq \sigma_2 \geq ... \geq \sigma_d \geq 0$. From this, it can be proved [2] that to minimize equation (24.11) the rotation matrix should be defined as:

$$R = VU^T \tag{24.19}$$

24.1.2 ICP Convergence

The ICP algorithm is guaranteed to converge to a local minimum [1]. The algorithm terminates when the matching step does not change any of the previous matchings. A paper by Helmut Pottmann shows that ICP has linear convergence in certain cases [3]. There is an upper bound of $|A|^{|B|}$ steps for matching point-set A to point-set B, but ICP has been proven to require $\Omega(|A \cup B| \log |A \cup B|)$ iterations for certain inputs [4]. Paper [5] provides a lower bound of $\Omega((|A \cup B|/d)^{d+1})$ iterations for inputs in d dimensions.

24.2 Earth Mover's Distance

The Earth Mover's Distance (EMD) is a measure of the dissimilarity between two multi-dimensional distributions [6]. Intuitively, suppose we are given two distributions. One distribution could be seen as a "mass of earth properly spread in space," and the other as "a collection of holes in that same space" [6]. Then, the EMD measures the least amount of work that needs to done in order to fill the holes with earth. In this case, "a unit of work corresponds to transporting a unit of earth by a unit of (ground) distance" [6].

The calculation of EMD is based on a solution to the "transportation problem" (described in [7]), a bipartite network flow problem that could be cast as the following linear programming problem: Let S be a set of suppliers, C a set of consumers, and c_{ij} the cost to ship a unit of supply from $i \in S$ to $j \in C$. The goal is to find a set of flows f_{ij} that minimize the overall cost, which is:

$$\sum_{i \in S} \sum_{j \in C} c_{ij} f_{ij} \tag{24.20}$$

The flows f_{ij} are subject to the following constraints:

1. $f_{ij} \geq 0, \quad i \in S,\ j \in C$
2. $\sum_{i \in S} f_{ij} = y_j,\ j \in C$
3. $\sum_{j \in C} f_{ij} \leq x_i,\ i \in S$

where x_i is the total supply of supplier i and y_j is the total capacity of consumer j. The first constraint permits the movement of supplies only from a supplier to a consumer (and not vice-versa). The second constraint "forces the consumers to fill up all of their capacities" [6]. Constraint 3 limits the supply that can be sent by a supplier to its total amount [6]. As the total demand cannot exceed the total supply, there is a feasibility condition that

$$\sum_{j \in C} y_j \leq \sum_{i \in S} x_i \tag{24.21}$$

The transportation problem has applications in "signature matching," where one signature could be defined as the supplier and the other as the consumer. The transportation problem could then be solved, where the cost c_{ij} is the ground distance between element i in the first signature and element j in the second [6]. If the total weights of the signatures are unequal (i.e., in case of partial matches), the smaller signature could be made the consumer so that the feasibility condition is satisfied.

The transportation problem described above can be solved by finding the optimal flow on a graph. The optimal flow on a graph $G(V, E)$, where V represents the set of vertices and E the set of edges, can be calculated in time,

$$O((|E| \log |V|)(|E| + |V| \log |V|)) = O((n^2 \log n)(n^2 + n \log n)) = O(n^4 \log n) \tag{24.22}$$

Having solved the transportation problem and found the optimal flow F, the earth mover's distance is:

$$EMD(S, C) = \frac{\sum_{i \in S} \sum_{j \in C} c_{ij} f_{ij}}{\sum_{i \in S} \sum_{j \in C} f_{ij}} = \frac{\sum_{i \in S} \sum_{j \in C} c_{ij} f_{ij}}{\sum_{j \in C} y_j} \tag{24.23}$$

In equation (24.19), the denominator is a normalization factor that prevents the favoring of signatures with smaller total weights [6]. The ground distance c_{ij} is typically defined based on the specific problem. Thus, the EMD is an extension of the idea of distance between single elements to "distance between sets of elements, or distributions" [6].

References

[1] P. J. Besl, N. D. McKay, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 239-256, 1992.
[2] K. S. Arun, T.S. Huang, S.D. Blostein, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 9, pp. 698-700, 1987.
[3] H. Pottmann, Q. Huang, Y. Yang, S. Hu, Geometry and convergence analysis of algorithms for registration of 3D shapes, 2004.
[4] E. Ezra, M. Sharir, A. Efrat, *Symposium on Compuational Geometry*, pp. 95-104, 2006.
[5] D. Arthur, S. Vassilvitskii, *Symposium on Foundations of Computer Science*, pp. 153-164, 2006.
[6] Y. Rubner, C. Tomasi, L.J. Guibas, *IEEE International Conf. on Computer Vision*, pp. 59-66, 1998.
[7] G.B. Dantzig, *Activity Analysis of Production and Allocation*, pp. 359-373, 1951.
[8] Jeff M. Phillips, ICP: A User's Guide, 2006.