CS 664 Machine Vision	February 24, 2000
Lecture 10	
Lecturer: Ramin Zabih	*Scribe: Kate Jenkins*

1 Preliminaries

1.1 Administrivia

Quiz 2 will be on Tuesday.

Problem Set 1A due next Thursday.

Problem Set 1B out next week. It will cover global methods of simulated annealing and graph cut methods. We will be given a max flow library with an interface to the PRF max flow algorithm.

1.2 Two views of graph construction

Recall that Multiway Cut (MWC) is solvable by swap moves, via the following algorithm:

For each label α, β, starting with input labelling f

Find $\hat{f} = \arg\min_{f'} E(f')$, where f' is within one $\alpha\beta$ swap from f.

If $E(\hat{f}) < E(f)$ then $f = \hat{f}$

In this algorithm, we said that $f(p) = \alpha$ if α is linked to p after the cut.

An alternate view is to say that $f(p) = \alpha$ if the link $\alpha \to p$ is severed by the cut.

Either approach can work when we have only two labels. We just set the link weights differently in each case. The latter approach will NOT give you the encoding you want when dealing with more than two labels simultaneously.

For this lecture, we will be using the second approach.

2 Generalization of Swapmove to $V \neq$ Potts Model

2.1 V revisited

Recall that in global image restoration, we attempt to minimize the function

$$\sum_p C(p, f(p)) + \sum_{p,q} V(f(p), f(q))$$

where p, q are adjacent pixels.

So far, we have used the Potts function for V. This has

$$V(\alpha, \beta) = \begin{cases} 0 & \text{if } \alpha = \beta \\ 1 & \text{otherwise} \end{cases}$$

The Potts model likes a world made up of flat surfaces all facing the camera, since all differences in label values, even small ones, are equally bad. To deal with sloped surfaces, we may want to consider other energy functions where V penalizes nearby labels less than labels that are very different from each other.

(In fact, the Potts model works surprisingly well in practice, since labels are usually discretized).

We always want our energy function to be piecewise smooth.

2.2 One Swapmove iteration

Luckily for us, Swapmove can be generalized to an arbitrary choice of V.

In an iteration of Swapmove, we consider only the pixels with label α or β and we find the relabelling of these pixels (to labels α or β) that minimizes the energy. This optimal relabelling is called the best $\alpha\beta$ swap.

Given two adjacent pixels p, q, there's a penalty $V(f(p), f(q))$ associated with the difference between their labels. (Within a swap move, $f(p)$ is α or β, as is $f(q)$). There are two interesting things that these pixels can do: they can either be given the same label or be given different labels.

In the Potts model, we have:

$$\text{Penalty} = \begin{cases} 1 & \text{if } p,\ q \text{ have different labels} \\ 0 & \text{if same} \end{cases}$$

For a general V, we have:

$$\text{Penalty} = \begin{cases} V(\alpha, \beta) & \text{if different labels} \\ 0 & \text{if same} \end{cases}$$

The n-link $p \to q$ is cut $\Rightarrow$ p and q are given different labels.

Suppose that on a later iteration, labels ρ and ω are fighting over the same link $p \to q$. Unlike in the Potts model, a general V may have $V(\alpha, \beta) \neq V(\rho, \omega)$.

To solve this problem, we just update the link costs each swap move iteration, depending on what the current labels (i.e. α, β) are.

2.3 The fatal flaw

There's a fatal flaw with the Swapmove algorithm for general V as described above.

During a Swapmove iteration, you only look at the part of the graph labelled α or β. However, pixels in this subgraph may also have neighbors in the full graph that have other labels. For example, suppose $p \to q$ is an n-link, $f(p) = \alpha$, $f(q) = \gamma$, and labels α and β are being compared. In the Potts model, the cost of the link $p \to q$ is the same regardless of whether p takes on the label α or β, since all that matters is that both labels are different than q's label γ (i.e. $V(\alpha, \gamma) = V(\beta, \gamma) = 1$). However, in general it is not the case that $V(\alpha, \gamma) = V(\beta, \gamma)$ when $\alpha \neq \beta$.

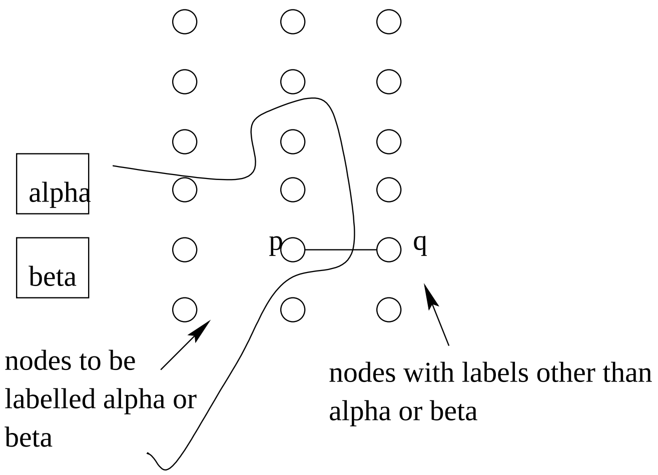

Figure 1: pixels in $\alpha\beta$ swap

We can solve this problem by updating the costs of the d-links in our subgraph to represent how unhappy outside pixels would be with each label assignment.

2.4 Constraints on C, V

This revised Swapmove algorithm works for general energy functions with cost C and penalty V, as long as a few constraints are satisfied. Namely, we need:

$$C >= 0$$

$$V(\alpha, \beta) = 0 \Leftrightarrow \alpha = \beta$$

$$V(\alpha, \beta) = V(\beta, \alpha) \text{ (can relax this constraint)}$$

Energy functions that violate these constraints show up in some other problem domains.

3 Expansion Moves

3.1 Definition and motivation

There's another kind of relabelling move that we can do that takes advantage of minimum cuts and which is very powerful. This is the expansion move.

Whereas a swap move considers the two labels α and β, an expansion move considers labels α and $\bar{\alpha}$ ($\bar{\alpha}$ denotes *not* α).

Definition 10.1 Given input labelling f, we say that f' is within a single α-expansion of f if all pixels p satisfy $f'(p) = f(p)$ or $f'(p) = \alpha$.

This means that each pixel can either remain what it is, or be labelled α.

A single expansion move can be very powerful, since it can change the labels of pixels that had had a number of different labels.

For example, suppose we have labels 1, 2, 3, 4, 5, 6 and α. Pixels with any of these labels could get labelled α within one move.

NOTE: the results of an α-expansion move are not necessarily the same as the results of doing multiple $\alpha\beta$ swaps (e.g. swap(α,1), swap(α,2), ..., swap(α,6)).

Suppose you start with a random image of N pixels and K labels. There are K^N possible labellings, but it is always possible to get to the global energy minimum in at most K expansion moves, since the i^{th} expansion move can correctly label all of the pixels that should have label α_i and there are K labels total. (The hard part is figuring out what these K expansion moves should be. This problem is NP-hard).

One common example is that you have a region where the boundary pixels have the same correct label, but you don't yet know what the interior pixel labelling should be. The interior pixels can all be assigned to the boundary label value in just one α-expansion, whereas it would take simulated annealing a very long time to find this minimum. (One can make an analogy between simulated annealing and a lake freezing).

3.2 Computing the α-expansion

Now we address the problem of quickly finding a good expansion move.

Suppose we have adjacent pixels p and q, and are using the Potts energy function.

In $\alpha\beta$ swaps:

$$p, q \text{ both same label } (\alpha \text{ or } \beta) \Rightarrow \text{cost} = 0$$
$$p, q \text{ different labels} \Rightarrow \text{cost} = 1$$

With expansion moves, things are more complicated:

$$p, q \text{ different labels } (\alpha \text{ and } \bar{\alpha}) \Rightarrow \text{cost} = 1$$
$$p, q \text{ both have label } \alpha \Rightarrow \text{cost} = 0$$
$$p, q \text{ both have label } \bar{\alpha} \Rightarrow \text{cost}=?$$

If p and q both have label $\bar{\alpha}$ we need to distinguish between the case where they have the same label in f, and where they have different labels in f, since this will affect the cost.

We need to modify the graph so that adjacent pixels in the current labelling f that are different both want to become α. We use the following construction:

For p, q adjacent with $f(p) \neq f(q)$, add a new node x in the middle of the n-link $p \to q$. Add a d-link $x \to \bar{\alpha}$. So $p \to q$ is replaced by three links $p \to x$, $x \to q$, and $x \to \bar{\alpha}$.

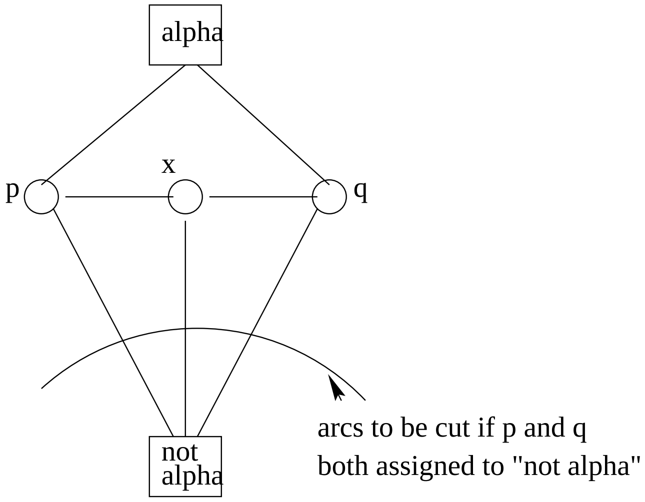

Figure 2: modified graph for $\bar{\alpha}$ and general V

If both p and q are going to be assigned to $\bar{\alpha}$, then link $x \to \bar{\alpha}$ will need to be cut, along with links $p \to \bar{\alpha}$ and $q \to \bar{\alpha}$. (Remember we assign a pixel to a label if the d-link between the pixel and that label is cut, since we are using the alternate construction today). Hence we can penalize assigning both p and q to $\bar{\alpha}$ by giving the link $x \to \bar{\alpha}$ cost=1.

This approach can be generalized to other V by updating the graph depending on which label α is being considered.

NOTE: There is a subtlety when considering expansion moves with a general V. We now require that V be a metric (hence satisfies the triangle inequality). The Potts model is a metric.

3.3 Optimality guarantee for Potts model

One can prove that with the Swapmove algorithm, the local minimum that is found can be far away from the global minimum. This is not the case with expansion moves.

Theorem 10.1 Let $\hat{f}$ be the local minimum computed by repeatedly applying expansion moves until convergence. Let f^* be the global minimum. Then $E(\hat{f}) \leq 2E(f^*)$.

Proof:

Let A be the region of f^* that has label α. This same region may have many different labels in $\hat{f}$. Define a new labelling f^α by

$$f^\alpha(p) = \begin{cases} \alpha & \text{if } p \in A \\ \hat{f}(p) & \text{if } p \notin A \end{cases}$$

Then f^α is within one α-expansion of $\hat{f}$. Since $\hat{f}$ is a local minimum, $E(\hat{f}) \leq E(f^\alpha)$.

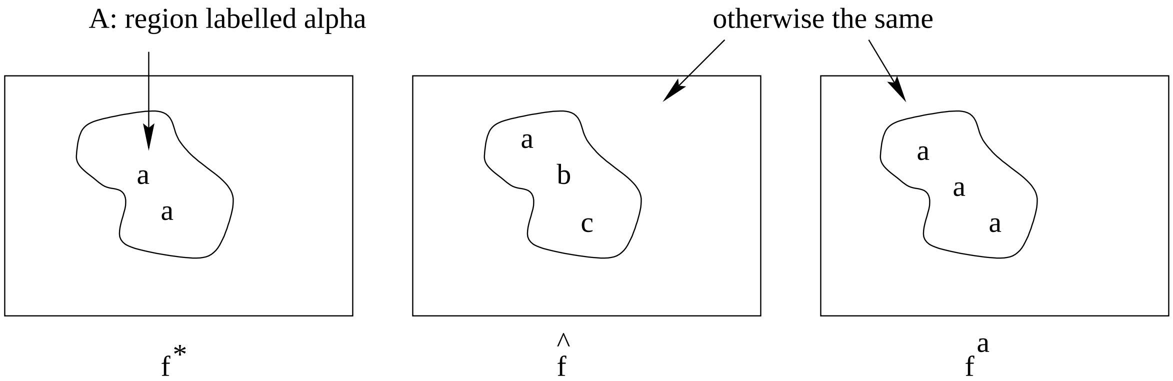

Figure 3: f^*, $\hat{f}$, and f^α

Let $E_A(f)$ denote the energy contribution from pixels in A. i.e.

$$E_A(f) = \sum_{p \in A} C(p, f(p)) + \sum_{p \in A, q \in A \cup \delta(A)} T[f(p) \neq f(q)]$$

where the second sum is over pixels $p \in A$ and adjacent pixels q. $\delta(A)$ denotes the border of A.

Then

$$E_A(\hat{f}) \leq E_A(f^\alpha) \leq E_A(f^*)$$

This last inequality holds because both f^α and f^* agree on A, but all connections between A and $\delta(A)$ are broken in f^* (since no border pixels have label α in f^*).

This inequality holds for an arbitrary label α (and corresponding set A). Summing over all α, we get:

$$E(\hat{f}) \leq E(f^*) + \text{ number of discontinuities}(f^*)$$

so

$$E(\hat{f}) \leq 2E(f^*)$$

for the Potts model.

$\square$

In fact the bound is much stronger than a 2-approximation, since evidence suggests that in the global minimum, most of the energy comes from the data terms, *not* the discontinuities. A guess is around 10% of the energy comes from discontinuities.

For an arbitrary metric V, define $C = \frac{\max V(\alpha,\beta)}{\min_{\alpha\neq\beta} V(\alpha,\beta)}$. Then one can prove the more general result:

$$E(\hat{f}) \leq 2CE(f^*)$$