Assignment 25/04/2018

Stereo Matching with Loopy Belief Propagation

Hard Deadline: 09/05/2018

Introduction

Stereo matching is the problem of finding corresponding pixels, and thus their depth, in a left-right stereo image pair. These input images are rectified so that epipolar lines are scan lines, i.e., the location of a pixel in the left and right image is shifted along horizontal axis, compared to one another, depending on its distance from the cameras. This offset is called disparity. For each image pair, you will ultimately seek to output a disparity map for the left image indicating the disparity to the right image. Figure 1 illustrates the stereo matching problem; for each pixel the optimal disparity d is to be estimated, where $d \in \{0, 1, 2, ..., d_{max}\}$.

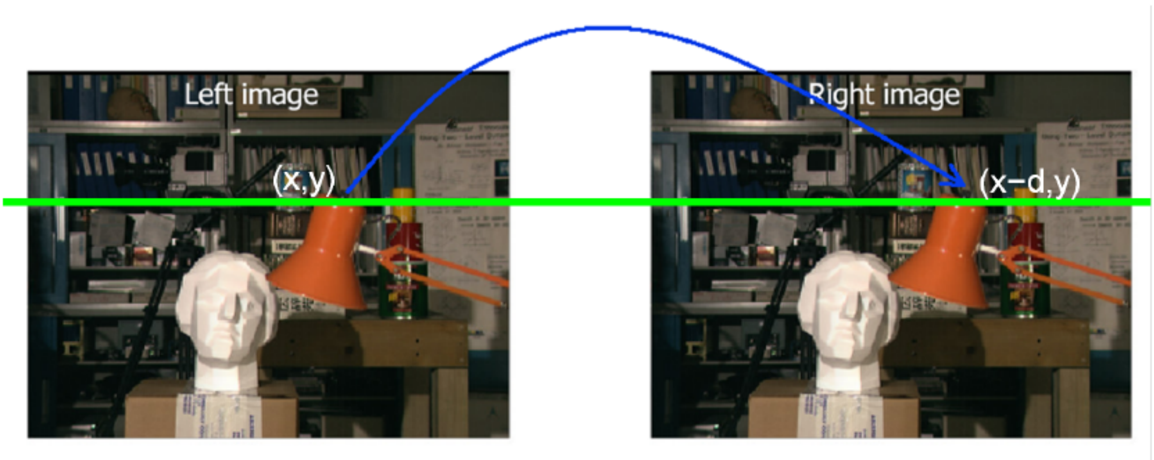

Figure 1: Stereo Correspondence: for each pixel $p = (x, y)$ of the left image, find the optimal shift $d(x, y)$ (an integer, its disparity) so that pixel $(x - d, y)$ in right image matches p.

The problem of recovering an accurate disparity map $\mathcal{L}$ can be posed as an energy minimization problem:

$$\mathcal{E}(\mathcal{L}) = \sum_{p \in \mathcal{P}} \mathcal{D}_p(l_p) + \lambda \cdot \sum_{p,q \in \mathcal{N}} \mathcal{V}(l_p - l_q)$$

where $l_p \in \{0, 1, 2, ..., d_{max}\}$ is the disparity value of pixel p, $\mathcal{P}$ is the set of pixels in the image, $\mathcal{N}$ is the set of undirected edges in the four-connected image grid graph, $\mathcal{D}_p(l_p)$ is the cost of assigning label-disparity l_p to pixel p, and $\mathcal{V}(l_p - l_q)$ is the cost of assigning labels l_p and l_q to two neighbor pixels.

In this assignment, as data cost $\mathcal{D}_p(l_p)$ you will use the truncated absolute intensity difference:

$$\mathcal{D}_p(l_p) = \min(|I_{left}(y, x) - I_{right}(y, x - l_p)|, \tau)$$

and as measure of difference between labels $\mathcal{V}(l_p - l_q)$ you will use the Potts model:

$$\mathcal{V}(x) = \begin{cases} 0 & \text{if } x = 0 \\ 1 & \text{otherwise} \end{cases}$$

For the parameter τ you will use the values 15 and for d_{max} you will use the value 15.

Finding the disparity map $\mathcal{L}$ that minimizes the above energy corresponds to maximum a posteriori (MAP) estimation problem for an appropriately defined markov random field (MRF). In this assignment you will use the min-sum Loopy Belief Propagation (LBP) algorithm that approximates a MAP solution.

Question 1

LBP is an iterative algorithm. At each iteration t, every node p of the graph sends messages to every other neighboring node $q \in \mathcal{N}(p)$, where $\mathcal{N}(p)$ are the four neighbors of pixel p: the left, right, up, and down pixels. Let $m^{t-1}_{p \to q}(l_q)$ for $\forall p \in \mathcal{P}$, $\forall q \in \mathcal{N}(p)$ and $l_p = \{0, ..., d_{max}\}$ be the messages that were exchanged at iteration $t-1$. Then:

i What is the formula for computing the vector of messages that node p will send to node q at iteration t if a general pair-wise cost $\mathcal{V}$ is used?

ii Explain what is the complexity of computing this vector.

iii What is the formula for obtaining the beliefs $b^t_q(l_q)$ of node q at iteration t and how do we get the MAP labels from them?

Question 2

If we use the Potts model as $\mathcal{V}$, the messages can be computed more efficiently. Derive a formula and explain how we can compute the vector of messages that node p sends to node q in a much faster way than the general case. What is the computational complexity of computing the vector of messages from p to q in this case?

Question 3

For this question you will have to implement the min-sum LBP algorithm that estimates the disparity map given a stereo image correspondence like the one in Figure 1. Use the message passing formula of Question 2. The functions that you will have to implement are:

comp_data_cost.m Computes the data cost term:

$$\mathcal{D}_p(l_p) = \min(|I_{left}(y, x) - I_{right}(y, x - l_p)|, \tau)$$

update_message.m Given the data cost term $\mathcal{D}_p(l_p)$ and the messages at iteration $t-1$, it computes the new messages $m^t_{p \to q}(l_q)$. For efficient implementation use the message passing formula of Question 2. For convenience you may not compute the messages that are sent from the pixels on the boundaries of the image.

comp_beliefs.m Computes the beliefs $b_q^t(l_q)$ for each node q at iteration t

comp_MAP_labeling.m Computes the MAP labeling $\mathcal{L}$ at iteration t based on the above beliefs

comp_energy.m Computes the global energy of the labeling $\mathcal{L}$ at iteration t

$$\mathcal{E}(\mathcal{L}) = \sum_{p \in \mathcal{P}} \mathcal{D}_p(l_p) + \lambda \cdot \sum_{p,q \in \mathcal{N}} \mathcal{V}(l_p - l_q)$$

Study the master function:

stereo_belief_propagation.m Is the code that calls the above functions at appropriate places and with the right arguments.

Question 4

Inside the **stereo_belief_propagation.m** function there is a call to a function named **normalize_messages.m**, which is already implemented. What is the purpose of the function? Theoretically, does it affect the labeling that is obtained at each iteration (and why)?

Question 5

Estimate with the Loopy Belief Propagation algorithm the disparity map for the following values of the parameter λ: 1, 10, and 1000. Report in each case:

i The obtained disparity map.

ii A plot of the energy at each iteration.

What do you observe regarding how the parameter λ affects the obtained disparity map?

Guidelines

Each student should submit a single zip file that will contain all the required code as well as a PDF document with the responses to the questions. For queries contact monasse@imagine.enpc.fr.