Correspondence Problem

Václav Hlaváč

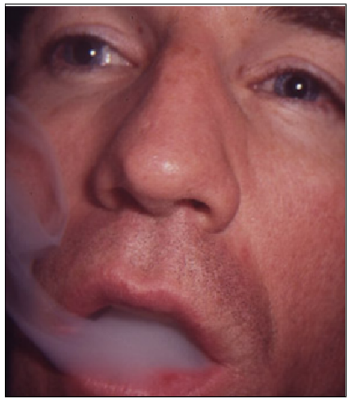
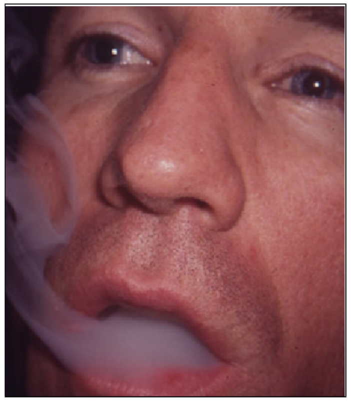
Czech Technical University in Prague
Faculty of Electrical Engineering, Department of Cybernetics
Center for Machine Perception
Czech Republic

http://cmp.felk.cvut.cz/~hlavac, hlavac@fel.cvut.cz

Problem formulation and introduction

Given: A pair of images of the same 3D scene
with known/unknown epipolar geometry.

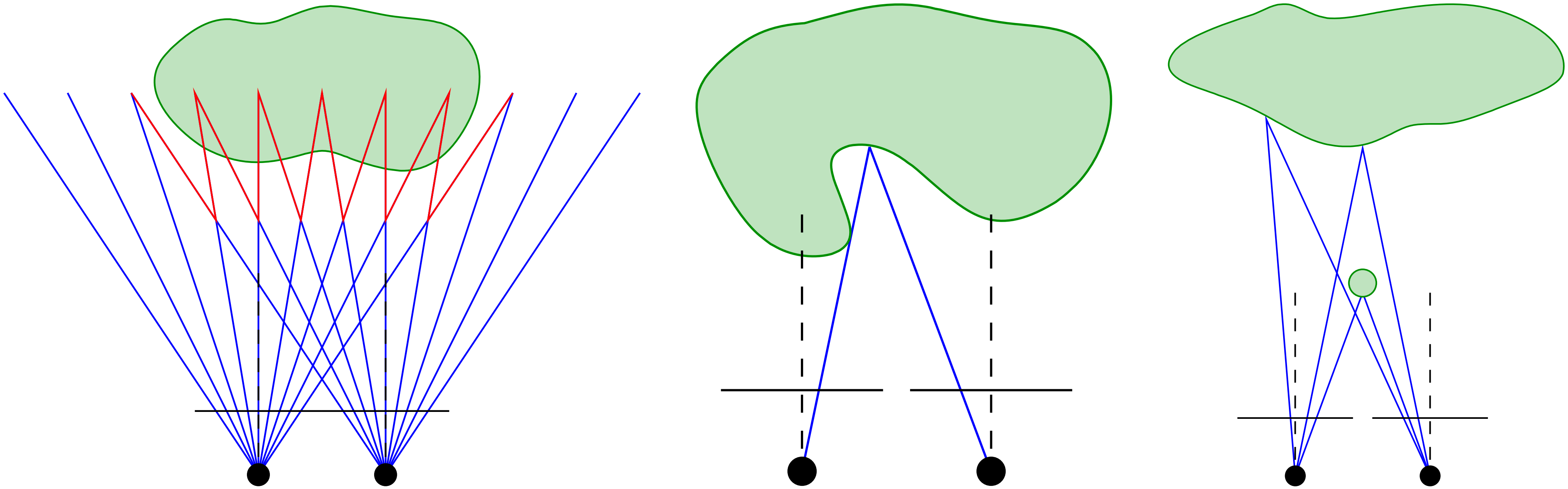
Find: A (maximal) set of pairs of corresponding points.

Troubles:

- The task is inherently ambiguous.
- Uniform, non-textured regions.
- Self-occlusions.
- Broken ordering constraint

Inherent troubles



Ambiguity Occlusion Broken order

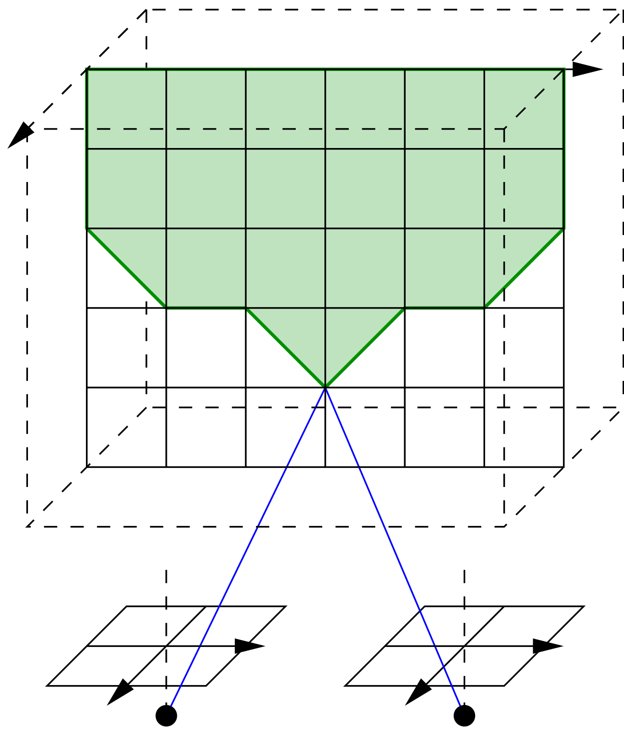
The pessimistic conclusion is that the task is not solvable in the general case at all!

How to overcome troubles?

- Humans use the high level visual semantics to disambiguate the task.
- What about computers? Model first!
  - One smooth (object) surface and many cameras $\Rightarrow$
    use a voxel representation and space carving (Kutulakos, Seitz, 2000).
  - If only a small number of corresponding points is needed $\Rightarrow$
    use special point detectors and SIFT descriptors or MSER-s. (Matas, 2002)
  - Two cameras, one smooth surface which is almost everywhere binocularly visible $\Rightarrow$
    use dynamic programming, energy minimization, ... (Schlesinger, Flach, 2004)
  - ...

Notation

- Left image: $f_l(u_l)$.
- Right image: $f_r(u_r)$, where u_l, u_r, $\in \mathbb{Z}^2$.
- Let $\mathcal{U} \subset \mathbb{Z}^2$ denote the discrete 2D-domain of the depth map.
- Depth map: $h\colon \mathcal{U} \to K$, where $K \subset \mathbb{Z}$ denotes the (discrete) depth range.



A smooth surface seen by two cameras

- One smooth surface (almost) everywhere seen by two cameras.
- Suppose (for simplicity) that the cameras are fully calibrated.
- Each discrete 3D-point (u, k) is mapped onto a pair of corresponding image points (u_l, u_r).
- We need a similarity measure in order to evaluate the quality of such a pair.

The similarity measure

$$S(u_l, u_r) = \|f_l(u_l) - f_r(u_r)\|^2$$

$$S(u_l, u_r) = \sum_{v \in W} \|f_l(u_l + v) - f_r(u_r + v)\|^2$$

$$S(u_l, u_r) = \min_{c} \sum_{v \in W} \|f_l(u_l + v) + c - f_r(u_r + v)\|^2$$

$$S(u_l, u_r) = \min_{c,b} \sum_{v \in W} \|b \cdot f_l(u_l + v) + c - f_r(u_r + v)\|^2$$

$$S(u_l, u_r) = \min_{c,b,T} \sum_{v \in W} \|b \cdot f_l(u_l + v) + c - f_r\big(T(u_r + v)\big)\|^2$$

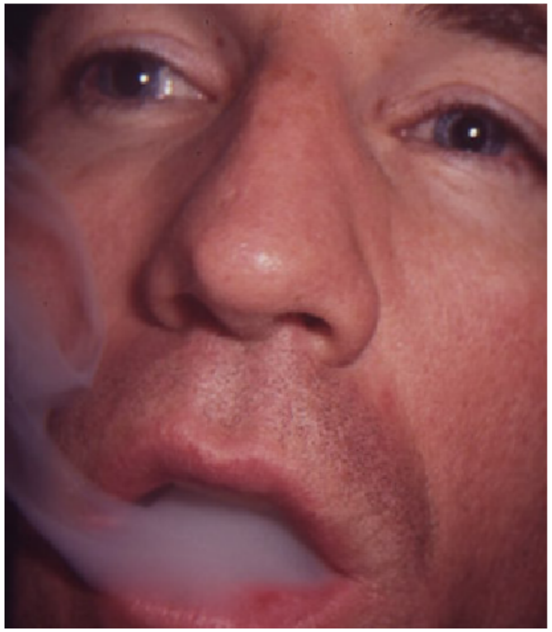
For each 3D-position $(u, k) \mapsto (u_l, u_r)$ we have a quality $q(u, k) = S(u_l, u_r)$

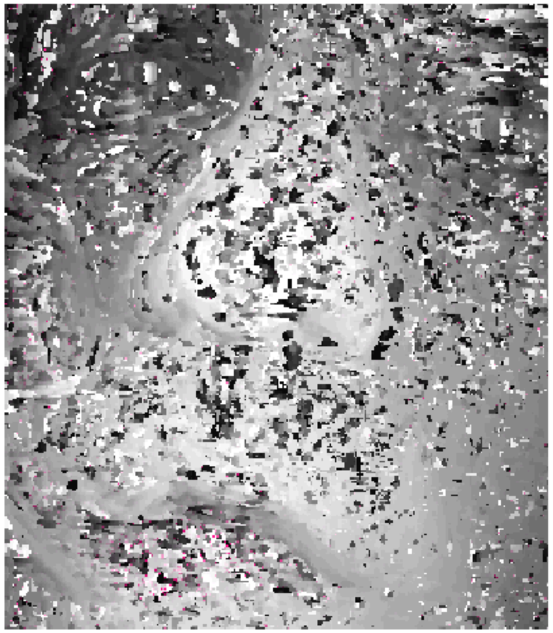
Block matching

Using the simple decision

$$h(u) = \arg\min_{k \in K} q(u, k)$$

gives:





We have not yet modeled the **smoothness assumption**!

Line by line methods

- Consider one row of the depth profile

$$h_i = h(u_1 = i, u_2)$$

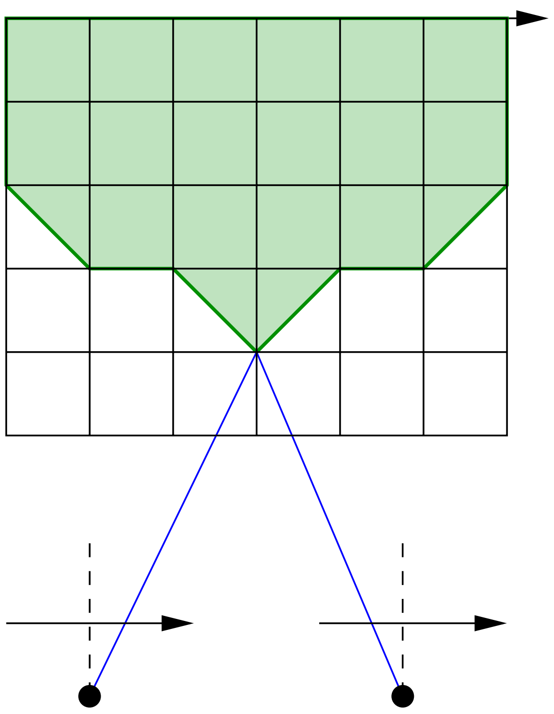
- Introduce penalties for big depth jumps, e.g.

$$g(h_i, h_{i+1}) \sim (h_i - h_{i+1})^2$$

- Consider the optimisation problem for $h = (h_1, \ldots, h_n)$

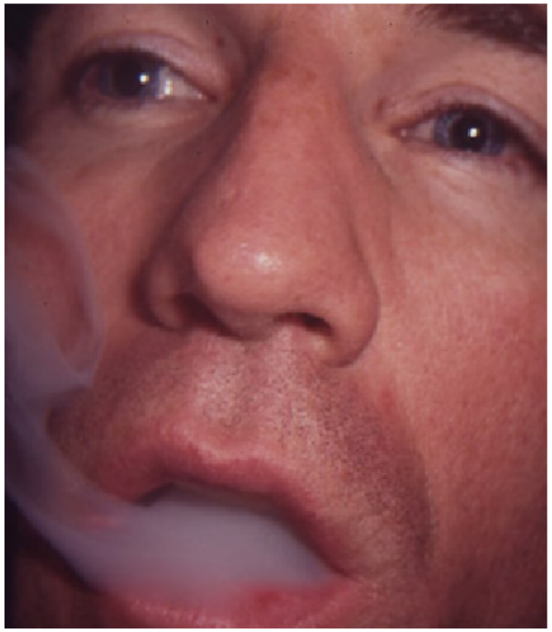
$$h^* = \arg\min_h \Big[\sum_{i=1}^{n} q_i(h_i) + \sum_{i=1}^{n-1} g(h_i, h_{i+1}) \Big]$$

How to solve it?

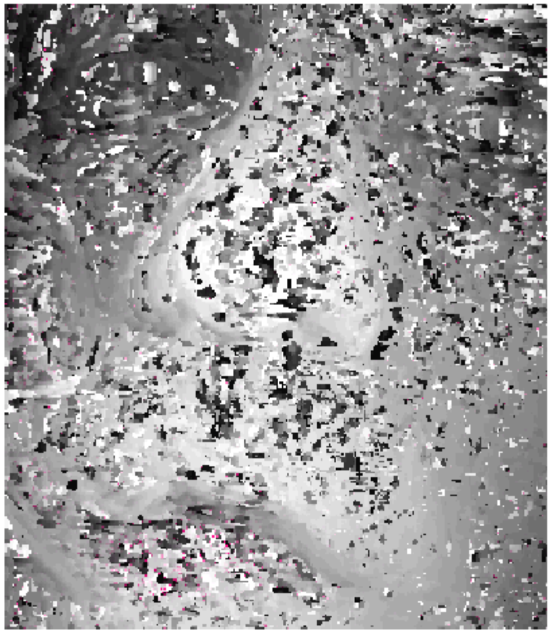


Line by line methods

By dynamic programming!



Image



Block matching

Dynamic programming

2D-optimization, from the chain to the lattice

- Nodes $u \in \mathcal{U}$ connect the nearest neighbors in the lattice by edges. Let us denote the set of edges by E.
- Define penalties $g\colon K \times K \to \mathbb{R}$ for depth jumps on edges, e.g.,

$$g(k,k') = \begin{cases} 0 & \text{if } |k-k'| \le \delta \\ +\infty & \text{otherwise} \end{cases}$$

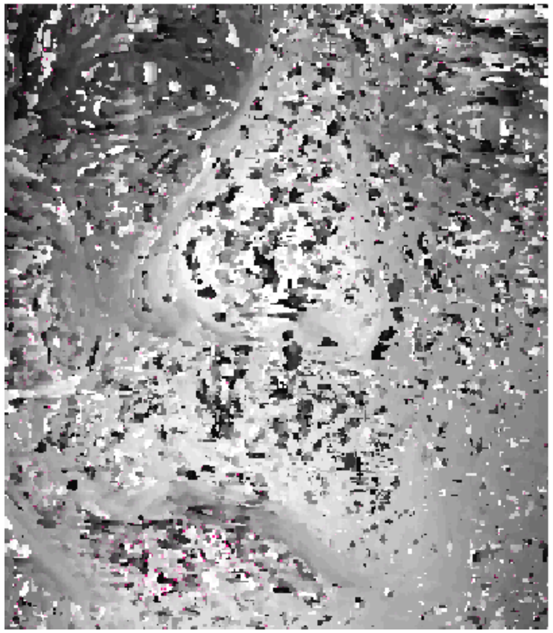
$$g(k,k') = c \cdot (k-k')^2$$

- The optimization task reads

$$h^* = \arg\min_h \Big[\sum_{u \in \mathcal{U}} q_u\big(h(u)\big) + \sum_{(u,u') \in E} g\big(h(u), h(u')\big) \Big]$$

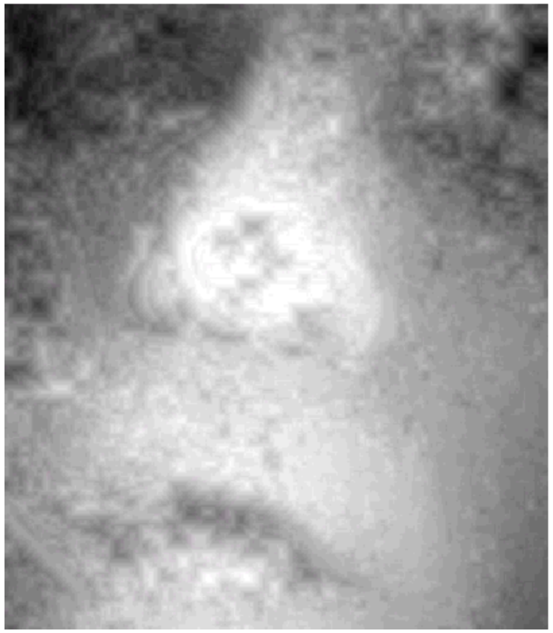
- (Min,+)-Task for graph labeling (a.k.a. Energy Minimization). These tasks are NP-complete in general. They can be solved in polynomial time if the edge functions are submodular.

2D-Optimization, example



Block matching

Dynamic programming



(Min,+) solution

A probabilistic model

Define the joint probability to observe images f_l, f_r and the depth map h by

$$p(f_l, f_r, h) = \frac{1}{Z} \exp\Big[- \sum_{u \in \mathcal{U}} q_u\big(h(u)\big) - \sum_{(u,u') \in E} g\big(h(u), h(u')\big) \Big]$$

The Bayes decision for h depends on the chosen loss function

$$h^* = \arg\min_h \sum_{h'} p(h' | f_l . f_r) C(h', h)$$

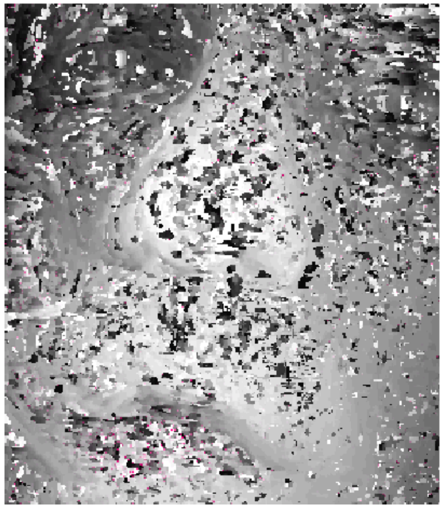
a) The loss function $C(h', h) = \mathbb{1}\{h' \neq h\}$ gives the (Min.+) problem

b) The loss function $C(h', h) = \sum_{u \in \mathcal{U}} \big(h(u) - h'(u)\big)^2$ leads to a different decision:

$$h^*(u) = \sum_{k \in K} k \cdot p\big(h(u) = k | f_l, f_r\big)$$

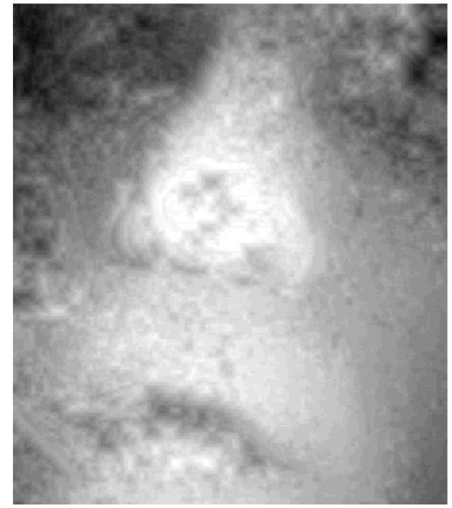
The needed marginal a-posteriori probabilities $p\big(h(u) = k | f_l, f_r\big)$ can be calculated approximately by a Gibbs-Sampler

The probabilistic model examples



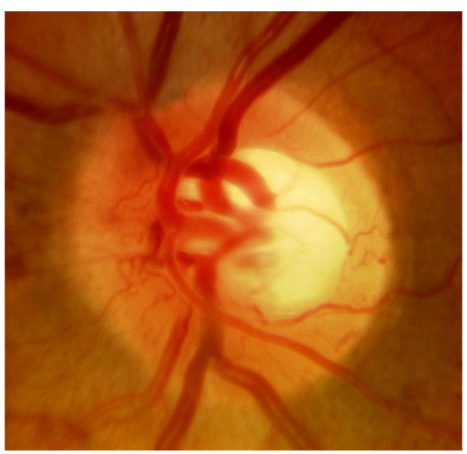
Block matching

Dynamic program.

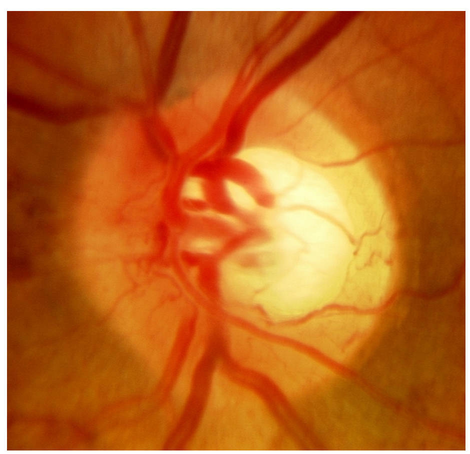


(Min,+) solution

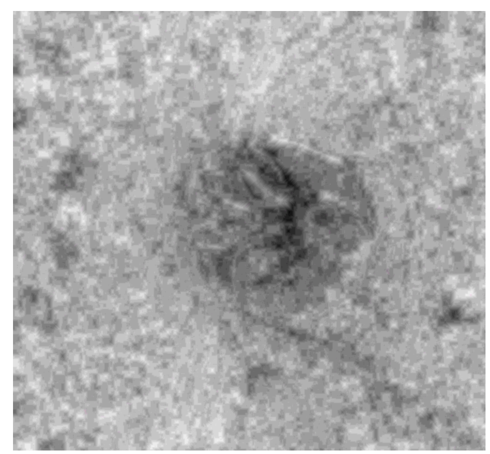
Locally additive loss



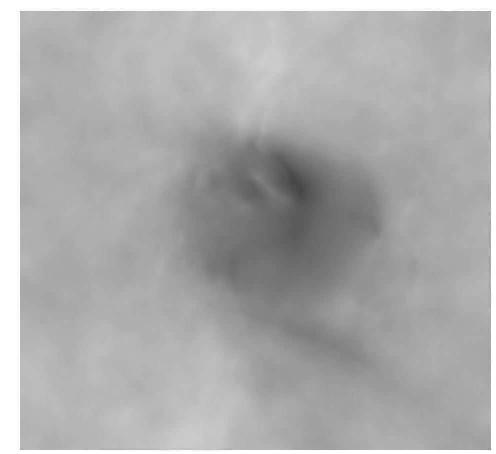
Right image



Left image



(Min,+) solution



Locally additive loss