Combinatorial Optimization in Computer Vision

Introduction

WS 2011/12

Ulrich Schlickewei
Computer Vision and Pattern Recognition Group
Technische Universität München

Plan for Today

1. Formalities
2. Examples of Combinatorial Optimization Problems in Computer Vision
3. Outline of the Lecture

Dates

- **Lecture**
  - Tuesday 14.15 – 15.45, room 02.09.023
  - Course
- **Tutorial**
  - approx. every other week
  - time tbd.
  - precise dates will be announced on the course web page

October 2011

Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
						1
2	3	4	5	6	7	8
9	10	11	12	13	14	15
16	17	18 Lecture 1	19 Lecture 2	20	21	22
23	24	25 no lecture	26 no tutorial	27	28	29
30	31					

November 2011

Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
		1 All Saints	2 no tutorial	3	4	5
6	7	8 Lecture 3	9 no tutorial	10	11	12
13	14	15 Student assembly	16 Lecture 4	17	18	19
20	21	22 Lecture 5	23 Tutorial	24	25	26
27	28	29 Lecture 6	30 Tutorial			

Graduation

- **Requirements for being admitted to the exam**
  - regular participation in lecture and tutorial
  - signs of activity in the tutorials
- **Exam**
  - oral or written exam
  - more information in the course of the semester

Slides

- slides are available on the course page
  - **user name**: tum
  - **password:** combinatorics
- let me know any typos, mistakes etc.

Course Page

- For further information (slides, tutorial dates, exam dates etc.) see http://cvpr.in.tum.de/teaching/ws2011/cocv2011
- Don‘t hesitate to contact me ulrich.schlickewei@in.tum.de

Plan for Today

1. Formalities
2. Examples of Combinatorial Optimization Problems in Computer Vision
3. Outline of the Lecture

Examples

a) Image Segmentation
b) Shape Matching

Image Segmentation

Decompose an image into foreground (lion) and background (forest)

image courtesy of C. Nieuwenhuis

Segmentation Approach: Markov Random Fields

1. Probability distribution on the set of all possible segmentations, based on
   - foreground-background probability distribution
   - prior-knowledge on objects‘ boundary
2. Optimization problem

Reminder: Random Variables I

- A **Random Variable** is a function

$$X : \mathcal{S} \to \Lambda$$

which assigns unique numerical values to all possible outcomes of a random experiment under fixed conditions.

(Mathematically, X is a measurable map from a probability space $\mathcal{S}$ to a measurable space Λ.)

- Example: throw the dice twice.

$\mathcal{S}$ = {(1,1),(1,2),...,(6,6)}, Λ = {2,...,12}

X = sum of the two results

Reminder: Random Variables II

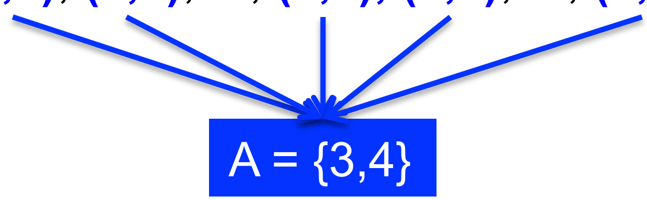
- A random variable $X : \mathcal{S} \rightarrow \Lambda$ induces a probability distribution P_X on Λ by setting

$$P_X(A) = P_{\mathcal{S}}(X^{-1}(A))$$

where $A \subset \Lambda$ and $P_{\mathcal{S}}$ is the probability distribution on $\mathcal{S}$.

- Example: throw a dice twice.

$\mathcal{S}$ = {(1,1), **(1,2)**, **(1,3)**, ..., **(2,1)**, **(2,2)**, ..., **(3,1)**, ..., (6,6)}



$$P_X(A) = P_{\mathcal{S}}(X^{-1}(A)) = \frac{|X^{-1}(A)|}{|\mathcal{S}|} = \frac{5}{36}$$

Reminder: Random Variables III

- Typically one is interested only in the distribution of a random variable and not its domain $\mathcal{S}$ nor in the actual map X.
- In our setup we assume that we took a photo of a real scene and we want to distinguish between the foreground and the background of the scene. In this setup we have
  - the space $\mathcal{S}$ of all possible scenes which could have led to the image measured.
  - a random variable $X_{i,j}$ for each pixel (i,j), $1 \leq i \leq n$, $1 \leq j \leq m$, with

$$X_{ij} = \begin{cases} 1 \text{ if } (i,j) \text{ is in the foreground} \\ 0 \text{ if } (i,j) \text{ is in the background} \end{cases}$$

  - the space $\Lambda = \{0,1\}$ of possible pixel labels

Modelling a Distribution for Image Segmentation

- In the following we will present a model for the distribution of the random field $X_{ij} : \mathcal{S} \rightarrow \Lambda$.
- For this, we need the notion of Markov Random Fields (MRFs).
- Recall that two random variables X_1, X_2 are independent (write $X_1 \sqcup X_2$) if the joint probability distribution decomposes in a product

$$p_{X_1,X_2}(A \times B) = p_{X_1}(A) \cdot p_{X_2}(B)$$

(intuitively, this means that X_1 and X_2 do not influence each other)

Markov Random Fields I

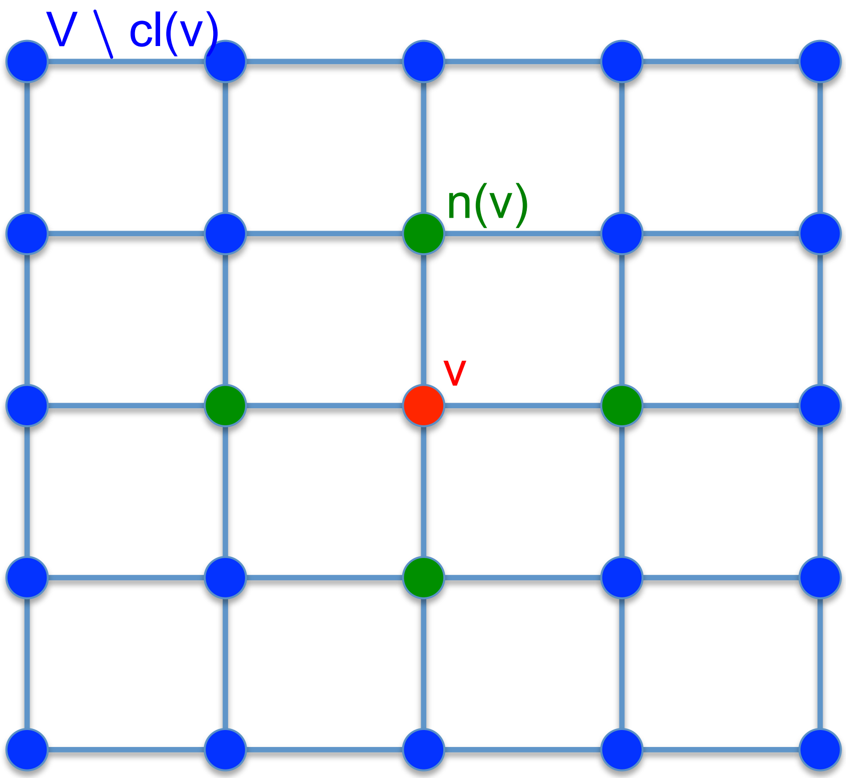
Definition: Let G = (V,E) be a graph. A Markov random field (MRF) indexed over G is a family of random variables $X = (X_v : \mathcal{S} \to \Lambda)_{v \in V}$ such that the following condition is fulfilled:

(Local Markov Property) A variable X_v is conditionally independent of all other variables if its neighborhood is given:

$$\left(X_v \bigsqcup X_{V \setminus \mathrm{cl}(v)}\right) \Big| X_{\mathrm{n}(v)}$$

Here, n(v) is the set of vertices which are connected to v by an edge and cl(v) = n(v) ∪ {v} is the closed neighborhood of v.

Markov Random Fields II



Consider a

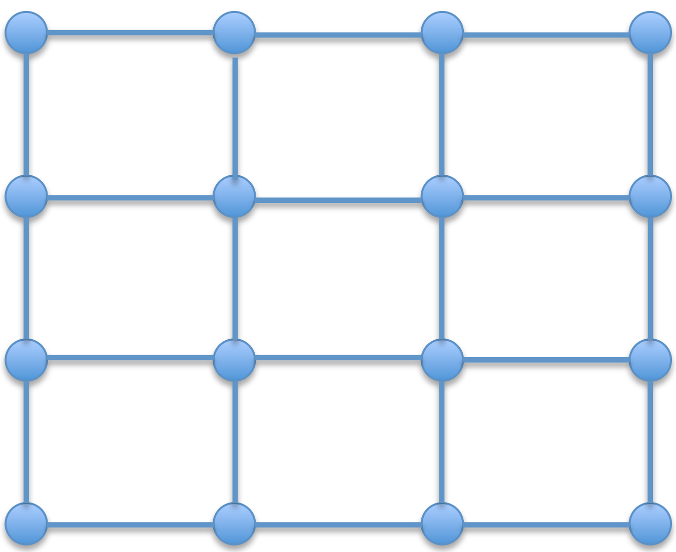
- vertex v
- its neighborhood $n(v)$
- complementary vertices $V \setminus cl(v)$.

Assume that $X_{n(v)}$ is fixed. Then the value of X_v only depends on $X_{n(v)}$ and it is independent of $X_{V \setminus cl(v)}$.

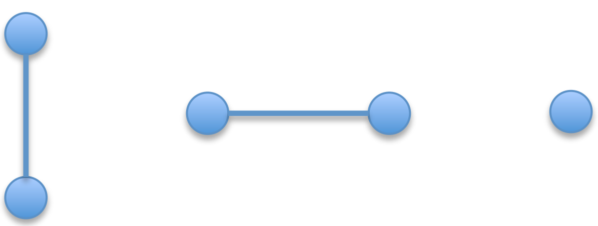
The Markov property of a random field $X = (X_v)$ has the following meaning:

Cliques

Definition: Let G = (V,E) be a graph (V = set of vertices, E = set of edges). A clique is a subset $C \subset V$ such that for all vertices $v,w \in C$ are connected by an edge $(v,w) \in E$.



In the example on the left, there are three types of cliques:



Factorization over Cliques

Theorem (Hammersley-Clifford):

Let X = ($X_v : \mathcal{S} \to \Lambda)_{v \in \mathcal{V}}$ be a random field over a graph G = (V,E). Assume that each possible value $x \in \Lambda^{|\mathcal{V}|}$ has positive probability.

Then the following are equivalent:

- X is a Markov random field.
- For each each possible value $x \in \Lambda^{|\mathcal{V}|}$ we have the following factorization property

$$p_X(x) = \prod_{\text{C clique in } G} p_{X_C}(x_C).$$

(Here, p is the probability distribution on Λ induced by X and x_C is the projection of x onto $\Lambda^{|C|}$.)

Gibbs Potential

- Under the hypothesis that every possible value x has positive probability there exists a function

  E : $\Lambda^{|V|} \rightarrow$ **R** with

$$p_X(x) = \frac{1}{Z} \exp(-E(x)) \text{ for all } x \in \Lambda^{|V|}.$$

  Here, Z = $\sum_x$ exp(-E(x)) so that all probabilies sum up to 1.

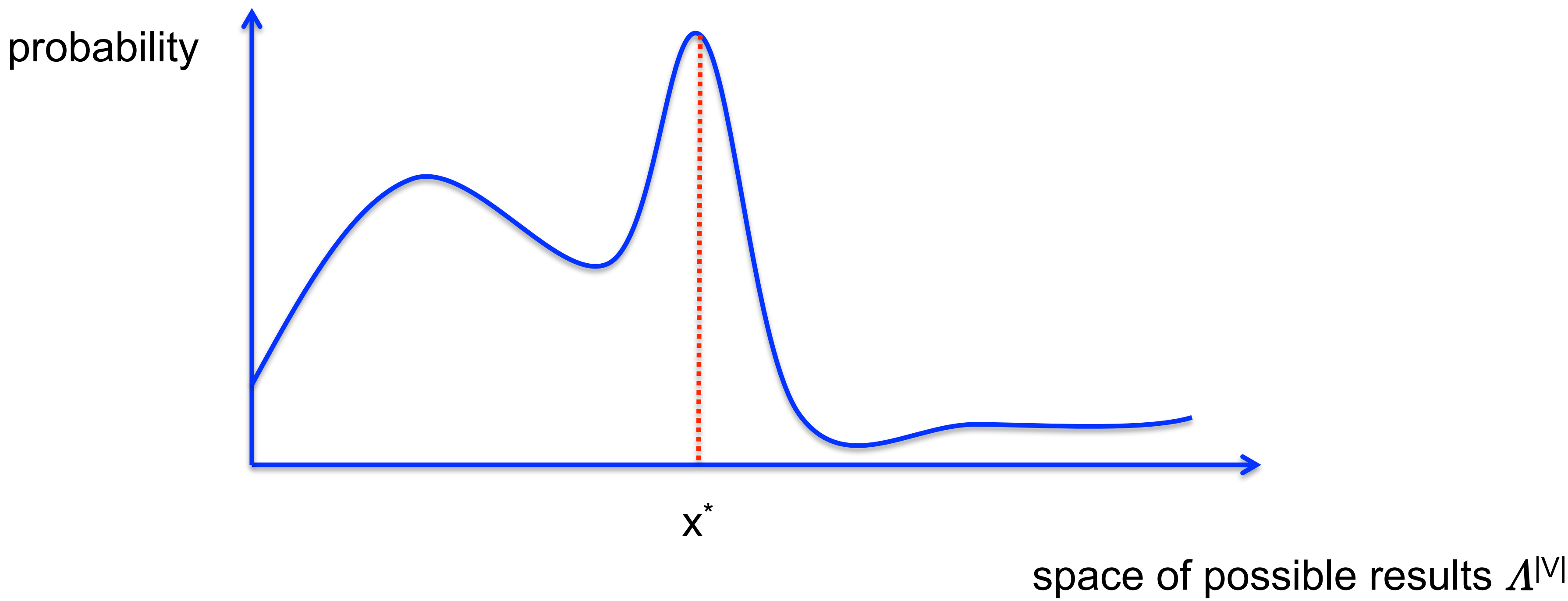
- If now X has in addition the Markov property, then E can be written as

$$E(x) = \sum_{C \text{ clique in } G} E_C(x_C).$$

- Such a potential is called a Gibbs potential.

MAP Inference for MRFs

Problem: Want to find the result $x^* \in \Lambda^{|V|}$ with the highest probability.



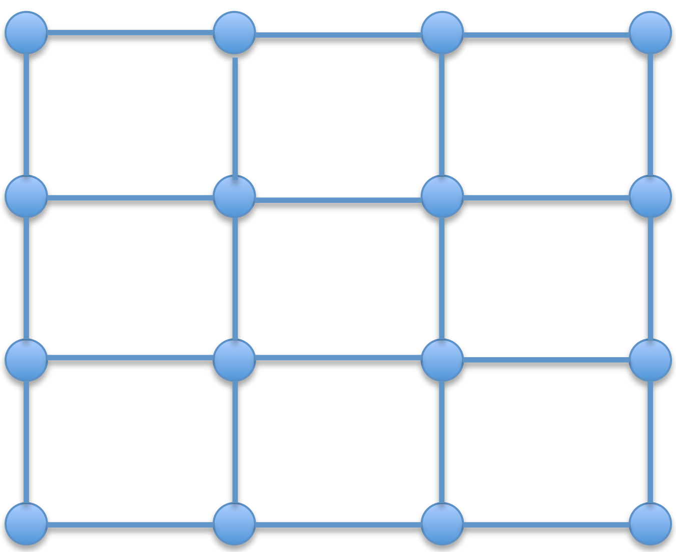
Minimizing the negative log-likelihood

Mathematically, the problem of finding the most probable x results in the following optimization problem

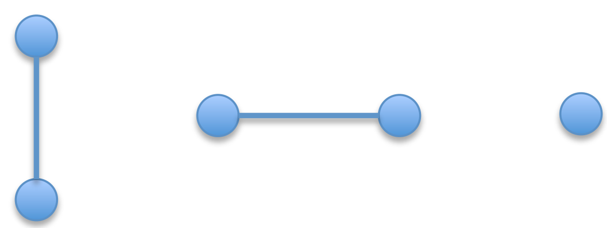
$$
\begin{aligned}
x^* &= \operatorname{argmax}_{x \in \Lambda^{|V|}} p_X(x) \\
&= \operatorname{argmax}_{x \in \Lambda^{|V|}} \frac{1}{Z} \exp(-E(x)) \\
&= \operatorname{argmax}_{x \in \Lambda^{|V|}} \exp(-E(x)) \\
&= \operatorname{argmax}_{x \in \Lambda^{|V|}} \log \exp(-E(x)) \\
&= \operatorname{argmin}_{x \in \Lambda^{|V|}} E(x).
\end{aligned}
$$

Clique Factorization in Image Segmentation

- In the following: study an MRF model for image segmentation
- 4-neighborhood leads to Gibbs potential with factorization over the three types of cliques



In the example on the left, there are three types of cliques:



MRF Model for Image Segmentation

- Markov random field X = (X_{ij}) with

$$X_{ij} = \begin{cases} 1 \text{ if } (i,j) \text{ is in the foreground} \\ 0 \text{ if } (i,j) \text{ is in the background} \end{cases}$$

- Next step: Determine potentials W_{ij} for singleton cliques and $W_{ij,kl}$ for pairwise cliques such that

$$E(x) = \sum_{ij} W_{ij}(x_{ij}) + \sum_{(i,j)\sim(k,l)} W_{ij,kl}(x_{ij}, x_{kl})$$

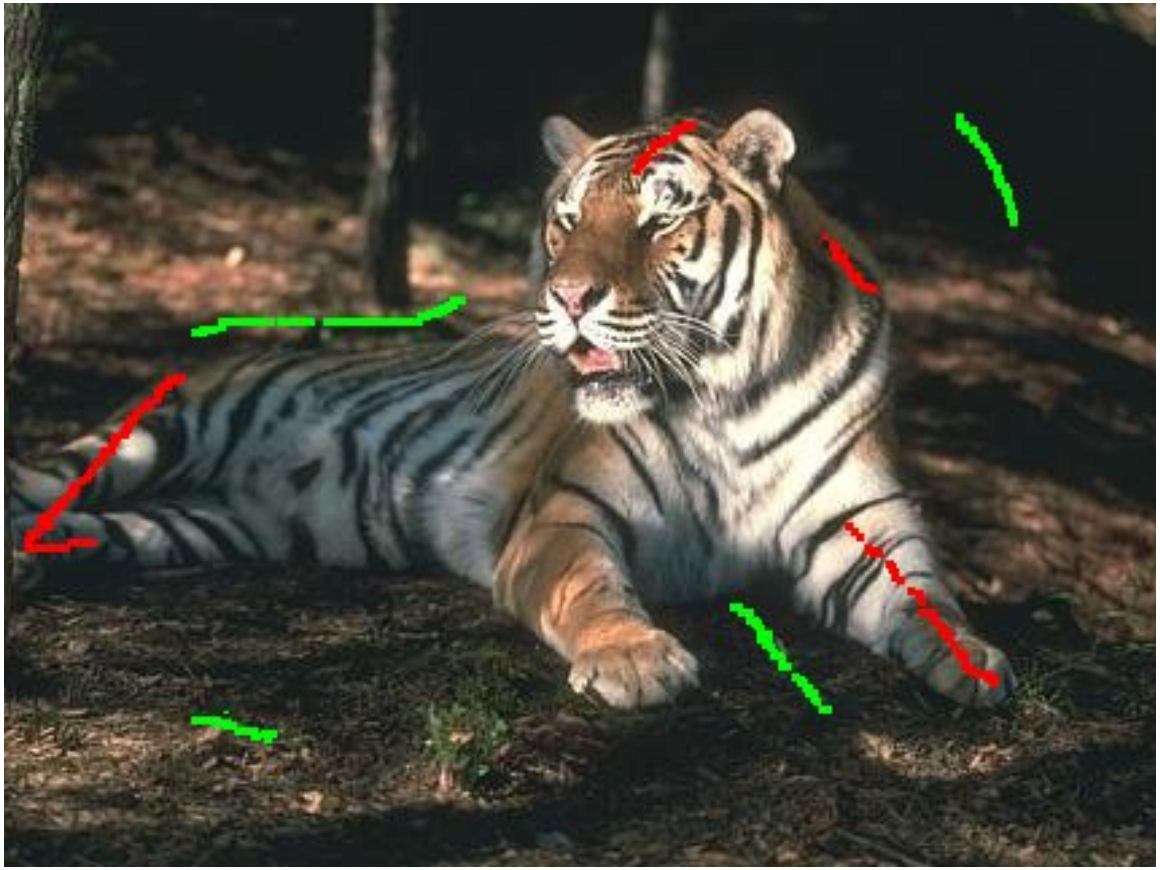
Unary Cliques Potential

- Estimate probability distributions p,q : $\Omega \rightarrow [0,1]$:

  p(i,j) = probability of prixel (i,j) belonging to the foreground

  q(i,j) = probability of pixel (i,j) belonging to the background.

- Such a distribution could be estimated for example on the base of user scribbles.



User Scribbles for Color Distribution

Based on user scribbles, a foreground-background probability distribution is estimated.

image courtesy of C. Nieuwenhuis

Data Term

- Once p and q are determined, choose $w_{ij}^{\mathrm{fg}}, w_{ij}^{\mathrm{bg}}$ proportional to –log(p(i,j)), -log(q(i,j)).
- Next, set

$$W_{ij}(x_{ij}) = w_{ij}^{\mathrm{fg}} x_{ij} + w_{ij}^{\mathrm{bg}} (1 - x_{ij})$$

- The unary potential is often referred to as data term

Pairwise Clique Potential: Boundary Prior

It is likely that the intensity values change at the object‘s boundary. Set

$$w_{ij,kl}^{\text{pairwise}} = \frac{\lambda}{(u(i,j) - u(k,l))^2}.$$

Then the pairwise potential

$$W_{ij,kl} = w_{ij,kl}^{\text{pairwise}} \delta(x_{ij}, x_{kl})$$

favours such object boundaries. Here,

$$\delta(x_{ij}, x_{kl}) = \begin{cases} 1 & \text{if } x_{ij} \neq x_{kl} \\ 0 & \text{else.} \end{cases}$$

Optimization Problem for Segmentation

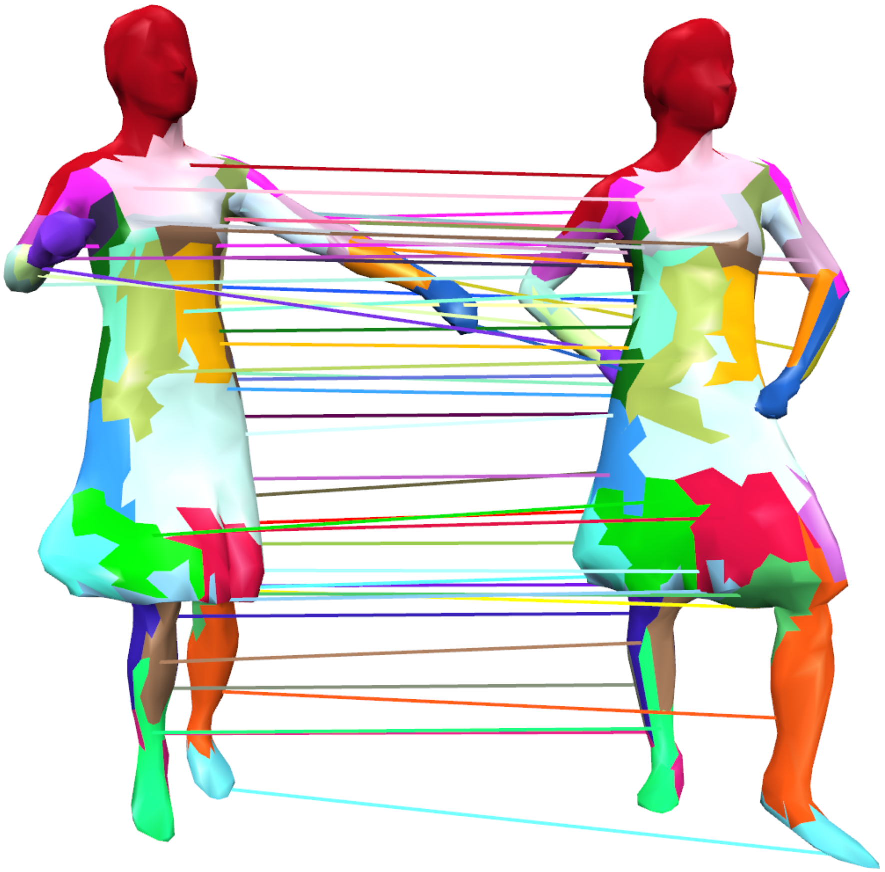
MAP inference for computing the most probable segmentation will then be an optimization problem

$$\min_{x\in\{0,1\}^{m\cdot n}} \underbrace{\sum_{i,j} w_{ij}^{\text{fg}} x_{ij} + w_{ij}^{\text{bg}}(1-x_{ij})}_{\text{unary cliques}} + \underbrace{\sum_{(i,j)\sim(k,l)} w_{ij,kl}^{\text{prior}} \delta(x_{ij}, x_{kl})}_{\text{pairwise cliques}}$$

where

$$\delta(x_{ij}, x_{kl}) = \begin{cases} 1 & \text{if } x_{ij} \neq x_{kl} \\ 0 & \text{else.} \end{cases}$$

Next lecture: Polynomial time solution via graph cuts!

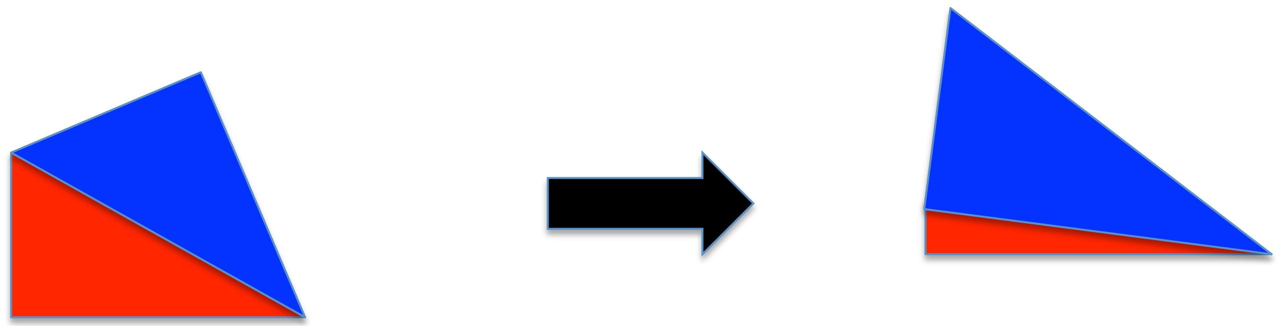


3D Shape Matching

The goal in 3D shape matching is to find meaningful correspondences of two 3D shapes (represented as triangle meshes)

Triangle Correspondences

- We search for optimal correspondences of triangles

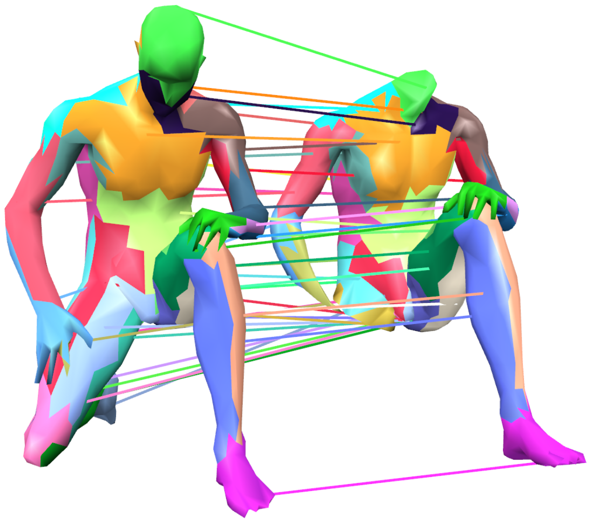


- Optimization over indicator vector $z \in \{0,1\}^N$
- linear cost vector c
- linear constraints granting that matchings are bijective and that neighboring triangles are matched to neighboring triangles

Integer Linear Program

Alltogether, this approach allows to formulate 3D shape matching as an integer linear program

$$\min_{z\in\{0,1\}^{|F|}} \quad c^t \cdot z$$
$$\text{subject to} \quad \begin{pmatrix} \partial \\ \pi_X \\ \pi_Y \end{pmatrix} \cdot z = \begin{pmatrix} \mathbf{0} \\ \mathbf{1} \\ \mathbf{1} \end{pmatrix}.$$



Plan for Today

1. Formalities

2. Examples of Combinatorial Optimization Problems in Computer Vision

3. Outline of the Lecture

Outline of the Lecture

- Combinatorial Optimization problems are roughly optimization problems over finite sets.
- We will see several instances of such problems arising in Computer Vision.
- The main focus lies on
  - identifying polynomial-time solvable problems (totally unimodular matrices)
  - learning efficient algorithms for solving the latter (graph cuts, linear programming)
  - learning efficient algorithms for approximately solving NP-hard vision problems (QPBO, dual decomposition)

Summary

- We have seen three examples of Combinatorial Optimization problems in ComputerVision: binary segmentation, feature matching, shape matching
- We have derived how maximum-a-posteriori (MAP) inference in Markov random fields (MRF) leads to a combinatorial minimization problem.