Document recto/verso separation with a dual layer Markov model

CORESA 2009

Christian Wolf

Université de Lyon, CNRS
INSA-Lyon, LIRIS, UMR5205, F-69622, France
http://liris.cnrs.fr/christian.wolf

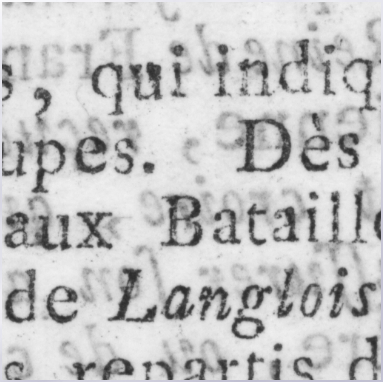
March 20^{th} 2009

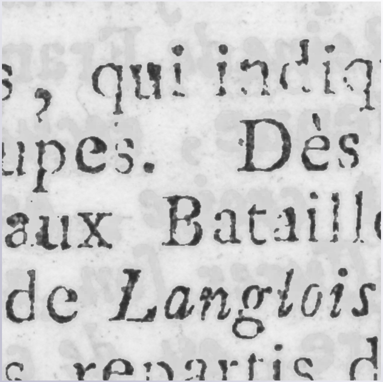
Plan

Plan

The problem

Blind ink bleed-through removal (scans of the verso side are not available).



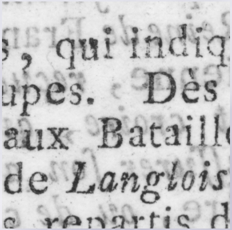


The verso contents must be replaced by background.

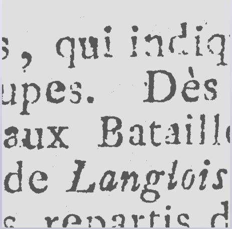
Ink bleed-through removal

Problem : image segmentation. The classification decision for each pixel is based on

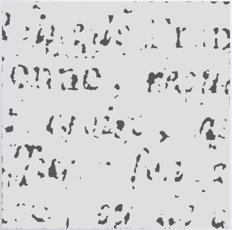
- the color/gray value of each pixel
- the local structure of the image.



input



recto labels

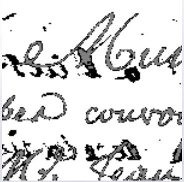


verso labels

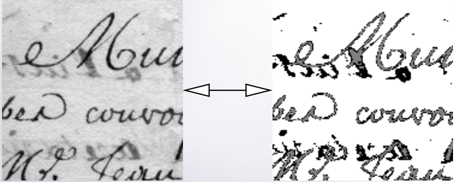
The framework : Bayesian estimation with two hidden Markov random fields (MRF)

Bayesian estimation

$P(f)$



$P(d|f)$



PRIOR KNOWLEDGE

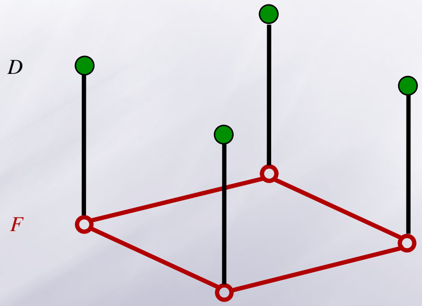
LIKELIHOOD OF THE DATA
GIVEN THE HIDDEN VARIABLES

ESTIMATION

$$\hat{f} = \arg\max_f p(f|d)$$

Plan

A classical hidden Markov random field



Observed variable

Hidden variable

Clique of the Prior model

Clique of the observ. model

$$P(f) = \frac{1}{Z} \exp\left\{-U(f)/T\right\}$$

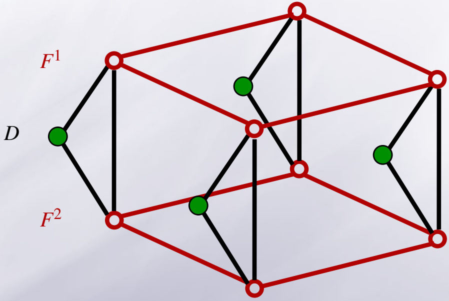
Why two hidden label fields ?

input

segmentation

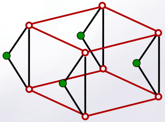
- The prior should regularize fields which directly correspond to the natural process "creating" the contents.
- A correct estimation of the covered verso pixels, through the spatial interactions encoded in the MRF, helps to correctly estimate verso pixels which are not covered by a recto pixel.

The double hidden Markov random field



Observed variable	Hidden variable	Clique of the Prior model	Clique of the observ. model

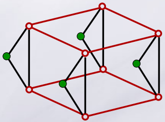
The joint probability : dependence of recto and verso



- w/o observations, f^1 and f^2 are independent
- with oberservations, they are dependent

$$
\begin{aligned}
& P(f^1, f^2, d) \\
= & \frac{1}{Z} \exp\left\{ - \left(U(f^1) + U(f^2) + U(f^1, f^2, d) \right) / T \right\} \\
= & \frac{1}{Z_1} \exp\left\{ -U(f^1, f^2) / T \right\} \cdot \\
& \frac{1}{Z_2} \exp\left\{ -U(f^1, f^2, d) / T \right\} \\
= & P(f^1, f^2) P(d | f^1, f^2) \\
= & P(f^1) P(f^2) P(d | f^1, f^2)
\end{aligned}
$$

The prior of a single field : the Potts model



$$U(f) = \sum_{\{s\}\in\mathcal{C}_1} \alpha f_s + \sum_{\{s,s'\}\in\mathcal{C}_2} \beta_{s,s'} \delta_{f_s,f_{s'}}$$

$$\delta_{i,j} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{else} \end{cases}$$

- Discontinuity preserving

Plan

The observation model (Likelihood)

Assumption

Two independent degradation processes for each side, followed by a combination of the two “virtual” observations.

$$D = \phi_c(\phi_1(F^1), \phi_2(F^2))$$

The combination process ϕ_c assumes 100% opaque ink : a recto pixel $f_s^1 = 1$ completely covers the verso pixel f_s^2, whatever its label.

A Gaussian observation model

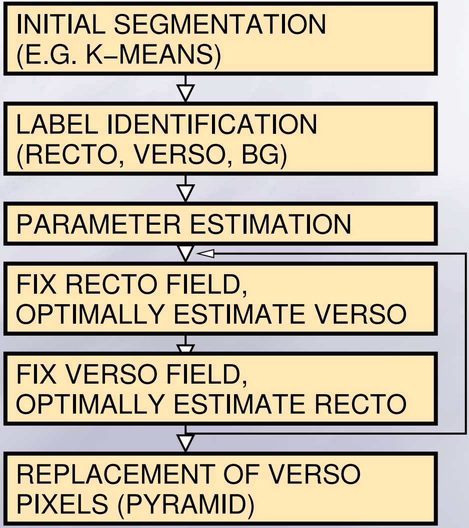
$$P(d|f^1, f^2) = \prod_s \mathcal{N}(d_s;\ \boldsymbol{\mu}_s, \boldsymbol{\Sigma}_s)$$

$$\boldsymbol{\mu}_s = \begin{cases} \boldsymbol{\mu}_r & \text{if } f_s^1 = \textit{text} \\ \boldsymbol{\mu}_v & \text{if } f_s^1 = \textit{background} \text{ and } f_s^2 = \textit{text} \\ \boldsymbol{\mu}_{bg} & \text{else} \end{cases}$$

$$\boldsymbol{\Sigma}_s = \begin{cases} \boldsymbol{\Sigma}_r & \text{if } f_s^1 = \textit{text} \\ \boldsymbol{\Sigma}_v & \text{if } f_s^1 = \textit{background} \text{ and } f_s^2 = \textit{text} \\ \boldsymbol{\Sigma}_{bg} & \text{else} \end{cases}$$

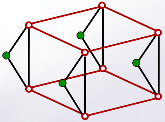
Plan

Outline of the whole method



Plan

The posterior probability and its maximization

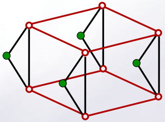


$P(d)$ does not depend on f^1, f^2 $\Longrightarrow$ We can maximize the joint probability instead of the posterior probability

$$
\begin{aligned}
P(f^1, f^2 | d) &= \frac{1}{P(d)} \quad P(f^1, f^2) P(d | f^1, f^2) \\
&\propto \quad P(f^1, f^2) P(d | f^1, f^2)
\end{aligned}
$$

$$
\begin{aligned}
U(f^1, f^2, d) & \\
= & \sum_{\{s\} \in \mathcal{C}_1} \alpha^1 f_s^1 + \sum_{\{s,s'\} \in \mathcal{C}_2} \beta^1_{s,s'} \delta_{f_s^1, f_{s'}^1} \\
+ & \sum_{\{s\} \in \mathcal{C}_1} \alpha^2 f_s^2 + \sum_{\{s,s'\} \in \mathcal{C}_2} \beta^2_{s,s'} \delta f_s^2, f_{s'}^2 \\
+ & \sum_{\{s\} \in \mathcal{C}_1} \frac{1}{2} (d_s - \boldsymbol{\mu}_s)^T \boldsymbol{\Sigma}_s^{-1} (d_s - \boldsymbol{\mu}_s)
\end{aligned}
$$

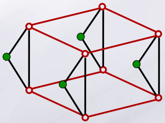
The posterior probability and its maximization



Rewrite of the energy in terms of unary functions U_1 and two types of binary functions U_2 and U_2'

$$
\begin{aligned}
& U(f^1, f^2, d) \\
= & \sum_{\{s\}\in\mathcal{C}_1} \left[\alpha^1 U_1(f_s^1) + \alpha^2 U_1(f_s^2)\right] \\
+ & \sum_{\{s,s'\}\in\mathcal{C}_2} \left[\beta^1_{s,s'} U_2(f_s^1, f_{s'}^1) + \beta^2_{s,s'} U_2(f_s^2, f_{s'}^2)\right] \\
+ & \sum_{\{s\}\in\mathcal{C}_1} U_2'(f_s^1, f_s^2; d_s)
\end{aligned}
$$

Minimization with minimum cut/maximum flow



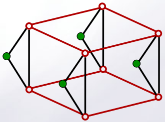
[Kolmogorov et al., PAMI 2004] :

E is *graph-representable* $\Longleftrightarrow$ each binary term $E(x, y)$ is submodular :

$$E(0,0) + E(1,1) \leq E(0,1) + E(1,0)$$

Depending on the observation d_s of site s this may or may not be the case for the observation model.

Minimization with minimum cut/maximum flow



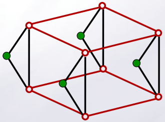
Iterative optimization algorithm :

Fix/estimate

- Initialize lables, e.g. with k-means
- repeat
  - Fix f^1, estimate optimal f^2
  - Fix f^2, estimate optimal f^1
- until happy/dead/running out of patience

- Sub problem can be solved exactly with minimum cut/maximum flow
- Equivalent to the α-expansion move algorithm for a single MRF with adapated interaction potentials.

Minimization with minimum cut/maximum flow

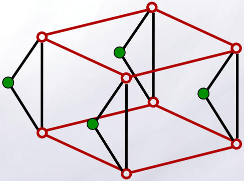


Joint estimation of a subset of the graph G

- Determine submodular sites :
  $H_s = 1 \iff U_2'(f_s^1, f_s^2, d_s)$ is submodular
- Initialize lables, e.g. with k-means
- repeat
  - Fix f_s^1 for all s and f_s^2 for $H_s = 0$, estimate optimal f^2 for all s and f_s^2 for $H_s = 1$
  - Fix f_s^2 for all s and f_s^1 for $H_s = 0$, estimate optimal f^1 for all s and f_s^1 for $H_s = 1$
- until happy/dead/running out of patience

$\Longrightarrow$ larger standard move !

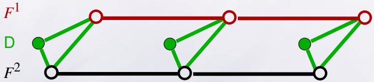
Minimization with minimum cut/maximum flow



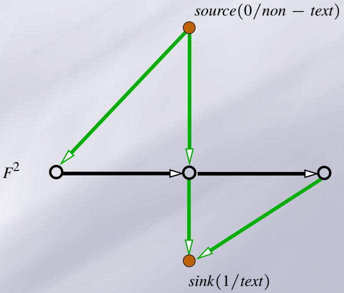
- Iterative algorithm : one field is fixed (e.g. f^1, the other field is estimated (e.g. f^2).
- Submodular sites ($H_s \longleftarrow 1$) are jointly estimated.

$$
\begin{aligned}
U(f^1,f^2,d) = & \\
= & \sum_{\{s\}\in\mathcal{C}_1:H_s=0} \alpha^1 U_1(f_s^1) \\
+ & \sum_{\{s\}\in\mathcal{C}_1:H_s=1} \alpha^1 U_1(f_s^1) \\
+ & \sum_{\{s\}\in\mathcal{C}_1} \alpha^2 U_1(f_s^2) \\
+ & \sum_{\{s,s'\}\in\mathcal{C}_2:H_s=0\wedge H_{s'}=0} \beta^1_{s,s'} U_2(f_s^1,f_{s'}^1) \\
+ & \sum_{\{s,s'\}\in\mathcal{C}_2:H_s=1\wedge H_{s'}=1} \beta^1_{s,s'} U_2(f_s^1,f_{s'}^1) \\
+ & \sum_{\{s,s'\}\in\mathcal{C}_2:H_s\neq H_{s'}} \beta^1_{s,s'} U_2(f_s^1,f_{s'}^1) \\
+ & \sum_{\{s,s'\}\in\mathcal{C}_2} \beta^2_{s,s'} U_2(f_s^2,f_{s'}^2) \\
+ & \sum_{\{s\}\in\mathcal{C}_1:H_s=0} U_2'(f_s^1,f_s^2;d_s) \\
+ & \sum_{\{s\}\in\mathcal{C}_1:H_s=1} U_2'(f_s^1,f_s^2;d_s)
\end{aligned}
$$

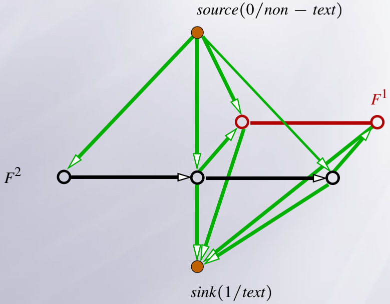
The cut graph



Dependency graph



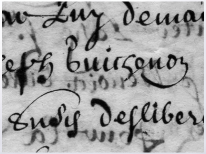
Cut graph α-exp move



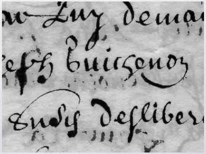
Cut graph proposed algorithm

Plan

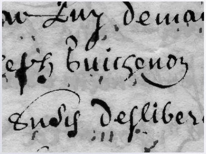
Examples



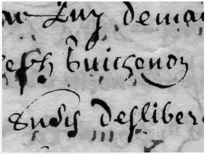
Input



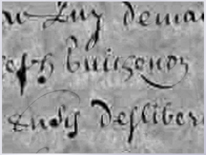
K-means



Single MRF

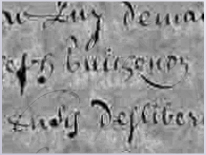


Double MRF

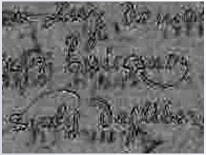


Tonazzini et al. (2007) src #1

Tonazzini et al. (2007) src #2



Tonazzini et al. (2004) src #1



Tonazzini et al. (2004) src #2

Experiments : synthetic & empirical

Synthetic images

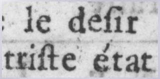
Noiselevel	σ=10	σ=15	σ=20
K-Means	0.25	1.40	3.56
Single-MRF	0.03	0.23	0.73
Double-MRF	**0.01**	**0.08**	**0.31**

Empirical evaluation (16 people $\times$ 4 test images)

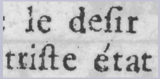
	complete			no single-MRF	
Method	1^{st}	2^{nd}	3^{rd}	1^{st}	2^{nd}
K-Means	18	10	36	21	43
Single-MRF	13	39	12	-	-
Double-MRF	**33**	15	16	**43**	21
Total	64	64	64	64	64

OCR improvement

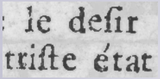
Method	Recall	Prec.	Cost	Dataset
No restoration	65.65	49.91	76,752	100
Niblack (segm. only)	-	-	-	-
Sauvola et al. (segm. only)	78.75	66.78	45,363	100
K-Means (k=3)	78.57	69.43	40,375	100
Tonazzini et al. meth. 1 - src #1	41.00	30.05	*74,819*	66
Tonazzini et al. meth. 1 - src #2	-	-	-	-
Tonazzini et al. meth. 2 - src #1	-	-	-	-
Tonazzini et al. meth. 2 - src #2	-	-	-	-
Tonazzini et al. meth. 2 - 3 src	50.52	33.90	*101,280*	89
Single MRF & α-exp. move	81.99	72.12	36,744	100
Double MRF (proposed method)	**83.23**	**74.85**	**32,537**	100


le deſir
triſte état

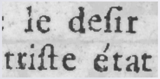
Ic déïïr triïïcfgitar
Input


le deſir
triſte état

le delîr t.tiPce état
k-means


le deſir
triſte état

le delîr ttiftc état
single MRF


le deſir
triſte état

le defir triûe état
double MRF

Conclusion and Outlook

- The double MRF model is more natural and more powerful than the single MRF model.
- The double MRF model using binary labels naturally leads to an efficient graph cut implementation.
- The MRF methods based on classification beat the source separation based methods in terms of OCR improvement.

Journal paper

More details in

C. Wolf, *Document Ink bleed-through removal with two hidden Markov random fields and a single observation field*, to appear in IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)